ACE Blog

Why Digital Sentience is Relevant to Animal Activists

posted on February 3, 2015 by Brian Tomasik (http://www.animalcharityevaluators.org/blog/author/briantomasik/) | Leave your thoughts (http://www.animalcharityevaluators.org/blog/why-digital-sentience-is-relevant-to-animal-activists/#respond)

Guest blogger and ACE board member Brian Tomasik (/about/meet-our-team/board-of-directors/#brian) joins us this week to discuss digital sentience. Currently founding a research institute in Switzerland, Brian has previously worked as a software development engineer at Microsoft performing large-scale data analysis. He graduated from Swarthmore College with concentrations in computer science and statistics, and has written a number of well-known articles (http://reducing-suffering.org/) about effective ways to reduce animal suffering.

Summary

When aiming to reduce animal suffering, we often focus on the short-term, tangible impacts of our work, but longer-term spillover effects on the far future are also very relevant in expectation. As machine intelligence becomes increasingly dominant in coming decades and centuries, digital forms of non-human sentience may become increasingly numerous, perhaps so numerous that they outweigh all biological animals by many orders of magnitude. Animal activists should thus consider how their work can best help push society in directions to make it more likely that our descendants will take humane measures to reduce digital suffering. Far-future speculations should be combined with short-run measurements when assessing an animal charity's overall impact.

Introduction

As animal advocates, it's important for us to consider both the short-term impact of our work on animal suffering that exists today as well as the longer-term side effects of our advocacy. The far future is hard to predict and even harder to reliably influence, but the amount of animal suffering in the future may astronomically exceed that in the present, so even a tiny chance of a small impact on humanity's future trajectory (http://lesswrong.com/lw/hjb/a_proposed_adjustment_to_the_astronomical_waste/) could have immense consequences for non-humans of the future.

The future of animal suffering

What will the future of animal suffering look like? Obviously it's hard to say for sure, but some trends are suggestive.

- In the nearer term, we can expect increased animal consumption as China and other developing nations grow wealthier.
- On the timescale of decades, we might see a greater shift toward meat substitutes, possibly including *in vitro* meat, for reasons of health, environment, animal welfare, and perhaps cost.
- Society may shift away from cows and pigs as the impacts of climate change become increasingly severe,

---

Search ACE

[Search]

Sign up for our newsletter

Email Address*

[   ]

[Subscribe]

ACE Blog

Latest posts | Popular posts

(http://www.animalcharityevaluators.org/blog/researcher-referral/) *Researcher Referral (http://www.animalcharityevaluators.org/blog/researcher-referral/)*

(http://www.animalcharityevaluators.org/blog/why-digital-sentience-is-relevant-to-animal-activists/) *Why Digital Sentience is Relevant to Animal Activists (http://www.animalcharityevaluato*

although this might actually increase total animal suffering if consumers substitute toward poultry, fish, and perhaps insects -- which need to be farmed in greater numbers to produce the same yield.

- With luck, people will show increasing concern for the suffering of wild animals (http://foundational-research.org/publications/importance-of-wild-animal-suffering/) due to natural causes like predation and disease, although this impulse may be counteracted by increasing reverence for pristine wilderness and ecosystem integrity. Tendencies for both animal welfare and environmentalism increase with education and GDP per capita, so it's not obvious which meme will win out. At the moment, environmentalism is more prevalent.

On longer timescales, humans may begin populating other planets, perhaps starting with Mars. They may attempt "terraforming" -- making other planets more Earth-like in their conditions so that they can support plant and then animal life. While this would take thousands to hundreds of thousands of years (https://web.archive.org/web/20130410153012/http://www.popsci.com/military-aviation-space/article/2007-01/hijacking-red-planet) if it's feasible at all, the end result would be a significant increase in the amount of wild-animal suffering (http://www.animalcharityevaluators.org/research/foundational-research/wild-animal-suffering/). Terraforming many planets would multiply animal suffering manyfold over what we see today.

However, terraforming scenarios assume that biological life will remain in control of humanity's future. The apparent advent of greater machine intelligence casts this assumption into doubt.

