Can Machines Have Rights?

David J. Gunkel
Northern Illinois University
dgunkel@niu.edu
Work in the new field of machine morality, machine ethics, and roboethics generally focus attention on the decision making capabilities and actions of machines and the consequences of this behavior for human beings and human social institutions. Absent from much of the current literature is a consideration of the other side of this issue—that is, the question of machine rights.
Can Machines Have Rights?

Today I want to take up and investigate whether it is possible for a machine (defined broadly and including artifacts like software bots, algorithms, embodied robots, etc.) to have or be ascribed anything like rights, understood as the entitlements or interests of a moral subject that need to be respected and taken into account.
Because there are a number of different and competing methods by which to examine and decide the question the rights, I will not supply one definitive answer, but will consider three related moral perspectives that, taken together, add up to an affirmative response to the question concerning machine rights.
From a traditional philosophical perspective, the question “can machines have rights?” not only would be answered in the negative but the query itself risks incoherence.
This is because traditional forms of ethics, no matter how they have been articulated have been anthropocentric. Under this human-centered conceptualization, technology, no matter how sophisticated its design or operation…
Instrumentalism

“The instrumentalist theory offers the most widely accepted view of technology. It is based on the common sense idea that technologies are ‘tools’ standing ready to serve the purposes of users.” – Andrew Feenberg 1991

….is considered to be nothing more than a tool or instrument of human endeavor.

From this perspective, the bar for machine rights appears to be impossibly high.

In order for a machine to have rights, it would need to be recognized as human or at least be virtually indistinguishable from another human being.
Although this is often the subject of science fiction, it is not limited to fictional speculation,
and researchers like Hans Moravec, Rodney Brooks, and Raymond Kurzweil predict human-level or better machine capabilities by the middle of the century.

Although achievement of this remains hypothetical, the problem is not whether machines will or will not attain human-like capabilities.
The problem rests with the anthropocentric criteria itself, which not only marginalizes machines but has often been mobilized to exclude others—including women, people of color, and animals.
Because of this, moral philosophy has been critical of its own procedures and has sought to articulate moral status in ways that are not dependent on spurious or prejudicial criteria.
Recent innovations disengage moral standing from identification with the human being and refer it to the generic concept of "person."

As promising as this innovation is, however, there is little agreement concerning what makes someone or something a person, and the literature on this subject is littered with different formulations and often incompatible criteria.
In an effort to contend with, if not resolve these problems, researchers often focus on the one "person making" quality that appears on most, if not all, the lists—consciousness.

In fact, consciousness is widely considered to be a necessary if not sufficient condition for moral standing,
and there has been considerable effort in the fields of philosophy, AI, and robotics to address the question of machine morality by targeting the possibility (or impossibility) of machine consciousness.
Determining Factors

1. Design and performance of actual systems and artifacts
2. How we understand and operationalize “consciousness”

This determination is dependent not only on the design and performance of actual artifacts but also—and perhaps more so—on how we understand and operationalize “consciousness.”

Unfortunately there is little or no agreement concerning this matter, and the concept encounters both terminological and epistemological problems.
Terminological Problems
“The term means many different things to many different people, and no universally agreed core meaning exists. ” – Max Velmans 2005

First, we do not have any widely accepted definition of “consciousness,” and the term means many different things to many different people. In fact, if there is any agreement among philosophers, psychologists, cognitive scientists, neurobiologists, AI researchers, and robotics engineers regarding this matter, it is that there is little or no agreement, when it comes to defining and characterizing the concept.
Epistemological Problems
Because consciousness is a property attributed to "other minds," its presence or lack thereof requires access to something that is and remains inaccessible.

Second, even if it were possible to define consciousness, we still lack any credible and certain way to determine its actual presence in another.

Because consciousness is a property attributed to "other minds," its presence or lack thereof requires access to something that is and remains inaccessible.

