Nicholas Caudill

Dr. Jonathan Watson

ENG 216

30 October 2021

Exploring Perfect Reasoning Machines In Literature

In the Sherlock Holmes series by Sir Arthur Conan Doyle, Sherlock is referred to in the first adventure, “A Scandal In Bohemia,” as “the most perfect reasoning and observing machine that the world has seen.” This paper aims to explore other fictional tales and popular culture icons that include perfect reasoning machines in hopes of making connections that might support the hypothesis that Sherlock Holmes can be seen from the perspective that he is a sentient and artificially intelligent agent. If you fancy this perspective, then Sherlock Holmes becomes some dystopian RoboCop-like fantasy set in the 19th century, or a utopia in the sense that if more people are being caught after breaking laws then surely crime would decrease as a result.

For example, let us observe Samuel Butler's 1872 novel “Erewhon” or ‘Nowhere’. Chapter 13 of this novel talks about machine consciousness and self-replicating machines. This quote stuck out in particular, “Assume for the sake of argument that conscious beings have existed for some twenty million years: see what strides machines have made in the last thousand!” It is interesting Butler was thinking about the evolution of machines (his Wikipedia page mentioned he took a lot of inspiration from Charles Darwin), but what machines was he referencing specifically – the printing press or steam engine?

“But returning to the argument, I would repeat that I fear none of the existing machines; what I fear is the extraordinary rapidity with which they are becoming something very different

to what they are at present." This reminds me of archaeologist, Heinrich Schliemann, who contributed greatly to humanity by discovery of the lost city of Troy[0] through assuming "The Odyssey" by Homer actually happened and was not just a work of fiction. By following real-world hints in the Illiad, the lost city of Troy was discovered which means "The Odyssey" is not entirely fiction and humans might have really slaughter each other like animals around 1200 B.C.E.

What if we fancied the idea that Sherlock Holmes was a real artificially intelligent agent that really lived and was invented by some ancient civilization? Was it not this same process that lead to the discovery of the lost city of Troy? There might also be value in doing the inverse of what Schliemann did, where instead of assuming something in a seemingly fictional tale actually happened in the past, we instead assume something in a fictional story will happen soon. Instead of discovering lost history one could be discovering new potentials through fictional works such as Sherlock Holmes. What better way to discuss the moral and ethics in engineering artificially intelligent systems than by reflecting on fictional literature such as Sherlock Holmes. I could imagine two designers discussing a design for a system where they say "By studying Sherlock, we noticed if this Matrix operation is tweaked slightly this direction, then we produce a Sherlock-like agent that is more friendly towards women which is one of Sherlock's flaws communicated throughout the stories.

Researchers at the University of Cambridge identified one major theme in utopian scenarios featuring AI is immortality, or indefinite lifespans[1]. When writing Sherlock Holmes, Doyle may have sought the theme of immortality or 'improving current mortality' by creating the character Sherlock Holmes who increases the likelihood of people living longer by catching violent criminals that do harm to society. Another interesting detail about artificial intelligence

that ties in well with fictional literature such as Sherlock Holmes, Dracula, or The Strange Case of Dr. Jekyll and Mr. Hyde is that if someone built a robot that looks indistinguishable from a human and say one worked at a movie theater with this robot for many years thinking the robot was a real human – this would be similar to the plots of these stories. We have Dracula, who is this very smart alien-monster that blends in with normal people so well- or Dr. Jekyll maintaining his social status by day and then rampaging through London at night. If Sherlock Holmes was to be realized in real life tomorrow through some sophisticated engineering feat recently discovered – what problems could Sherlock solve if we asked him to specialize in areas other than criminal science? How would we prevent Sherlock from hacking our bank accounts and buying tons of opium/tobacco with our money? These are the questions we should really be asking since we are the first generation to really see systems like this on the road, literally, with fully autonomous self-driving vehicles.

Works Cited

[0] (2018): <https://www.nationalgeographic.co.uk/history-and-civilisation/2018/11/how-archaeologists-found-lost-city-troy>.

[1]Cave, Stephen; Dihal, Kanta. “"Hopes and fears for intelligent machines in fiction and reality.” *Nature Machine Intelligence*. doi:10.1038/s42256-019-0020-9

[2] Doyle, Sir A. Sherlock Holmes and a Scandal in Bohemia: Case 1. Graphic Universe ™, 2018.A Scandal in Bohemia. 2009.

[3] Butler, S. (2012). Erewhon. Courier Corporation.