Advantages of digital minds

Even today we see machines replacing humans in the field of space exploration. Robots are hard to build, but they can go places like Mars where it would be more expensive and more risky to send humans. Computers need power, but this is easier to generate (https://en.wikipedia.org/wiki/Radioisotope_thermoelectric_generator#Space) in electrical form than by creating a supply of human-digestible foods that contain a variety of nutrients. Machines are easier to shield from radiation (https://en.wikipedia.org/wiki/Radiation_hardening), don't need exercise to prevent muscle atrophy, and can generally be made more hardy than biological astronauts.

But in the long run, it won't be just in space where machines will have the advantage. Biological neurons transmit signals at 1 to 120 meters per second (http://people.eku.edu/ritchisong/301notes2.htm), whereas electronic signals travel at 300 million meters per second (the speed of light). Neurons can fire at most 200 times per second, compared with (https://en.wikipedia.org/wiki/Superintelligence#Feasibility) about 2 billion times per second for modern microprocessors. While human brains currently have more total processing power than even the fastest supercomputers, machines are predicted to catch up in processing power within a few decades. Digital agents can also be copied quickly, without requiring long development times. Memories and cognition modules can be more easily imported and distributed. Software code is transparent and so is easier to modify. These and other considerations are outlined by Kaj Sotala's "Advantages of Artificial Intelligences, Uploads, and Digital Minds (http://intelligence.org/files/AdvantagesOfAIs.pdf)".

Scenarios in which machines supplant biological life may sound preposterous, perhaps because they are common tropes in science fiction. But from a more distant perspective, a transition of power toward machines seems almost inevitable at some point unless technological progress permanently stagnates, because the history of life on Earth is a history of one species being displaced by another. Evolved human brains are far from optimally efficient, and so seem likely to be displaced (http://www.bbc.com/news/technology-30290540) unless special effort is taken to prevent this. Even if biological humans remain, it's plausible they would make heavy use of machine agents to spread to the stars in ways that are dangerous or impossible for collections of cells like us.

Do digital minds matter ethically?

If, as seems plausible, most of the minds of the far future will be digital rather than biological, where does this leave the cause of reducing animal suffering? Biological animals may remain after the machine transition, or they may go extinct as people and machines appropriate their habitats in order to create more machines. But either way, they may eventually be vastly outnumbered by digital intelligence, especially if machines colonize space (http://www.amazon.com/Superintelligence-Dangers-Strategies-Nick-Bostrom/dp/0199678111) in order to acquire more computing substrate and energy.

A few philosophers (https://en.wikipedia.org/wiki/The_Emperor%27s_New_Mind) maintain that artificial intelligence (AI) cannot be conscious, so if we care only about conscious suffering, AIs would not show up on our moral radar. But most philosophers (http://consc.net/papers/qualia.html) and scientists (https://en.wikipedia.org/wiki/Consciousness_and_the_Brain) agree that AIs could in principle experience conscious emotions if they were constructed in the right ways, and much of the debate concerns where the boundary between unconscious and conscious machines lies.

I think it's pretty likely that most people would grant human rights to very human-like AIs, such as perfect "uploads" of human brains -- if only because people could see such AIs generate poignant reports about the depths of their emotional lives. Many science-fiction movies have already popularized the idea of caring about humanoid robots. Indeed, people often express sympathy even for simple robots (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2044797) that have far less intelligence than a fruit fly.

perhaps because the robots look like humans or pets. We might compare this trend to the way in which humans care more about "cute" animals than ugly or disgusting animals of comparable cognitive ability.

I think the more neglected moral issue with digital minds is whether people will care about agents that don't appear human-like, don't have robotic bodies, or can't communicate. Most of the artificial, harnessable computation that happens on Earth today takes the form of silent, invisible processes within personal computers, cluster computers, and other digital devices. Operating systems, monitoring routines, data processing, large-scale machine learning, and network management are more representative of computation today than a few embodied robots or virtual chatbots. The same may hold in the future insofar as any advanced civilization will need large amounts of infrastructural computation to manage its resources and development.

