Abstract

It is the goal of this paper to provide a suitable argumentative framework to defend a claim to varying levels of personal rights for artificial agents of varying capabilities. In reviewing relevant debates concerning the requisites of personhood through history and its application to non-persons and non-human entities I hope to identify some common driving forces present in those arguments and develop a unified theory of a spectrum of personhood. In the first section of this paper, I will introduce the artificial agent, and the urgency of this discussion, followed by a brief history of the act of extending quasi or full personal rights to agents, human and non-human. The second section will explore a materialist argument on personhood followed by a review of the functional relationships considered in persons. The third section will provide possible avenues for refutation and my response to such arguments and in my final section, I will summarize what we have accomplished in the paper and recapitulate how the paper contributes to the dialogue of personhood in the 21st century and to our society at large.

Outline:

Section 1: Introduction

Intro

1. The urgency of this discussion.

Chapter I

1. A brief history of legal persons and the act of extending legal rights.

Section 2: Two possible avenues of approach

Chapter II

1. A review of a beaten path, so to speak, within the forum of this topic. This chapter will discuss the prospects of an argument based on a materialist perspective of the person. Here we will explore what has been argued before but also try reconcile this argument with the earlier stated model of a spectrum of personhood needing a spectrum of legal responses.

Chapter III

1. This chapter will explore a theoretical, functional relationship necessary in considering something a person. It will also review historical precursors to any extension of rights involving any agents not previously considered legitimate persons, with the larger goal of integrating commonalities into an argument for a varied approach to extending rights to non-persons.

Section 3: A Refutation

Chapter IV

1. This chapter will investigate possible avenues of discourse against the earlier propositions made in the paper.

Section 4: Conclusion

Chapter V

1. This chapter will summarize my findings and restate the urgency of this debate.

Introduction

The date was July 9th, 2015. Selmer Bringsjord, chair of New York’s Rensselaer Polytechnic Institute’s cognitive science department, had assembled three small *NAO* robots in her robotics lab, each similarly programed and each tasked with solving a puzzle which at its core required an element of self-awareness to solve. The ubiquitous wise men puzzle was being simulated among these robots, each one given the information that they had received a one of two kinds of pills, a “dumbing” pill, which would prevent them from speaking, or a placebo, which would have no effect (techxplore). In reality, the effects of these “pills” would be triggered by pressing a button atop the robot’s heads which when triggered would prompt the beginning of the puzzle--by having the robots believe they were given the one of the two pills. From here, the robots are asked verbally to deduce which pill they received. The simulation begins quietly as each robot attempts to respond to the question. With the other two robots unable to respond the “placebo” robot gets up and reports that it does not know what pill it received, but then, after hearing its own voice it retracts its earlier response and states: “sorry, I know now. I was able to prove that I was not given the dumbing pill.” Not many AI have been capable of passing this test, which relies not only on the agent’s ability to think critically and inductively but also relies on the agent’s ability to recognize itself distinctly from its environment. On the subject of this project’s significance, Bringsjord herself states that “we…[are] talking about a logical and a mathematical correlate to self-consciousness” (techxplore).

Bringsjord is not alone in her interests. The tech world is experiencing a boom in AI-related projects and research--across the private and public sectors. In 2016, some of the US’s largest technology firms including Google, Facebook, Microsoft and Baidu spent some $8.5 billion on deals aimed at the brightest talent across the AI research space, by one *Economist* report (economist). The same report shows that specialists in machine learning are most in demand. The growth of this industry is moving so quickly that its professionals are already making efforts to draw the reigns on their work and control the direction of research and development to ensure that products of AI benefit humankind. To this cause, many of the earlier mentioned tech firms have also come together to devise “real ethics” for the development of artificial intelligence (New York Times). In these meetings, concern for how much change AI can bring to the world, good or bad, is recognized practically as clearly as the tech industry’s newfound and immense interest in the field itself. While the conglomerate has yet to establish a name, and has yet to decide on how it function or what it will do they understand their broad goal—to protect humankind from the dangers aspects of their research, to provide some arbitration as to what is a bad idea and what is a good idea insofar as development of AI.

It is in the same spirit that I am writing this thesis. It is not so easy to predict what trajectory AI will take in the next fifty or one-hundred years. What we do understand is that the artificial agents we know from science fiction are no longer just the stuff of fantasy novels and movie depictions. They are today, the objects of intense research and the ultimate ambition of an ever-increasing amount of public and private institutions. Their work is powered by a philosophical, technical, and intellectual human hunger to accomplish the impossible, but perhaps more concerning, it is driven by the relentless force of the market, which, as many have already noted, has its eyes on immediate, and long term future of artificial intelligence. The progression of this technology has moved not only myself but also others interested in the capabilities of AI and its intersection with philosophy, to reenter forums discussing what it means to be human in the first place; forums like those discussing what it means to have an identity and in this case, what it means to be a person. Such discussions I feel will have a meaningful and well-needed impact on the direction developers chose to move in going forward in this field.

