Science, technology, and morality

Jeremy Maitin-Shepard
Michael Rectenwald
76-101 D

December 10, 2004

Recent advances in biology and psychology have provided dramatic new insights into the understanding of biological processes, specifically the operation of human and animal brains, and have allowed explanation of human behavior and psychological conditions in terms of environmental and genetic factors; at the same time, technological advances, particularly in biotechnology and artificial intelligence, raise the possibility for genetically engineered humans, animals, and hybrid organisms, direct interfacing between brains and machines, and development of machines capable of highly intelligent, sophisticated and human-like behavior. These advances, in blurring the distinction between humans, animals, and machines, and in suggesting biological causes for individual ability and behavior, pose problems with common conceptions of morality and personal responsibility, which depend on such distinctions and on uncertainty regarding the causes for individual ability. Specifically, consideration of these possibilities exposes the common classification of humans, animals, and other objects for differential treatment as largely arbitrary.

Ethical controversy over technological innovations is nothing new; military technology, life-support systems, birth control and abortion technology, and neurosurgical procedures have always spawned such controversy. But these recent advances do more than create ethical controversy; they expose fundamental problems with common moral theory. Regarding brain-related biotechnology and genetic engineering, although ethical responses to specific

technologies and procedures vary widely, there is general consensus that certain applications of these technologies are clearly unethical, and that society must place certain limits on the use of such technology. Although 'unethical' situations can be avoided by effective prohibition of such uses, prohibition does not resolve the problems such possibilities pose to the foundation of moral theory. In the case of artificial intelligence, because the technology in question is not generally believed to be unethical, some philosophers simply deny that human-like behavior will ever be achieved, while others attempt to distinguish in some way between human behavior and the behavior that could possibly be achieved through 'artificial intelligence;' detailed analysis, however, shows flaws in such arguments that artificial intelligence poses no challenges to the bases of moral theory. Specific examples of recent developments and future possibilities in genetic engineering of humans, artificial intelligence, and the interfacing between brains and machines show that such advances pose unique challenges to moral theory.

The theoretical possibilities for genetically engineering humans and animals are nearly limitless. The relevant forms of genetic engineering include so-called gene therapy, or insertion of genes into somatic, namely non-reproductive, cells of existing humans or animals, and so-called germ line manipulation, which involves either prenatal selection of embryos based on their genetics or insertion of genes into embryos or the reproductive cells of existing humans or animals; of particular note regarding germ line manipulation is that any changes made affect all of the descendants of the human or animal.[4, 5]

The extent to which a particular instance of genetic engineering is controversial depends largely on the type of genes affected. Changes involving insertion of specific genes in order to prevent or reduce the risk of specific diseases are less controversial compared to the more controversial insertion of genes for the purpose of improving fitness of various sorts, or for the purpose of affecting behavioral traits or physical traits. (Although behavior may be substantially affected by the environment, there is good reason to believe it is at least

partially controllable through genetic manipulation.) Far more controversial possibilities exist, however, such as insertion of animal genes into human cells or insertion of human genes into animal cells. The effect of this, to produce a human-animal hybrid organism, could possibly also be achieved by adding human stem cells to animal embryos or adding animal stem cells to human embryos. Even more controversial is the possibility of designing new proteins for specific purposes, and inserting into human or animal cells the DNA sequence to produce the protein, and thereby allow humans to function in ways that neither humans nor animals have functioned before.

Many of the challenges posed by the possibilities of genetic engineering to common or traditional morality are readily apparent. In the case of human children with a selected or 'designed' genome, depending on the extent and precision of the selection or design, the children could easily be viewed as inferior to their designers, as is argued by the organization Human Genetics Alert.[3, p. 7] Furthermore, depending on the extent to which behavior of the children is predictably controlled, it could well be reasonable to morally hold the designers partially responsible for the actions of the children, even after such children become adults. Assuming that no 'abnormal' genes are inserted, the children would 'objectively' be no different biologically from other people, which would imply that, according to common moral theory, they should not receive differential treatment, and specifically they should be considered no less autonomous than other people; in this way, common moral theory proves to be incoherent in this case.

