The Philosophy of Computer Games Conference, Bergen 2013

Representation and virtuality in computer games

Rune Klevjer

In this paper I want to discuss the nature of participatory real-time graphical environments, which is the general form that makes a broad range of action game genres possible, from *SpaceWar!* (1962) to contemporary 3D action-adventures, including the FPS genre. Real-time graphics is also of course established outside videogames, found in anything from creativity software to fully immersive VR applications, but is nevertheless, I would argue, associated primarily with videogames and videogame culture. In public discourse as well as in academia, the general category of "videogames" is even sometimes used – however incorrect – synonymously with "games with real-time graphics".

I will suggest that real-time graphical environments, when engaged with in the habituated and intuitive manner demanded by games, is best understood as a simulation of a physical world, which is aptly accounted for by the notion of the *virtual*. Responding to John Richard Sageng's (2012) suggestion that this notion is an empty in-between category and a philosophical dead end (an escape route for a phenomenon that seems to fit uneasily between the real and the fictional), I will suggest a positive definition of virtuality, and argue that it should be conceptualised as an ontologically irreducible category.

Let us begin by assuming that real-time graphical environments, of the kind I am concerned with in this paper, belong to the broader category of simulations. Broadly defined, a simulation is the implementation of a model. A model is a "functional representation", according to Umberto Eco (1976, 209). It reproduces some functional aspect that which it models. A broom, for example, may be used as a model in so far as it is able to reproduce some functional aspect of a horse or similar creature. As a rough categorisation, a model can be either concrete (a tin soldier, a crash test dummy) or abstract (a climate model, a path-finding algorithm). Mimetic games like Risk or chess also rely on models, which are implemented as simulations during play, although this aspect may not necessarily be a significant part of the experience.

In terms of their reality status, game simulations, and especially computer game simulations, seem to belong somewhere in-between *reality* in the ordinary sense (outside the simulation) and *fictions* in the traditional sense: stories, novels, films. Espen Aarseth uses "the virtual" and "the simulated" synonymously:

In short, games are not fictions, but a different type of world, between fiction and our world: the virtual. There are also other worlds: dream worlds, thought experiments, religious perceptions, mirror worlds, etc. All these are different alternatives to our own world, and as different from fiction as they are from each other (Aarseth 2007, 39).

It seems clear that the simulated worlds of computer games are quite unlike fictions, at least as the notion of "fiction" is commonly understood: as worlds and events that are projected in our imagination. However, we should note that Aarseth does not distinguish between computer-generated simulations and other simulations in this respect; computer game worlds are "virtual" not because they are computerised, but because they are simulations. In contrast, Philip Brey defines virtuality as a product specifically of the capacities of the digital computer, even if his observation is otherwise quite similar:

At first glance, the ontological status of virtual entities is puzzling. They resemble fictional objects like characters in novels or movies because they do not have physical existence: they have no mass and no identifiable location in physical space. However, virtual entities are not just fictional object because they often have rich perceptual features and, more importantly, they are interactive: they can be manipulated, they respond to our actions, and may stand in causal relationships to other entities. So in our everyday ontology, virtual entities seem to have a special place: different from physical entities, but also different from fictitious or imaginary entities. (Brey 2003)

Although Brey emphasises that virtual entities of this particular kind belong to the category of simulation[1], he is not concerned with the nature of simulations as such. What he is interested in is the way in which the computer, unlike any other technology, is able to simulate physical reality, by generating objects that are not physical entities (no mass, no location), but which yet behave *as if* they were. It is this simulated physicality which gives virtual entities a "special place" in our "everyday ontology", not the simple fact that they are simulated entities.

However, even if we accept, in broad terms, that the capacity to simulate physical environments is key to computer game worlds' status as a "third place" in relation to fictions and ordinary reality, as I think we should, this does not in itself address the difficult question of *representation* in real-time graphical environments: When we interact with "quasi-physical" ladders, guns, water or whatever, in ways that intuitively resemble some important aspect of how we would interact with the corresponding objects and elements in real life, does this imply that our own interactions are also merely *representations* of actions rather than actual actions, just like, presumably, computer-simulated water is a representation of real water?

