PICTURING MIND MACHINES

Cable Room #1 & #5, 2008

Front page: Cable Room #4

The new field of Artificial Intelligence called Machine Ethics is concerned with ensuring that the behaviour of machines towards human users and other machines is ethical. The edited collection entitled 'Machine Medical Ethics' aims to provide an interdisciplinary platform for researchers in this field to present new research and developments in Machine Medical Ethics. In the epilogue chapter 'Picturing Mind Machines' Neuroscientist Simon van Rysewyk (S) and Visual Artist Janneke van Leeuwen (A) discuss whether it is possible to artistically represent machine self-direction or autonomy and what must be shown in such a representation. The unfolding dialogue is interwoven with examples of Janneke van Leeuwen's work, who creates photographic Mind Models of psychological states. Starting out with the question what constitutes a 'normal human being', the scientist and artist try to come to a mutual understanding of how value systems of the human mind could be pictured and translated into intelligent machines.

Scientist: Machine ethics concerns whether there ought to be machine decisions and behaviours in the real world, where such decisions and behaviours have real world effects. We researchers think the machines in question will need the capacity for self-direction or autonomy in decision and behaviour, just like normal human beings. Do you think it is possible to artistically represent machine self-direction or autonomy? What must be shown in such a representation?

Artist: I hope you don't mind if I ask you a few questions first in response to your introduction. To begin with, what do you consider to be a 'normal human being'? I would already find that very challenging to define. Furthermore, for a machine to act like a 'normal human being', you would have to create models that could describe not only concepts like self-direction or autonomy in mechanical terms, but every possible psychological process that goes on in a human mind. You asked me whether I think it's possible to artistically represent machine self-direction or autonomy [but first]

'Do you as a scientist think the human mind could be faithfully represented as a machine?'

Scientist: Minds are brains and so part of the natural world. There is robust evidence and argument that the mind relies on the approximate laws of physics, chemistry, and biology, and on mathematics and logic; it is not gifted from some supernatural realm. I think that's news for great optimism and excitement, because it implies there's no theoretical reason why Mind cannot be simulated, 'faithfully represented', in a machine.

Artist: I agree the human mind is generated by the brain and hence rooted in the natural world and its physical laws. Still, the way the combined activity of all these natural forces creates our sense of personal identity remains a magical phenomenon. On the one hand, human brains process every kind of information in predictable patterns. On the other hand however, the personal experience of and reactions to these processes are completely different for everyone. This trade off between unique experiences and predictable patterns of brain activity is I think the most fascinating characteristic of the human mind. It's also the point of departure in my artistic practice, in which I create photographic mind models of psychological states. Although everybody has different representations of the world inside their minds, on a more abstracted level they overlap with the internal worlds of other people and it's these abstracted internal worlds that I try to visualize.

Scientist: Interesting. Can you describe to me how these pieces achieve your stated goals of co-representing individual differences and similarities in mental states? Are the patterns shown significant?

Artist: The series 'Cable Rooms' was the first mind model I made and I regard it as the basic model from which I still build on. It was inspired by the statement made by Descartes that a human was an automaton, a type of spring-loaded machine, similar to animals, except in one respect, that it contained a 'rational soul' that could initiate volition of the body by its own accord. Especially that last part intrigued me, because how exactly would that 'rational soul' initiate the movement of the human automaton if you would look into its inner mechanisms? I decided to explore this question by approaching my own mind as an automaton, using photography as an analogy for internal representations. Could I represent the mechanisms of my own mind in such a way it would relate to other people as well? I felt this would be strongly dependent on the level of contextual grounding the model would have. The more abstracted spaces and objects are from a recognizable time and place, the less personal they become. Photography's ability to distort perception of space and dimensions is also an important aspect of my work. The mind models I create are all build in miniature scale, but by using a wide-angle lens they look like much larger constructions in the photographs. Careful observation of the elements used in the 'Cable Rooms' will provide hints on how they could be understood, but there is no right way of 'getting it'. I aim to expose the mechanisms of my mind through the eyes of the camera, in the hope others can intuitively connect to the images as well. However, it is impossible for me to declare to which degree I succeeded in making these models extend beyond my own mind. That's a question only others can answer.

Scientist:

'Is it a question an intelligent machine could answer?'

Artist: That's a very interesting thing to ask! I think that depends on how you would want to approach this. If you would want to find out whether my mind models represent some kind of mathematical truth about the functioning of the human brain, I'm sure an intelligent machine could test this. This is a completely different question however from asking the machine how it values the photographs as an artistic expression. Even though education and familiarity strongly influence the validation of a work of art, the aesthetic appreciation and emotional evaluation of it is an ultimately intuitive judgment, which I think would be much harder to program.

