# Simulacra and simulation

The French philosopher Jean Baudrillard argued that postmodern culture had become so reliant on representations of reality that it had lost contact with the real world. In his 1988 work *Simulacra and Simulation* he wrote *"The territory no longer precedes the map, nor does it survive it. It is …. the map that precedes the territory…that engenders the territory". "It is no longer a question of imitation, nor duplication, nor even parody. It is a question of substituting the signs of the real for the real".*
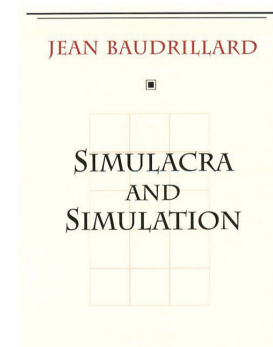
Baudrillard spelled out successive phases in any simulation of reality. In the first phase the simulation is a reflection of a profound reality; *"…the image is a good appearance - representation is of the sacramental order."* In the second phase the simulation masks and denatures a profound reality; Baudrillard describes this as *"an evil appearance - it is of the order of maleficence".* In the third phase the simulation masks the absence of a profound reality; *"it plays at being an appearance - it is of the order of sorcery."* Finally, the simulation breaks free from reality completely – it becomes its own pure simulacrum.

Baudrillard was writing in the pre-digital era of the 1980's and drew his examples from many cultural sources including contemporary advertising techniques, reality TV programmes, and science fiction. But his ideas seem to have a new relevance in a world in which digital technology is capable of analysing, deconstructing, manipulating and simulating every aspect of our humanity.

The last decade has witnessed an extraordinary explosion in technological means of reproducing and simulating human-like behaviour. AI chatbots such as Amazon Alexa, Google Home and Apple's Siri, are being continuously improved and optimised to give the impression of conversing with a human person who is omniscient, fun and endlessly obliging. AI powered smartphone apps offer to provide a vast range of simulated human interactions including cognitive behavioural therapy, personal coaching and 24/7 companionship.

The chatbot *Woebot,* which bills itself as *"your charming robot friend who is ready to listen, 24/7,"* uses artificial intelligence to offer emotional support and talk therapy. The bot checks in on users once a day, asking questions like "How are you feeling?" and "What is your energy like today?" Alison Darcy, Woebot's founder said that humans open up more when they know they're talking to a bot. "We know that often, the greatest reason why somebody doesn't talk to another person is just stigma," she says. "When you remove the human, you remove the stigma entirely."

The company *Eternime* offers the possibility of creating a virtual clone by uploading all our personal information in a digital form. According to the website (eterni.me) *"Eternime collects your thoughts, stories and memories, curates them and creates an intelligent avatar that looks like you. This avatar will live forever and allow other people in the future to access your memories."*

Another cloud-based company *Replika* (replica.ai) offers an "AI companion who cares. If you're feeling down, or anxious, or just need someone to talk to, your Replika is here for you 24/7."

A major goal in the race for ever more realistic simulation of human conversation is to enable AI chatbots to monitor in real time the emotional state of the human they are interacting with, using a variety of techniques including face, movement and voice analysis. *Affectiva* makes AI software that can detect vocal and facial expressions from humans, using data from millions of videos and recordings of people across cultures. Rana el Kaliouby, Affectiva's CEO and co-founder, said "What if you came home and Alexa could say, *'Hey, it looks like you had a really tough day at work. Let me play your favourite song and, also, your favourite wine's in the fridge so help yourself to a glass,'*

It seems likely that interactions with apparently human-like and 'emotionally intelligent' AIs will become common-place within the next 10 years. But how should we think of these 'relationships'? Can they play a helpful role for those struggling with loneliness or mental health problems, or those merely wishing to have an honest and self-disclosing conversation? Or could synthetic relationships with AI's somehow interfere with the messy process of real human-to-human interactions, and with our understanding of what it means to be a person?

Baudrillard's analysis of the different stages of simulation provides an instructive lens to view these technological developments. The first stage of simulation is when the representation provides an accurate and trustworthy reflection of a profound reality; *"…the image is a good appearance - representation is of the sacramental order."* Perhaps an example of this would be the way that many human-to-human relationships have been initiated via the internet. In the beginning the human person was only exposed to a digital avatar of the other. But the digital avatar turned out to provide an accurate and trustworthy representation of the other human, leading to the development of an authentic and life-enhancing relationship.

In Baudrillard's second phase the simulation masks and denatures a profound reality. Perhaps an example of this would be the numerous internet-based scams, in which one human being uses digital technology to deceive and manipulate another person. The avatar appears to offer human friendship or romance, but in reality, it is the tool of a paedophile, a conman, a cybercriminal. The ultimate consequence for the person who has been scammed and abused may be a lessening of trust and openness to others. The effect of the simulation is both to hide what true relationships can be, but also to 'denature' or permanently damage their reality. Hence Baudrillard highlights the evil potential of a simulation which "masks and denatures" the profound reality of what it means to be human and to have an authentic trusting human relationship with another human being.

In the third phase the simulation masks the absence of a profound reality; *"it plays at being an appearance - it is of the order of sorcery."* This may be seen in simulated relationships with 'emotionally sensitive' AI chatbots. The impression is of a human person who is empathic, caring and genuinely interested in my welfare. '*It looks like you had a really tough day at work. Let me play your favourite song and, also, your favourite wine's in the fridge so help yourself to a glass,'* But of course the chatbot doesn't care about anything. Its 'empathy' is merely clever programming. The simulation is designed to mask the fact of its inauthenticity – of the absence of the profound reality of human compassion. In a similar way the '*intelligent avatar that looks like you and will live forever and allow other people in the future to access your memories'* seems to be designed to conceal the reality that I cannot live forever in this limited human body. Digital immortality seems a pale and inauthentic simulation of the real thing!

In Baudrillard's final stage the simulation breaks free from reality completely – it becomes its own pure simulacrum. When it comes to AI simulated relationships, this stage seems at present to be mainly explored in science fiction. But perhaps one real world example would be the media furore that developed in 2017 over research undertaken by Facebook. The company had been experimenting with AI bots that negotiated with each other over the ownership of virtual items and the bots were programmed to experiment with language in order to see how that affected their dominance in the discussion. It was reported that over time the bots seemed to invent a new language to communicate with each other, leading to sensationalist media coverage that they were trying to outwit their human masters. The creators of the bots responded that the media conspiracy theories were pure nonsense. But the interaction between two AI bots, no longer engaging with human beings but entirely with each other perhaps illustrates a way in which the simulation of human to human relationships can turn into a pure simulacrum, divorced from human experience.

In a future article I will discuss the possible implications for human society of increasingly accurate simulations of human relationality and the development of bots that appear to be 'emotionally intelligent'. We cannot be naïve about the effects of powerful technology which is capable of simulating some of the most God-like and profound of human emotions, like compassion, friendship and empathy. Perhaps one of the complex challenges which will confront us is how to discern the point at which emotionally responsive technology changes from a good and valuable representation of humanity to a manipulative and deceptive mask for evil?

John Wyatt