But would these "business-oriented" information-processing computations have moral significance? Probably their micro-level behavior and even collective dynamics would be different from that of human brains. Some artificial algorithms, such as reinforcement learning (http://arxiv.org/abs/1410.8233), do show surprising similarity with biological cognition, but even in these domains, there are often artificial tweaks to the algorithms that render them less reminiscent of their biological inspirations. It's possible that even very sophisticated computational systems of the future will show only modest resemblance to human-style emotion.

At the same time, one lesson that animal advocates have learned is that minds can matter morally even if they're built differently from human minds. For example, birds lack the mammalian neocortex, but they have a different structure (http://www.sciencedaily.com/releases/2012/10/121001151953.htm) to perform comparable functions. Insect brains look even less like those of mammals, but some scientists and animal advocates feel that even insect suffering (http://reducing-suffering.org/do-bugs-feel-pain/) should be averted. How far away from human resemblance we want to extend moral consideration is a deep and difficult question, but it's one that needs to be explored more seriously, and animal advocates have an important place in this discussion. For instance, it was only because I already cared about animal welfare that I even considered the idea that powerless subprocesses of a future intelligent civilization's computations might have moral standing. The focus by animal activists on suffering by voiceless creatures who are often hidden out of sight has relevance to the case of digital sentience.

Some amount of instrumental computation will be necessary for any intelligent civilization, for the same reason as many computers are necessary for business, governance, and maintenance today. But even if we can't prevent suffering computations from being run, we can apply the same Three Rs (https://en.wikipedia.org/wiki/The_Three_Rs_(animals)) as in the context of animal welfare:

1. Replace more suffering-like computations with less conscious or more happiness-like computations.
2. Reduce the number of suffering-like computations run.
3. Refine suffering-like algorithms to be "more humane", even if doing so sacrifices efficiency or performance.

Virtual animals

While it seems relatively clear that any spacefaring advanced civilization will need lots of instrumental computational subprocesses, such a civilization would also have a vast scope for discretionary computation, in a similar way as consumers in wealthy countries have money to spend on luxuries, hobbies, and entertainment. There are many possibilities for what these discretionary computations might look like. For instance:

- They could support a vast population of human-like digital minds (http://www.nickbostrom.com/astronomical/waste.html) in advanced virtual societies.
- Computational power might be parceled out to each person to use as s/he pleases. Perhaps some people would create vast video-game worlds, others would explore the depths of mathematics, and others would run lots of copies of themselves.
- Artificial mind architectures might replace human-like minds altogether and create a vast, unified, galaxy-wide agent pursuing various goals.
- An AI might spend all of its discretionary computation trying to solve a mathematical problem or optimize the layout of the physical universe.
- ...and there are many other plausible outcomes.

At least some of these scenarios entail large numbers of complex digital agents in lifelike virtual worlds. Perhaps some environmentalists would build animal-rich virtual habits to a high level of precision. Biologists might develop artificial animals (https://en.wikipedia.org/wiki/Artificial_life) and simulate their evolution (https://en.wikipedia.org/wiki/3D_Virtual_Creature_Evolution), including the pain of being eaten alive or dying from parasitism. Video gamers might create vast arrays of sentient creatures to be mowed down with machine guns. The animals in these scenarios might or might not look like the fauna we see on Earth, but some of them might have comparable intelligence and emotional depth as chickens or cows.

With respect to these scenarios, animal advocates help push the future in a better direction by nudging society toward more humane, anti-cruelty norms, which should help slightly decrease the probability that our descendants tolerate vast amounts of virtual animal suffering.

descendants tolerate vast amounts of virtual animal suffering.

Don't stop thinking about the future

These scenarios, and perhaps others that we haven't yet imagined, are significant but are also distant and abstract. Do they have practical implications for the present?

One implication is that it would be valuable for more people to explore and write about the future of animal suffering in general. Insofar as the values and dynamics of future civilization are influenced by what we do in the present, even small changes to steer humanity in better directions now may prevent literally astronomical amounts of suffering down the road.

