

Summary of Masters Thesis:
“Man vs. Machine – An Exploration of the Concept ‘Continuity’”
Johnny H. Søraker

Submitted to Dept. of Philosophy, Norwegian University of Science and Technology, Summer 2003

Published in *Filosofisk institutts skriftserie #47*, Trondheim: NTNU University Press

The purpose of my Masters thesis was to develop a conceptual framework for analysing the relation between human beings (moral persons) and other entities that share a subset of our properties. The background for this project was MIT historian Bruce Mazlish's claim that humans are *continuous* with machines, in the same way that we are continuous with animals and the world at large. Rather than focusing explicitly on whether humans are indeed unique or not, my aim was to reach a better understanding of what such a claim amounts to. This is important since the uniqueness of man has important implications for our conception of ourselves, our relation to other entities, and our corresponding moral obligations. The sometimes paradoxical claims about human uniqueness is perhaps best illustrated by animal experimentation, which on the one hand rests on the assumption that there is a continuity between man and animals (i.e. the experimental results are relevant for humans) but, on the other hand, that there is a discontinuity between man and animals when it comes to moral obligations.

The analysis is carried out on the background of Mazlish's claim (which is inspired by Freud) that there have been four dramatic scientific revolutions in the history of mankind, and that these revolutions are best described as the establishment of *continuities*; From Copernicus, through Darwin and Freud, mankind has come to acknowledge that we do not occupy a special place in the universe, on the earth, or even in our own minds. In other words, there is no sharp discontinuity between our planet and the rest of the universe (Copernicus), between humans and animals (Darwin), or between rational and irrational humans (Freud). Mazlish argues that we should also overcome what he terms *the fourth discontinuity*; that there is no sharp distinction between humans and machines.

A fundamental problem with Mazlish's analysis, however, is that the arguments are sometimes merely metaphorical, the concepts are vaguely defined, and it is unclear whether the concept of 'continuity' has any significant meaning at all. Furthermore, Mazlish fails to address the controversial debate on anti-reductionism in philosophy of mind. On this background, my aim was to get a better understanding of what it means for something to be "discontinuous" (qualitatively different), and what the criteria for establishing discontinuities ought to be.

After having analyzed Mazlish's concept, it soon became clear that the concept of discontinuity have different meanings in different contexts. For instance, claiming that mainstream science regard humans as continuous with other animals is

completely different from claiming that most members of a given culture regard man and animal as continuous. On this background, chapter 1 singles out four distinct meanings to the concept: historical, scientific, theoretical and common-sense (dis-)continuities. The importance of the latter type of discontinuity is evidenced by Freud's analysis of the major "blows to the self-esteem of mankind", which Mazlish fails to incorporate. Mazlish, in an almost Heideggerian manner, regards continuity as a harmonic acceptance of being-there, of being one with the world, whereas Freud regards them as harmful blows to our self-esteem, thereby underlining the tension between science and popular opinion, as evidenced by the ridicule faced by Copernicus (and Galileo), Darwin and Freud himself. In order to encompass the insights of both Mazlish and Freud, I proceeded by separating between two further notions of continuity, which I termed *upgrading* and *degrading* continuities. In an *upgrading* continuity, the domains that have been previously regarded as distinct are now continuous in the sense that the domain that was previously regarded as inferior is *upgraded* to the level of the other domain. A *degrading* continuity is a continuity in which the domain that was previously regarded as superior is *degraded* to the level of the other domain. For instance, a continuity between humans and animals can be either upgrading (animals are worth as much as humans) or degrading (man is worth as little as animals). This difference can also be seen in the debate on artificial intelligence, where one can either ask "is man nothing but a machine?" (*degrading*) or "can machines do the kind of things that humans do?" (*upgrading*). In the full thesis, this is elaborated on the background of Hegelian dialectics.