In his article "In-Game Action" (2012), John Richard Sageng discusses this problem, and proposes a solution. The heart of the problem, he argues, is that that action and representation are mutually incompatible categories; they articulate opposite "directions of fit" in the relation between the subject and the world[2]. Representations are directed at affecting our perceptions and understandings about the world, which means that they have a mind-to-world direction of fit (2012, 7). In contrast, actions have world-to-mind direction of fit:

Action is also a concept that that relates a subject to the world, but the notion is tied a different set of components involved in the success of an opposite direction of fit. The concept of an action is tied to the role of assigning the proper place of an autonomous subject in the causal order of happenings. Cars stop, bodies move, glass break, but what makes some of

[1] Brey opposes the category of virtual simulations to virtual "reproductions" like money, documents or colour – entities which straightforwardly have the same reality status in computer-mediated form as in any other.
[2] Anscombe (1957)

these happenings actions and others not, is the active participation of a subject in making the world fit with his wants (Sageng 2012, 7).

So it appears that you can't have it both ways. Any given act – say, drinking a glass of water – must either be evaluated on the basis of its intention to represent something, in other words as what we may call a *mimetic* act, or it must be given status as an action in the proper sense, which in this case would presumably be intended to cause reduced thirst. You are either causally motivated, and *do* something, or you *represent* something. You either kill someone or you pretend that you kill someone.

With respect to action in computer game environments, therefore, Sageng's solution is that we must separately account for two different types of acts which take place simultaneously: the causally motivated action, *and* the act of representation. In order to qualify as proper action, the former cannot be directed at guns or ladders, but must be causally directed at something that is actually existing on the screen – which is a stream of "non-representational graphical happenings" (2012, 1). The representational acts must be understood as a specialised context, an *interpretation* of events, mirroring the player's real actions in some relevant respects. So even if we do, often unavoidably, interpret on-screen action in representational terms, *action* can only proceed given a "reference shift" from mimetic events to the real objects in the game, the shapes on the screen. Hence, representational acts in computer games are an "optional" dimension (2012, 20), piggybacking on actions that are straightforwardly real. The player is *literally* manipulating dynamic shapes on the screen, while at the same time having the option to *pretend* to be shooting enemies with a gun.

Once the categorical incompatibility of action and representation is established, based on their opposite directions of fit, there is, as Sageng points out, no use for notions of simulation or "virtuality", of the kind proposed by Aarseth or Brey. Simulation does imply that representation happens through action – but this does not solve Sageng's problem, in so far as it merely states that such a double articulation is possible, and does not address the puzzle of actions versus simulated actions.

Indeed the built-in tension in the virtuality account was addressed quite comprehensively – although from a cultural history point of view – by Roger Caillois in his classic *Man, Play and Games* (2001), where he explains how and why games of make-believe and rule-based competitive games do not mix well. Theoretical considerations as well as empirical evidence shows that any particular gaming activity will fall down on either one or the other side of the divide: either as "mimesis" or as "agon". A dress-up football match, for example, is either not *really* about the dress-up, or, as would mostly be the case, not *really* about winning. In Sageng's terms, we could say that one of them, either the action or the representation, would have to take the back seat, as a kind of parallel commentary track.

The incompatibility thesis seems to be unavoidable when considering a range of non-computerised mimetic activities, which for the purpose of this paper I will roughly divide into two categories. They are either *enactments* (acting, performing, role-playing, many kinds of ritualistic practices), or *mimetic games*, like Monopoly or paintball. Whereas the former type of activity is defined by acts of pretence (pretended motivations, goals and effects), the latter type is defined through real actions, which are optionally interpreted or "themed" through make-believe contexts in various ways – as has been extensively analysed by Juul (2005), Salen&Zimmerman (2004), and Järvinen (2009), among others. In Torben Gragh Grodal's terminology we would say that whereas enactments, like stories and movies, offer its participants indirect or second-hand experiences, games offer first-hand experiences (Grodal

2003). Whereas enactments use props as tools of expression, mimetic games use *models* as enablers of action and simulation.