Chapter I

A history of extending legal rights

I find myself in agreement with those theorists who acknowledge man’s developing social capacities. Charles Darwin observed these “social instincts,” which he believed were responsible for the pleasure some animals feel in fulfilling their social capacities. *Decent of Man* discusses a theoretical evolution of man’s social nature. It correlates man’s development of “intellectual power” to an ability to track the more distant consequences of his actions. When man became capable of experiencing impulses like sympathy or empathy, they were largely limited by the range of this intellectual power. As this ability grew in scope, man’s concerns bloomed outward onto the welfare of not just his cohorts but also to other less useful members of his society (e.g. “the imbecile [and] the maimed”). His ability to sympathize grew widely and abstractly, as today we find ourselves able to exert sympathy towards animals, and even plant life (Decent of Man, CH 3). Christopher Stone, in his work *Should Trees Have Standing?—Toward Legal Rights for Natural Objects,* makes an interesting observation in reviewing Darwin’s work. He shows us that the history of the law seems to corroborate Darwin’s account of a dynamic and progressive human social-sympathetic behavior.

Within the immediate social groups of early men, we have a record of changing social relationships among our familial actors as well as familial outsiders. In the past many persons who today we recognize as worthy of individual rights, relatively speaking, had none. Women and children make up much of this historical record. Early accounts of infanticide give a clearer picture of this. Not much is known about the public sentiments surrounding the earliest cases of infanticide, but the archeological evidence does suggest that it was a “commonplace” practice among ancient cultures (INFANTICIDE). The bulk of this evidence, however, hails from the records of ancient Greece and Rome, and within these societies, it seems that the practice was legitimized in a variety of ways “ranging from population control to eugenics to illegitimacy.” Whether given legitimate reason or not it is clear that the practice of infanticide puts the desire to fulfill some mandate ahead of any conception of rights for the child. The children of antiquity were regarded more like extensions of a man’s property. Paternal figures in ancient Rome had far-reaching control of their families, and could choose to give their daughters away in marriage, give their son a wife, divorce their child, or even sell their child (Stone [Maine]).

Of course, the relationship we share with young children has gone through much change since the days of ancient Rome. Our legal system has made dedicated legal provisions to the cause of protecting young persons. Consider the formation of the Department of Health and Human Services, also known as the “Children’s Bureau” in 1912 (Cornell #1). Also note Title 42, U.S. Code Chapter 67, detailing the laws surrounding child abuse, prevention, and the reform of adoption procedure (Cornell #2). Stone notes that each “successive extension of rights to some new entity has been, theretofore, a bit unthinkable.” Despite pushback from what some would consider “unthinkable” history has shown that under certain circumstances we are capable of evolving on what and whom we view as worthy of protection, rights and indeed, who is worthy of being called a person.

We find a powerful example of this within the animal rights movement, which has seen much development in recent history. As early as 1641, colonial America began enacting legal code preventing “tirrany or crueltie” towards animals kept for “man’s use” (Animal Protection). Prevention of cruelty to animals was of particular importance to the humanist activists of the 19th Century England, who campaigned for protection to animals from cruelty as well as for a number of human rights proposals of the time including anti-slavery and woman’s suffrage (movement animal). Our early human predecessors likely could not have conceived of a future where whole government institutions are dedicated to the protection of animals and yet our own U.S. government has several in place to do just this. The U.S. Fish & Wildlife Service, the National Park Service, and the Environmental Protection Agency are just some out of large list of agencies burdened with some duty to the well-being of animals and the environment. These progressively considerate legal provisions through time are but one testament to the growing liberality of human sympathy Darwin and Stone refer to.

Work cited (Introduction):

<https://www.nytimes.com/2016/09/02/technology/artificial-intelligence-ethics.html?_r=0> (New York Times)

<http://www.economist.com/news/business/21695908-silicon-valley-fights-talent-universities-struggle-hold-their> (economist)

(techxplore) <https://techxplore.com/news/2015-07-robots-wise-men-puzzle-degree-self-awareness.html>

<https://www.youtube.com/watch?v=95aYLmG0wt8>

<https://www.slashgear.com/nao-robot-becomes-self-aware-very-briefly-20393569/>

<http://www.sciencealert.com/a-robot-has-just-passed-a-classic-self-awareness-test-for-the-first-time>

<https://www.newscientist.com/article/mg22730302-700-robot-homes-in-on-consciousness-by-passing-self-awareness-test/?gwaloggedin=true>

Works cited (Chapter I):

<http://www.sonoma.edu/users/w/wallsd/animal-rights-movement.shtml> (movement animal)

<http://tah.oah.org/november-2015/the-history-of-animal-protection-in-the-united-states/> (Animal Protection)

<http://darwin-online.org.uk/content/frameset?itemID=F937.1&viewtype=text&pageseq=1> (Darwin)

<https://books.google.com/books?hl=en&lr=&id=i4XqOuxjzvIC&oi=fnd&pg=PA3&dq=history+of+infanticide&ots=AHG5nOKEfv&sig=aqDpy0VEb4V18hO7y9cuvgFTQ40#v=onepage&q=history%20of%20infanticide&f=false> (INFANTICIDE)

<https://www.law.cornell.edu/uscode/text/42/191> (Cornell #1)

<https://www.law.cornell.edu/uscode/text/42/chapter-67> (Cornell #2)