With these conceptual clarifications in place, I proceeded by analyzing what kinds of criteria that could determine whether or not two domains are continuous. Mazlish seems to claim that this is determined by whether or not the inner workings of two objects can be explained within the same scientific framework. For instance, if behaviourism can successfully explain the actions of both humans and animal, this would suggest that humans and animals are continuous. If computationalism can explain both computers and the brain, this would suggest a continuity between man and machine. In light of the principle of charity, I tried to establish a coherent theory on the basis of Mazlish's sketchy remarks, before rejecting this approach for a number of reasons. My main line of criticism is that almost anything can in some sense be explained within the same scientific framework, thus Mazlish's concept of continuity becomes at best a metaphorical device. For instance, it is probably possible *in principle* to explain the workings of man and a light bulb by physics alone. Or, as Daniel Dennett has claimed, anything can be explained *as if* it is an intentional agent (the intentional stance). Due to these problems, significantly elaborated in the full thesis, I proposed an alternative way of defining 'continuity'. In what I termed the *differential* approach, 'continuity' can be defined as two domains that share the same set of scientific levels of explanation. That is, an entity is continuous with another entity if and only if those entities require the same set of scientific levels of explanation. For the sake of the argument, I took the three main levels of scientific

explanations to be the physical level, the behavioural level, and the phenomenological level. My proposal was that entities that *require* the same levels of explanation are continuous. I also underlined that such a claim can be interpreted as an epistemological, or an ontological claim. For instance, it might be impossible *in practice* to fully explain humans without a phenomenological account (which makes that level of explanation epistemologically necessary), or it might be impossible *in principle* (which makes it ontologically necessary). Thus, if humans cannot be fully explained by physical and behavioural explanations, and requires a phenomenological explanation as well, an entity must require the same levels of explanation in order to be regarded as continuous with humans – either in practice (epistemological continuity) or in principle (ontological continuity). Again, such a continuity can be regarded as upgrading or downgrading. If an animal or a machine were to *require* a phenomenological explanation (i.e. the same set of levels of explanation), this would upgrade them to the level of humans (upgrading continuity). Conversely, if we were to discover that humans can be fully explained *without* a phenomenological level of explanation, this would degrade humans to the level of animals (degrading continuity). In the full dissertation, I defined more precisely what a ‘level of explanation’ really is, drawing on the work of Jaegwon Kim.

In light of these considerations, the rest of the dissertation argues that the phenomenological level of explanation is indeed necessary in order to fully explain humans, and that it is (currently) not needed in order to explain machines and the majority of animals. This argumentation is in part based on the debate between Daniel Dennett and John Searle on whether or not man can be fully explained from a third-person point of view.

In order to define the concept of continuity even more precisely, I also needed to address the Sorites paradox, which illustrates the problem with regard to drawing a line between two parts of a spectrum. Humans have undeniably evolved in gradual, small steps from one-celled organisms, so how can we justify that man is discontinuous from animals? This problem was addressed by adopting a three-valued logic in which a part of the spectrum is simply ‘undeterminable’. From this perspective, it becomes less controversial to claim that there is a discontinuity between two ends of a spectrum if a sufficient part of the “undeterminables” does not, as a matter of fact, exist. Although man has evolved gradually from apes, the gradual, small steps between are no longer existing, hence the notion of discontinuity is not threatened by the Sorites paradox (compare with a line-up of 50 people ranging from 150cm to 200cm. If you remove the persons with a height of 170-180 it becomes less controversial to separate the group into tall people and short people).

As a final reflection – and as an argument to the necessity of the phenomenological level of explanation for humans – I analyzed philosopher Mortimer Adler’s *The Difference of Man and the Difference It Makes* (1967) in which he argues that man is different in kind from animals due to having a sufficiently

different kind of behaviour. In other words, Adler tries to show that man is discontinuous from animals by, incidentally, using an approach similar to the one endorsed by Mazlish, i.e. that two domains are discontinuous if they cannot be explained by the same scientific framework. By highlighting Adler's arguments, I show that he (unwillingly) finds himself defining the difference between man and animal on the basis of the phenomenological level, which exemplifies my claim that discontinuities can and ought to be established on the basis of a difference in the set of explanations needed to fully explain an entity.