Genetic engineering also poses problems with common moral theory to the extent that there is an association made between merit and ability; evidence of such an association can be seen, among other things, in the favorable connotations of meritocracy. It seems contrary to general belief that a person who uses genetic engineering to enhance certain of his abilities should consequently be considered to have greater merit, particularly since such enhancements would surely be more available to those with more money. At the same

time, however, education seems to be commonly associated with merit, even though better education is available to those with more money. More fundamentally than merit, however, common conceptions of morality place humans in a class of their own, in so far as they claim humans have certain inalienable rights, while claiming that all other things, including machines and animals, have few or no rights. If a group of humans were genetically engineered with various enhancements, such that they become superior in ability to ordinary humans, and also diverge from ordinary humans to the point that they cannot reproduce with humans, it would be reasonable to say that these individuals are not humans, and yet it would also seem unreasonable to deny them the same rights as humans, particularly since they have superior abilities; in this way, classifying individuals for differential treatment, upon which common conceptions of morality depend, proves problematic. Richard Hayes, former assistant political director and national director of volunteer development for the Sierra Club, comes close to this point in stating that "[d]evelopment and use of these technologies would irrevocably change the nature of human life and human society. It would destabilize human biological identity and function."[2] In fact, though, a stronger point can be made: that the mere possibility of the development and use of these technologies is sufficient for exposing incoherences in common morality and its associated conception of a "human biological identity."

The concept of artificial intelligence similarly challenges conventional moral theory. The field of artificial intelligence covers a broad range of topics, unified only in that they all relate to developing algorithms for producing intelligent and useful behavior. Of particular interest, however, is the so-called Turing Test, proposed by Alan Turing in an article *Computing machinery and intelligence*[8]; he proposes the test as a suitable alternative to the question, which he argues is meaningless, of whether machines can think; the purpose of the test is to determine whether a particular algorithm can reliably imitate a human in a natural language dialogue; more specifically, an algorithm passes the test if a human interrogator is

unable to reliably distinguish between responses given by an actual human and responses given by the algorithm. As the field of artificial intelligence has matured over the past fifty years, and researchers have realized the need to develop specialized approaches to specific problems, the usefulness of the Turing test as a practical benchmark has come into question. Nonetheless, the test is of great significance in suggesting the theoretical possibility of a machine indistinguishable through verbal communication with a human, and in this way presenting certain philosophical questions. Clearly, it would be inconsistent with common moral theory to give any sort of rights to a machine, but a machine capable of passing the Turing test could, by definition, argue as convincingly as any human that it should receive the same rights as humans for the same reasons that members of previously oppressed groups have argued that all humans should receive the same rights.

A common rationale given for distinguishing between human and machine intelligence is that machines lack 'consciousness,' which humans, it is stated, surely have. This claim is in fact considered in Turing's original article; essentially, because a precise definition of consciousness cannot be given, the claim boils down to nothing; there is no reason to assume other humans are intelligent or conscious except by communicating with them, and since a machine passing the Turing test is indistinguishable from a human in verbal communications, consciousness of such a machine could just as well be inferred. John R. Searle argues, however, that such a response is unsatisfactory: he argues that human consciousness is rooted in the biological nature of the human brain.[6] Although Searle's claim can be dismissed as arbitrary and unfounded, recent developments in brain-machine interfaces allow for a much stronger response.

There has been substantial development recently in brain-machine interfaces, which depend on usefully associating the electrical activity in the brain with higher-level mental processes. Electrical signals in the brain can be monitored using electrodes either placed outside the skull or implanted inside the skull; through a calibration procedure, the readings

of these signals can be correlated to some degree with certain slightly higher-level mental processes and states; devices have been developed that allow humans to learn to control a computer cursor through such an interface; under a more refined system, sophisticated brain control of computers might be possible. Communication in the opposite direction is also possible. By sending small electrical impulses to specific regions of the brain, specific sensations can be induced. The recent success in 'remotely controlling' rats using a brain-machine interface is of particular note. Electrodes were implanted into the brains of rats such that electrical signals could be used to separately give the rat a rewarding sensation, simulate stimulation of the left whisker, and simulate stimulation of the right whisker. A training procedure was designed such that the rats would learn that simulated stimulations of the left or right whisker would be followed by the rewarding sensation if the appropriate movements were made. The rats were thus trained to behave in certain ways based on the electrical signals, and in this way they could be remotely controlled.[7]

