

"What Can A Robot Teach Us about Kantian Ethics?," in process.

### Abstract

In this paper, I examine a variety of agents that appear in Kantian ethics in order to determine which would be necessary to make a robot a genuine moral agent. However, building such an agent would require that we structure into a robot's behavioral repertoire the possibility for immoral behavior, for only then can the moral law, according to Kant, manifest itself as an *ought*, a prerequisite for being able to hold an agent morally accountable for its actions. Since building a moral robot requires the possibility of immoral behavior, I go on to argue that we cannot morally want robots to be genuine moral agents, but only beings that simulate moral behavior. But then, if that is what we want for robots, why should we want something different for human beings? Robot ethics, it seems, presents something of a *reductio* of Kant's ethics that points to hidden assumptions that hide in the very fabric of the Kantian moral enterprise, not the least of which is that Kant presumes humans to be *fallen* creatures. Religious doctrine, in other words, infects Kant's attempt to derive morality from reason. This paper will demonstrate that this is so.

Anthony F. Beavers, Ph.D.  
Professor of Philosophy / Director of Cognitive Science  
The University of Evansville  
<http://faculty.evansville.edu/tb2/>