Another implication is that even if you do continue to focus on the very urgent cause of animal suffering in the here and now, you can learn more about which short-term campaigns are likely to have more long-term payoff. For instance:
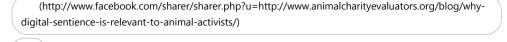
- In general, encouraging greater concern for animals seems more important than we might have thought, since these social values may be somewhat sticky. (That said, we shouldn't ignore the power of technologies and laws to change moral attitudes.)
- Raising awareness of the extent of suffering in nature and challenging ecocentric values seem more important than we might have thought, since even if humans can't do a lot about wild-animal suffering on Earth, they could much more easily avoid spreading wild-animal suffering into space and virtual ecosystems.
- Pushing the boundaries of animal moral patiency to include fish, crustaceans, and possibly insects could help set the stage for wider circles of compassion toward low-level, voiceless, invisible digital minds.
- Technological, economic, political, structural, and institutional dynamics may matter a lot because pushing on these may have an outsized influence on the far future. The influence of values on history was arguably small compared with the influence of politics, economics, and especially technology. While economic and technological evolution may be harder to change than values, we can make some tweaks on the margin, such as by promoting technologies likely to have better impacts sooner or by generating more social discussion of certain technologies so that they have more oversight during development.

Considering these points may temper some apparently stark differences among animal charities. For instance, it's plausible (http://www.veganoutreach.org/articles/chart.html)that farm-animal charities help tens or hundreds more animals per dollar than lab-animal charities, but the far-future impacts of the two seem less likely to differ by tens or hundreds of times. Maybe it's actually better to encourage concern for lab rats than for cows because rats are traditionally more disliked, so this work has more effect of challenging superficial prejudices. Maybe it's easier to get people on board with lab-animal replacements because this requires less personal sacrifice. Or maybe the farm-animal domain is more memetically effective because it's harder to argue that lab animals are unnecessary than that meat is unnecessary, or because veg outreach touches far more people per dollar than does lobbying scientists and corporate executives behind closed doors. The upshot of this particular debate is unclear, but we can see how the metrics for assessing impact become broader and less clear-cut when consequences for the far future becomes an important component of the equation.

Continuing to care about the short term

Low-probability, high-impact effects of our work on the far future seem very important, but they don't have to dominate our calculations. Personally I feel a "spiritual" need to prevent some animal suffering in the short run, in addition to considering outcomes in the very long run. It somehow feels "heartless" to focus exclusively on speculations about a far future that may very well not happen for one reason or another. And while encouraging greater concern for animals in general seems positive across a wide array of future scenarios, one could imagine cases where it backfires.

So when assessing the impact of our work on animal (and other forms of non-human) suffering, I think we should look both at more tangible and more speculative considerations. We have to decide for ourselves where to strike the balance between the two.

(http://reddit.com/submit?url=http://www.animalcharityevaluators.org/blog/why-digital-sentience-is-relevant-to-animal-activists/&title=Why+Digital+Sentience+is+Relevant+to+Animal+Activists)

Posted in Philosophy *(http://www.animalcharityevaluators.org/blog/category/philosophy/)*

Previous:

Leave a Reply

Name *

Email *

Website

Comment

Post Comment

☐ Sign me up for the newsletter!

The ACE Blog

Researcher Referral (http://www.animalcharityevaluators.org/blog/researcher-referral/) | *Feb 7, 2015*

Why Digital Sentience is Relevant to Animal Activists (http://www.animalcharityevaluators.org/blog/why-digital-sentience-is-relevant-to-animal-activists/) | *Feb 3, 2015*

Community Comments

- Raphael on Breaking Down the Leafleting Study (http://www.animalcharityevaluators.org/blog/breaking-down-the-leafleting-study/#comment-6497)
- Erica on Why the Situation of Animals in the Wild Should Concern Us (http://www.animalcharityevaluators.org/blog/why-the-situation-of-animals-in-the-wild-should-concern-us/#comment-6471)

ACE Newsletter

Sign up for the ACE Newsletter (/about/media/newsletter/) to receive monthly updates. View the latest newsletter (http://us6.campaign-archive1.com/?u=b4291a948c00a2b15c0615e45&id=44b3dc4056).