Scientist: Humans and machines understand the world differently: the objects and shapes in human art cannot be understood by machines in the form in which they are presented in a gallery exhibition. In the same way, the codes that are understandable to an intelligent machine are merely incomprehensible symbols to human aesthetic understanding. Do you think the gap between human and machine understanding can be somewhat bridged? If not, why?
Imagine taking an image of any canvas painting and reducing it to its fundamental hexadecimal code. If this code is then read by an intelligent machine, it would understand what the code is and thus what the painting represents or symbolizes. Is it possible to use code to create simple geometrical objects, representations of mathematical constants and recreations of masterpiece paintings that both humans and machines can understand? This would amount to creating objects in two forms: as the traditional material object (e.g., a cube) and as the hexadecimal and binary codes that symbolize them.
Thus, abstraction becomes material, the meanings for humans and machines are somewhat bridged.

Artist: To begin with, I would like to stress that there is a big difference between understanding a work of art as a formal object and appreciation of it on an aesthetic and emotional level. In my opinion, the example you give above only relates to the understanding of a work of art as a formal object. More importantly even, the reduction of a masterpiece into hexadecimal codes, which is then turned into both geometrical forms and binary code is a very subjective intervention in itself; an artistic interpretation of a masterpiece rather than a truthful representation. The masterpiece could have just as well been left out to create an artwork as such and in my opinion this approach would still give no meaningful insight into the aesthetic experience of a work of art. It would be more helpful if the intelligent machine would be programmed with a set of aesthetic values, measured against its personal beliefs and a database of historical and contemporary art knowledge. It would be interesting to see how the intelligent machine would integrate and draw conclusions on these often contradicting sources of input. The series 'Golden Rules' I made in 2012 alludes to the ambiguity of knowledge and rules when it comes to understanding art. It is a reference to the Golden Ratio, a mathematical proportion, which is regarded as the key to understanding beauty in nature. In my 'Golden Rules' images, however, there is no such immediate understanding, since it is unclear what the golden instruments are supposed to do. In this way I wanted to emphasize [that]

'not everything can be determined.
A certain level of uncertainty is always present and actually is the essence of beauty.'

Scientist: I think your insight about uncertainty and human creation can be generalized: uncertainty is always present when human beings create something. Is that true? Consider human engineers creating an intelligent machine.
The problem here is uncertainty in the form unforeseen side effects: when two or more systems, well-engineered in isolation, are put into a super system, this can produce interactions that were not only not part of the intended design, but also harmful; the activity of one system inadvertently harms the activity of the other. A way to help avert this particular problem is to design the subsystems to have fairly impenetrable boundaries that coincide with the epistemic boundaries of their human designers. That last thought makes me think: successful human-machine interactions may well require that the thought boundaries of machines should be fairly close or even coincident in some cases to the thought boundaries of their human creators. If we are going to trust machines in such interactions, to see in them an authentic interaction partner, they will have to incorporate our thought and value systems, including ethics.

Golden Rule #2, #4 & #6, 2012

Section Room #3, 2013

Back page: Touch Screen #1, 2013

Artist: I think that when intelligent machines would be capable of this deeper level of understanding, it would be because they would have the ability to abstract patterns from specific situations. This would also allow them to reflect on their own processes, form a Theory of Mind of other intelligent beings similar to them and operate independently.
As long as they would use this 'mind reading' with a morally sound intention it would be highly beneficial to humans to co-operate with these machines as it would expand both our physical and mental capacities. However, if they lacked moral awareness, humans would be facing intelligent machines with an antisocial 'mind-set' and it is not difficult to imagine how that could have very harmful consequences. The series 'Section Rooms' is a representation of that possibly harmful force in a single mind. This rather cold mental space is translucent, yet not fully transparent and could fold out in many possible ways. The colour purple stands for nobility, but here it is presented in a fluid and ominous form, about to slide over the edge.

Scientist: Vivid. Here is another representation of the same point: the boundaries of some or all 'section rooms' in machines and humans should coincide. That might help to regulate human-machine interactions. This is essential if humans are going to trust machines, since trust relies on the feeling that those you are interacting with share your basic ethics and concerns; or minimally, will act within the constraints prompted by those ethics and concerns.

Artist: I agree! We find ourselves in a very interesting moment in history [where]

**'we as human beings are merging more and more with technology,
that is becoming increasingly intelligent and independent.'**

It is our responsibility to do all we can to ensure that these technologies will behave morally once they are beyond our control. This also means we have to be keenly aware that our personal desires and impulses might be harmful to others, which many of us so easily forget when we're engaging with the internet for instance. The series 'Touch Screens' is perhaps a good work to conclude with in that respect. This work is a reference to the lure and seduction of the virtual worlds behind the smooth surfaces of the many computer screens we engage with everyday. Yet no matter how convincing the illusions are these virtual worlds create, I hope there will never come a time this will completely substitute genuine human interaction.

This publication has been based on the Epilogue in 'Machine Medical Ethics', an edited collection by van Rysewyk, S.P. & Pontier, M. (Eds.), p. 425-430, Springer Publisher, New York, October 2014.

Artist website: www.jannekevanleeuwen.com

Contact: info@jannekevanleeuwen.com

Janneke van Leeuwen is represented by Beers Contemporary in London.
Gallery website: www.beerscontemporary.com

Enquiries: info@beerscontemporary.com