The notion of *framing*, as introduced by Gregory Bateson and developed by Erwin Goffman (1974), is a useful way of explaining the optional status of representational events in mimetic games or other kinds of model-based activities.

Roughly, a "frame" is a social convention that governs people's mutual understanding of what any particular social situation is about. People's interaction at a weekend party, for example, will be framed by a set of more or less standardised conventions, and this frame enables the participants to orient themselves and be reasonably in tune with each other with respect to how to behave, what things mean, and generally "what it is that is going on" (Goffman 1974, 247). The boundary of a given frame can be standardised or in flux, fixed or unstable. More important in our context: The meanings of social interactions are often governed by a *composite* of frames, in which a given frame may operate as a superimposition or "lamination" on another frame, operating as a modifier that changes the original meaning of what is going on. A typical everyday superimposition, for example, would be the adding of a "joking" frame, which turns a wide range of interactions and meanings, like for example aggression or surprise, into friendly fun. The act of adding of such a "laminating" frame, either through language or through actions, is called "up-keying" by Goffman.

In mimetic games, the keying of a superimposed frame happens through the implementation of models. The formal rule-system articulates a unifying abstract model, which may define that, for example, pawns can only move forward on the board, and that the game is over when the "King" piece is taken. When (or if) taken as an abstract model, this can be interpreted as the peasants on the battlefield not being allowed to retreat, and that the battle is over when the king has fallen. In chess, there are also concrete models, the board and the pieces, whose visual appearance as well as, to a limited degree, basic physical affordances (such as size) help to articulate certain meanings and certain structure of actions. Obviously, the world of battlefields and kings and peasants can be peeled off, as it were, and this would be what Goffman calls *down-keying*, removing the laminations of the superimposed frame.

The conceptualisation of mimetic games as a matter of framing resonates well, it seems to me, with Sageng's account of in-game action, although he suggests notion of "reference shift" rather than framings and keyings as a solution to the action-representation problem.

We may note in this context that *depictions*, or at least a broad class of them, do not follow this paradigm, as they cannot reasonably be said to depend on "up-keyings" for their meaningful realisation. A portrait painting or a piece of animation, in so far as they are intuitively and unavoidably recognisable at the level of perception, have default status as representations, in the primary framework; they cannot be "down-keyed" in the way that superimposed representational events can. One might attempt to break down or "peel off" the representational content from the supposedly *real* act of looking at a painting, but this strategy would require that we consider the act of looking at the painting as merely a framed and structured surface, independently of recognising its representational features, as an "action" in the proper, intentional sense. It seems to me, however, that this kind of "looking" would rather be an abstraction, indeed a secondary frame that is being added to the default action of looking at the painting, which is a perceptual act that implies that we recognise its representational features. I will conclude, therefore, that visual representations always operate in the primary frame; there is no meaningful frame of action to which they could be down-keyed.

In contrast, mimetic activities are always *grounded*, in the sense that the representational frame is superimposed on a primary and non-mimetic frame of interaction, to which the activity can be down-keyed; enactments can be down-keyed to the *acting* that is going on, and mimetic games can be down-keyed to the *gaming* that is going on.

Computer games may appear to fit the same model: the player plays with graphical shapes on the screen, via buttons and sticks, in such a way that the structure of actions as well as the visual dynamics are able to persuasively up-key the real actions of the player to a frame of mimetic events, superimposed onto the domain of real action.

However, let us go back to Philip Brey's virtuality account. He does not conceptualise the user's actions as directed at shapes on the screen, but as directed at computer-simulated entities. The question, then, is if the latter kind of account excludes the former. Does the possibility of action in graphical environments imply that on-screen objects are "always already" perceived as simulated physical objects? If that is the case, then the action-representations problem will be, in an important sense, unsolvable. We can only solve it if we are able to distinguish between two different levels of events, so that Brey's simulation level (action directed at already-simulated entities) can be denied status as action in the proper sense.

