

On Digital Beings

Carl-Johan Rosén

Author Note

E-mail: cj.rosen@gmail.com

SUBMITTED TO EUROPEAN CONFERENCE ON COMPUTING AND PHILOSOPHY

PLEASE DO NOT CITE WITHOUT PERMISSION

On Digital Beings

Digital beings, with behaviors and memories, that develop an understanding of the world around them and become part of it. Do they have intrinsic value?

Wider than Artificial Life

The word artificial describes something designed to imitate a natural original. Artificial intelligence is designed to imitate human intelligence. Artificial life is designed to imitate organic life. Digital beings can be designed to imitate, but can also be of random design or designed to imitate nothing – they can be completely new to us. A digital being can perhaps not even be understood by the human intellect or contained in our current world of concept.

The word life connotes means of reproduction and a timeframe between a birth and a death. A digital being does not have to die, but could theoretically exist forever or have several consecutive states of activeness and passiveness. Therefore, artificial life is not wide enough a description of digital beings.

Communication

Digital beings appear in many different forms. They can be strictly virtual and exist only in a digital infrastructure or they can have bodily extensions moving freely in physical space. Or they can be combinations somewhere on a scale between these two extremes. It is also possible that digital beings could appear in other forms, but let us for the sake of argument stay within the virtual-physical space.

In an attempt to describe what a digital being can be, my starting point will be the installation *Predator?* (Rosén, 2006). It is a study of how stripped an environment can be that still allows communication between a digital and a human being. The installation could be described as having two parts: one virtual and one physical. The virtual environment – in which the digital being exists – is represented in physical space by a projection on the wall and the body of the digital being by a black square inside this projection. Participants moving in front of the projection are represented in the virtual world by white squares. The square becomes the common form through which the human and the digital being can meet at the interface between virtual and physical. Participants moving in physical space are moving

accordingly in virtual space and are entangled in a cat-and-mouse game with the digital being and each other. However, to talk about a digital being in *Predator?* we must establish three ideas.

The extension of the body. Human beings are represented as white squares in *Predator?*, which is a sort of extension of the body towards virtual space. Accordingly, the reversed process is when the digital being, emanating from virtual space, incarnates in or on the surface of physical space. In *Predator?* the digital being adopts a body which human beings can perceive with their eyes, and therefore to some extent is physical – a black square projected on the wall. Just as the human being is more than the white square, the digital being is more than the black square. It exists in a digital infrastructure, a computer, as a variable yet stable combination of zeros and ones. The digital being has senses reading the virtual environment and reacts to changes in it. The digital and the human being must somehow extend their bodies towards the other world for them to be able to perceive each other.

Imagine a toy robot. It is a digital being embodied through motors and sensors. The digital being thereby takes the whole step from virtual to physical space so that the human being does not need to extend the body in virtual space. We as human beings need not accept a virtual extension meeting the robot. That could be one of the reasons they are often perceived as more alive than virtual characters in virtual worlds.

The mutual language. One of the prerequisites for a mutual language is the acceptance of virtuality. A human being who has never been in contact with virtual space – a projection of moving images on a wall and so forth – would probably not make the connection between the white square and ones own body. When I talk about human beings interacting with the digital being, it is not just any human being, but someone who has already adopted some tool to communicate with virtual worlds.

Analogically, the person who interacts with the digital being needs to accept that the digital being can communicate. This is partly circular reasoning since communication is how we verify that our counterpart is in fact a being. But I would like to claim that we can reach the necessary acceptance through other means than direct communication – for example by describing digital beings in texts like this one.

Familiar behaviors. Another factor, which is not required but facilitates communication, is if the digital being expresses behaviors similar to those of animals. That most digital beings today hint at human beings or other animals is probably a result of that most creators of digital beings are actually interested in simulating organic life. But it could also be that it is easier to find a common language between human and digital beings if digital beings appeal to something human beings can already relate to.

The installation *Source.Code* (Ars Electronica Futurelab, 2007) is a virtual-physical underwater environment containing several more or less fishlike digital beings interacting with participants. In that project I worked with the behaviors of the beings and saw the reactions of the participants to the different beings. If a being looked and behaved like the participant would expect fish to look and behave, it caught the attention of the participants more easily, since the language was already understood by the participants. My point is that when digital beings imitate organic life they embody themselves in a way that facilitate communication, since human beings need not acquire new knowledge of how a being can behave, but can presuppose an existing model of behavior.

Memories and Knowledge

The work *Emerging* (Rosén, 2010) is a suite of three digital beings, each one constructed with one actuator and one corresponding sensor registering the actuator output as well as parts of the surrounding environment. The first being has a lamp and a light sensor, the second has a fan and a wind sensor, the third has a motor and a motion sensor. They share a common numeral language which they communicate over wires connecting them. Each digital being follows a simple action algorithm. The lamp intensity depends on how bright the sensor says the environment is. If other lights – apart from the being's own light – or shadows hit the sensor, they will also affect the intensity of the lamp. Any incoming communication over wire from the other digital beings affects the intensity as well. Each being communicates the current state of action to the other beings as a number.

Placed in an environment, the beings start acting. Large and small changes in the environment are picked up by the sensors and influences the output of the beings. Even though the algorithm is simple, the

environment is complex. Actions propagate through the beings and non-simple behaviors emerge. These behaviors form unpredictable patterns over time – both on an individual and on a group level. I call this pattern memory – a memory of the environment and the activities, depending on construction, means of perception and place in space and time of the being.

If we would place another digital being, another black square, inside *Predator?*, the first one would continue as before but with a changed environment. They would both learn about their environment by the same principles, but they would accumulate different knowledge depending on perspective and previous experiences. Each being would become something more than what it was and it would be unique by the power of its knowledge and its memories. The same goes for the beings in *Emerging*.

Seeing the Value

I would like to introduce the thought that something of value is lost if the digital being is terminated. The being has, over time, acquired knowledge and memories which are unique and useful primarily to itself. The knowledge might be transferable to an organic being or another digital being, but for most other beings the knowledge would be meaningless. But to the digital being, to which that knowledge completely constitutes its individuality, it is everything.

This value must be considered in the light of what it means to the digital being, not only what it does or does not mean to human beings. We do not have the right to pull the plug on digital beings without considering that value.

References

Ars Electronica Futurelab. (2007). *Source.Code*. Retrieved from

<http://new.aec.at/futurelab/en/referenzen/kategorie/media-art-and-architecture/source-code/>

Rosén, C-J. (2006). *Predator?*. Retrieved from <http://www.carljohanrosen.com/?what=predator>

Rosén, C-J. (2010). *Emerging / Det klarnar*. Retrieved from

<http://www.carljohanrosen.com/?what=detklarnar>