

https://dspace.org/

Artificial intelligence, machines and understanding - a critique of John Searle’s Chinese Room thought experiment

Alexander, Gregory

2019
https://hdl.handle.net/10539/29525

Downloaded from a DSpace repository. https://dspace.org/

Abstract

Can machines think? In this paper, I will argue against John Searle's *Chinese Room Thought Experiment*. The *Chinese Room Thought Experiment* argues against the claim that "the appropriately programmed computer literally has cognitive states and that the programs thereby explain human cognition" (Searle 1980, 417). The central thesis that I intend to pursue is that it is/will be possible for machines to exhibit cognition. I will engage with various replies to Searle's argument, such as *The Systems Reply*, *The Virtual Mind Reply*, *The Robot Reply* and *The Brain Simulator Reply*, to argue for my thesis. I argue for an externalist account of "understanding" and I argue that if a machine were created that combines these responses then it would be able to understand the world around it. I then move on to argue for the theory of mind known as connectionism. I argue that machines should be programmed with a connectionist system as connectionist systems are able to learn from experience. In this paper, I intend to argue against Searle, put forward a concept of cognition and understanding, and argue for the theory of mind known as connectionism to conclude that a machine will be able to exhibit understanding of the world around it.