Although philosophers, psychologists, and neuroscientists throw considerable argumentative and experimental effort at this problem, it is not able to be fully resolved.
Epistemological Problems
“There is no proving that something that seems to have an inner life does in fact have one” – Daniel Dennett 1998

Consequently, not only are we unable to demonstrate with any certitude whether animals, machines, or other entities are in fact conscious (or not) and therefore legitimate moral persons (or not), we are left doubting whether we can even say the same for other human beings.

And it is this persistent doubt that opens the possibility for extending rights to other entities like machines or animals.
Which brings us to the second issue…animal rights.
Animals have not traditionally been considered moral subjects, and it is only recently that the discipline of philosophy has begun to approach the animal as a legitimate subject of moral concern.
The crucial turning point in this matter is derived from a brief statement provided by Jeremy Bentham: "The question is not, Can they reason? nor Can they talk? but, Can they suffer?"

Following this insight, the crucial issue for animal rights philosophy is not to determine whether some entity, like an animal, can achieve human-level capacities with things like speech, reason, or cognition; “the first and decisive question would be rather to know whether animals can suffer”

This change in perspective—from an agent oriented to a patient oriented ethics—provides a potent model for entertaining the question of the rights of machines.
bête-machine

“If any such machine had the organs and outward shape of a monkey or of some other animal that lacks reason, we should have no means of knowing that they did not possess entirely the same nature as these animals” – René Descartes 1637

This is because the animal and the machine, beginning with René Descartes, share a common ontological status and position—marked, quite literally, by the hybrid term bête-machine. Despite this similitude, animal rights philosophers have resisted efforts to extend rights to machines, and they demonize Descartes for even suggesting the association. This exclusivity has been asserted and justified on the grounds that the machine, unlike an animal, is not capable of experiencing either pleasure or pain.
“A stone does not have interests because it cannot suffer. Nothing that we can do to it could possibly make any difference to its welfare. A mouse, on the other hand, does have an interest in not being kicked along the road, because it will suffer if it is.”— Peter Singer 1975

Like a stone or other inanimate object, the machine would have nothing that mattered to it and therefore, unlike a mouse or other sentient creature would not be a legitimate subject of moral concern. Although this argument sounds rather reasonable and intuitive, it fails for at least three reasons.
First, it has been practically disputed by the construction of various mechanisms that now appear to exhibit emotional responses or at least provide external evidence of behaviors that looks undeniably like pleasure or pain.
2. Epistemological Issues

“We cannot directly experience anyone else’s pain, whether that ‘anyone’ is our best friend or a stray dog. Pain is a state of consciousness, a ‘mental event,’ and as such it can never be observed.” – Peter Singer 1975

Second, it can be contested on epistemologically grounds.

Because suffering is typically understood to be subjective, there is no way to know exactly how another entity experiences unpleasant sensations.

Like “consciousness,” suffering is also an internal state of mind and is therefore complicated by the problem of other minds.
Furthermore, and to make matters even more complex, we may not even know what “pain” and “the experience of pain” is in the first place.

This point is something that is taken up and demonstrated in Daniel Dennett's, "Why You Can't Make a Computer That Feels Pain."

In this provocatively titled essay, published decades before the debut of even a rudimentary working prototype, Dennett imagines trying to disprove the standard argument for human (and animal) exceptionalism “by actually writing a pain program, or designing a pain-feeling robot.”
2. Epistemological Issues

“There can be no true theory of pain, and so no computer or robot could instantiate the true theory of pain, which it would have to do to feel real pain” – Daniel Dennett 1998

At the end of what turns out to be a rather protracted and detailed consideration of the problem, he concludes that we cannot, in fact, make a computer that feels pain.

But the reason for drawing this conclusion does not derive from what one might expect.

The reason you cannot make a computer that feels pain, Dennett argues, is not the result of some technological limitation with the mechanism or its programming.

It is a product of the fact that we remain unable to decide what pain is in the first place.
3. Moral Problems

“If (ro)bots might one day be capable of experiencing pain and other affective states a question that arises is whether it will be moral to build such systems—not because of how they might harm humans, but because of the pain these artificial systems will themselves experience. In other words, can the building of a (ro)bot with a somatic architecture capable of feeling intense pain be morally justified?” – Wallach & Allen, 2009

Third, all this talk about the possibility of engineering pain or suffering in order to demonstrate machine rights entails its own particular moral dilemma…expressed rather succinctly by Wallach and Allen…
Consequently, moral philosophers and engineers find themselves in a curious and not entirely comfortable situation.