In non-computerised mimetic games this is not hard to do, because the activities will always have a firm grounding. When we peel of the representational layer, the physical space of natural embodiment will guarantee that there is a primary and non-representational frame of interaction to fall back on, philosophically as well as empirically; we move tokens with our hands, run around in the woods, move our bodies to express ourselves. Indeed, it is not controversial to say that the possibility for action is part of the definition of what naturally embodied space *is*.

Which brings me to my central concern: in computer games, it is unclear how our interaction with real-time graphical environments has a similar grounding in naturally embodied space, *independently* of our perception and manipulation of simulated objects. While this may be true for some games (Tetris?), it seems more doubtful in the case of games that rely on elaborated real-time simulated environments, for example racing games. If there were to be such a frame, within which we could secure the independent status of proper actions, the intentional object domain would either have to be the physical interface itself (controllers, sticks, buttons), or it would have to be a set of non-representational but playable audio-visual patterns, recognisable on the screen in front of the player. Let me first consider this problem in terms of a phenomenological notion of perceptual intentionality, before returning to Sageng's account.

The basic phenomenological argument that I want to employ here, drawing on Maurice Merleau-Ponty's *Phenomenology of Perception* (2002), is that our actions and perceptions are necessarily articulated in terms of their intentionality. When we grasp something with our hands, what qualifies as action is the grasping of the object, not the moving or our arm. And what qualifies as perception is not perceiving the movement of the arm, but perceiving the grasping of the object. Therefore, in a First Person Shooter, what we see is three-dimensional space, not the two-dimensional surface of its projection. Our intention, in the perceptual as well as in the common sense of the term, is not directed towards moving the analogue sticks, but it is directed towards moving our subjective point of view in three-dimensional space.

As I have argued elsewhere (Klevjer 2012), the standard interface that has emerged with the action-adventure genres of computer games – mouse and keyboard or buttons and sticks – requires, as an absolute imperative, that the hardware interface is incorporated as a prosthetic extension of the player's own phenomenal body, and as such no *longer qualifying* as an intentional object in perception. In navigable 3D environments, such incorporated bodily extensions typically also include not just the hardware interface, but also an on-screen controllable marionette and/or vehicle, as well as, more importantly, the navigable camera itself, to the extent that it is directly controllable.

An illustrative exception to this model would the actions of an incompetent and unsuccessful player. Because computer game play must be learned, a player who is totally unfamiliar with the basic interface of interaction will direct his or her actions at the physical interface itself, and not be able to succeed in any action as directed beyond this interface; buttons will be pressed and sticks moved around, and this would indeed qualify as proper (and successful) action, on any account.

However, to the extent that players able to engage successfully beyond this level, in the intuitive kind of way that is required by standardised pc or console interfaces, perception and action is no longer articulated in terms of the player's relation to buttons, sticks and screen, that is: no longer directed towards objects in the space of natural embodiment. The prosthetic nature of successful engagement with real-time graphical environments implies that, to the extent that our perceptions and actions are intuitively directed and anchored elsewhere, our default frame of naturally embodied action – our phenomenal *space* as seated in front of the screen – will be perceptually suspended.

Indeed this is precisely how motion-controlled casual games on consoles (Wii, Kinect, Playstation Move) clearly differ from the standard paradigm; as Jesper Juul points out (2010), such interfaces articulate a play space *here*, in front of the screen, rather than "in there", in computer-simulated space.

Sageng's account of in-game action shares the general premise that action must be intentional in order to qualify as proper agency: "The subject, having certain wants or desires, forms an intention, and through deliberation brings the intended result about by means of movements of his body" (2012, 8). The key difference from my own account seems to relate to the notion perception, even if not made explicit:

"(…) the effect of introducing action is that the players intentional object change from the fictional happenings originally represented to the things he now interacts with, which are simply the computer generated spatio-visual graphical shapes that he sees before him on the screen" (Sageng 2012, 17).