More recently, an advance in far more sophisticated brain-machine interaction was made. 25,000 brains cells taken from a rat were cultured in a petri dish. Over a short period of time, the cells formed neural connections. Electrodes at the bottom of the dish allowed the network of neurons to interact electrically with a computer. By connecting the network of neurons to a flight simulator, and providing certain feedback signals, the network of neurons was made to learn in only a short amount of time to interact with the flight simulator such that the plane achieved the desired behavior.[1] Thus, the neurons acted as a living computer. Although the task to which the living neural network was applied in this particular experiment was of limited complexity, it is straightforward to see that a larger number of brain cells could be used to achieve more sophisticated tasks. In particular, if an algorithm for passing the Turing test were developed, such a living neural network could be trained to execute it, which would invalidate Searle's claim; common conceptions are morality can at best give ambiguous answers regarding the sort of treatment a biological computer capable of passing

the Turing test should receive, particularly if human brain cells were used. Furthermore, in so far as such devices could be used on fully developed human brains, they raise the possibility of some sort of human-computer hybrid, which similar to human-animal hybrids, is highly problematic for moral theories.

Ultimately, the theoretical possibilities suggested by recent technological and scientific advances challenge the bases on which moral theory classify the group of things or individuals that should receive certain 'rights'. Classifications based on biological similarity to humans prove problematic in considering the possibilities of genetic engineering; such classifications inevitably exclude genetically engineered humans that it would seem should intuitively receive the same 'rights' as ordinary humans. Classifications based on intelligence of behavior alone, however, are problematic for several reasons; they exclude humans with certain mental disabilities, while theoretically including machines capable of passing the Turing test. Even classifications based on some combination of biological and behavioral criteria, such as the criteria given by Searle, prove problematic when the possible applications of brain-machine interface are considered. Given these problems with common existing moral bases, the only coherent basis for a moral theory may be essentially an economic one, specifically a theory that dictates treatment based solely on expected benefit to the society; even this definition leaves open the question of coherently defining the relevant group to consider as the society.

References

[1] Carolyn Gramling. Uf scientists: "brain" in a dish acts as autopilot, living computer. *University of Florida News*, October 2004. URL: `http://www.napa.ufl.edu/2004news/braindish.htm`.

[2] Richard Hayes. The quiet campaign for genetically engineered humans. *Earth Island Journal*, 16(1), Spring 2001. URL: `http://www.earthisland.org/eijournal/new_articles.cfm?articleID=98&journalID=44`.

[3] David King. Why should i be concerned about human genetics? Briefing, Human Genetics Alert, March 2002. URL: http://www.hgalert.org/briefings/briefing1.PDF.

[4] Rachel Massey. Engineering humans, part 1. *Rachel's Environment & Health News*, 720, March 2001. URL: http://online.sfsu.edu/~rone/GEessays/engineerhumans.html.

[5] Rachel Massey. Engineering humans, part 2. *Rachel's Environment & Health News*, 721, March 2001. URL: http://online.sfsu.edu/~rone/GEessays/engineerhumans.html.

[6] John R. Searle. Minds, brains, and programs. *The Behavioral and Brain Sciences*, 3, 1980.

[7] Sanjiv K. Talwar, Shaohua Xu, Emerson S. Hawley, Shennan A. Weiss, Karen A. Moxon, and John K. Chapin. Behavioural neuroscience: Rat navigation guided by remote control. *Nature*, 417:37–38, May 2002. URL: http://www.nature.com/cgi-taf/DynaPage.taf?file=/nature/journal/v417/n6884/full/417037a_fs.html.

[8] A. M. Turing. Computing machinery and intelligence. *Mind*, 59(236):433, October 1950.