If it were in fact possible to construct a device that “feels pain” in order to demonstrate the possibility of machine moral rights, then doing so might be ethically suspect insofar as in constructing such a mechanism we do not do everything in our power to minimize its suffering.

Or to put it another way, positive demonstration of machine rights might only be possible by risking the violation of those rights.
Third, when evaluating these different methods for determining the rights of others, the task might not be deciding which approach is better or which moral theory is more or less exclusive. The problem may be with the very procedure itself.
Irrespective of how it is defined, standard approaches to moral reasoning take what Mark Coeckelbergh calls a “properties approach.”

That is, they first define criteria for inclusion and then ask whether a particular entity meets this criteria or not. Proceeding in this fashion has at least two difficulties. First, one needs to be able to identity the property or properties that will be considered necessary and sufficient for deciding moral status—and there is considerable debate about and indecision concerning this very matter.
Second, this decision is necessarily a normative operation and an exercise of power over others. In making a determination about the criteria for moral inclusion, someone or some group normalizes their particular experience or situation and imposes this decision on others as the universal condition for moral considerability.
“The institution of any practice of any criterion of moral considerablity," Thomas Birch argues, "is an act of power” over others.
In response to these problems, philosophers have advanced alternative approaches that can be called “social-relational ethics”

These efforts do not endeavor to establish ontological criteria for inclusion or exclusion but begin from the existential fact that we always and already find ourselves in situations facing and needing to respond to others—not just other human beings but animals, the environment, organizations, and machines.

In fact, recent debates concerning the moral status of corporations turn on the question whether rights derive from intrinsic properties or are, as Anne Foerst, describes it, a socially constructed and conferred honorarium.
Thinking Otherwise

Ontological Question

Can machines have rights?

Moral status is decided on the basis of what something is or is not.

What is important here, is that these alternatives shift the focus of the question and change the terms of the debate.

Here it is not longer a question of “can machines have rights?” which is an ontological question.
Thinking Otherwise

Can machines have rights?
Moral status is decided on the basis of what something is or is not.

Should machines have rights?
Moral status is decided by how we relate and respond to others.

Instead it is “should machines have rights?” which is a moral query and one that is decided not on the basis of what things are but on how we decide to relate and respond to them.
These efforts put ethics before ontology, or as Emmanuel Levinas describes it “Morality is not a branch of philosophy, it is first philosophy.”

And proceeding in this fashion opens all kinds of new opportunities for considering the rights of machines…even if Levinas, for himself, would have resisted such a move.
Conclusion

In the end, what we have is not one reason for extending rights to machines but what we might call a "cumulative argument" for seriously considering the question of machine rights.
My objective, in this, has not been to resolve the question “can machines have rights?” once and for all but to position this inquiry in such a way that we make it the subject of investigation.

This is not, I should point out, a dodge or a cop out...it is the one thing that philosophers and philosophy are good for.
“I am a philosopher,” Dennett writes at the beginning of one of his books, “not a scientist, and we philosophers are better at questions than answers. I haven't begun by insulting myself and my discipline, in spite of first appearances. Finding better questions to ask, and breaking old habits and traditions of asking, is a very difficult part of the grand human project of understanding ourselves and our world” – Daniel Dennett 1996

"I am a philosopher," Dennett writes at the beginning of one of his books, “not a scientist, and we philosophers are better at questions than answers.

I haven't begun by insulting myself and my discipline, in spite of first appearances.

Finding better questions to ask, and breaking old habits and traditions of asking, is a very difficult part of the grand human project of understanding ourselves and our world"
And it is for this reason—in an effort to formulate better and more precise questions regarding the social position and status of the machine—that my own research in this area has been situated under the title “the machine question.”