The problem will be to what extent we can say that the player actually perceives "spatio-visual graphical shapes" during play. It is true that he will be acting in such a way as to affect and respond to the behaviours of such shapes on the surface of the screen (as well as, we should add, sounds coming from the speakers). However, the sticky point is whether he "sees before him on the screen" a series of non-representational graphical shapes, such as to be able to form any possible intention towards them, either consciously or intuitively. In the default mode of competitive engagement that I have described above, I do not see how that is possible. It is especially hard to understand with respect to three-dimensional environments, which not only simulate mass, solidity and physical causality, but also spatiality and

positionality – which is projected as graphical shapes drawn on a flat surface, but hardly intelligible to the player as such.

I will suggest, therefore, that the best "solution" to our problem is to accept that the simulation of physical reality in computer games, unlike the abstract and concrete models of non-computerised mimetic play, is able to constitute its own irreducible ground of perception and action. The player's actions have no independent grounding within the phenomenal parameters of naturally embodied physical space, but must be directed towards objects and events in *virtual* space. Like depictions, our interactions with virtual objects necessarily operate in the primary frame, as there is no meaningful frame of action to which they could be down-keyed.

This is of course a philosophical problem. The rule of "direction of fit" is apparently broken; no "down-keying" or "reference shift" is possible, which could have directed us at a simulation-free domain underneath the simulated events. In other words, we just have to deal with it: computer technology has generated a truly paradoxical space for interaction and expression, in which action and simulated action cannot be separated. Let me finally attempt to also sketch a brief positive account of virtual objects and virtual space.

What distinguishes real-time graphical environments from game boards and toy guns as well as from the formal rule-systems of games, is that they are abstract models – algorithmic models – which are perceived and interacted with as concrete models. By "concrete" I do not only mean that we are able to interact with them in a general sense, but more particularly that we can touch them. The digital computer has the unique capacity to turn algorithmic models into objects of *tangible* interaction. In David Sudnow's classic analysis of videogame play from 1983, *Pilgrim in the Microworld*:

There's that space over there, this one over here, and we traverse the wired gap with motions that make us nonetheless feel in a balanced extending touch with things (Sudnow 1983).

In a certain sense, therefore, we can say that virtual space is enchanted space. In real-time computer games, when evaluated from an objective outside view, spatio-visual graphical models are second-order models (models of models). However, in the phenomenal perspective, the subject is able to, and indeed required to, engage with them as first-order models, as though they are concrete rather than informational objects and environments. So the abstract models of programmers have been, in a sense, reified. Information is made nature. In a wider cultural and historical context, such *programmed reified objects* are maybe not all that unlike enchanted mountains or Victorian automatons.

The question of the exact ontological status of virtual space, understood as programmed but reified and non-irreducible phenomenal space, must remain for now. Maybe we could say that, from the point of view of the participant, real-time participatory graphical simulations not only simulates a separate ontology but creates it. In any case, the parameters of embodied interaction are different, and the relationship between action and representation that we have come to take for granted, is recast. So I think Philip Brey is broadly correct when he suggests that computer-simulated physical environments are "perhaps less as an extension of ourselves than as an extension of our world" (Brey 2005, 395). A philosophical commitment to the intentional nature of action and perception leads us to conclude that the computer-generated worlds of *SpaceWar!* and *Grand Theft Auto V* (2013) are not, as one may assume, superimposed *frames* of action, but separate and alternative embodied spaces, engaged with as

second nature. “Virtual reality” is here quite a fitting term, even when mediated via a framed surface.

References

Aarseth, Espen. 2007. "Doors and Perception: Fiction vs. Simulation in Games." *History and Theory of the Arts, Literature and Technologies* (9):35-44.

Anscombe, Elizabeth. 1957. *Intention*. Oxford: Basil Blackwell.

Brey, Philip. 2003. "The Social Ontology of Virtual Environments." *American Journal of Economics and Sociology* no. 62 (1):269-282.

Brey, Philip. 2005. "The Epistemology and Ontology of Human–Computer Interaction." *Minds and Machines* (15):383-398.

Caillois, Roger. 2001. *Man, Play and Games*. Translated by Meyer Barash. Chicago: University of Illinois Press. Original edition, 1961.

Eco, Umberto. 1976. *A Theory of Semiotics*. Bloomington: Indiana University Press.

Goffman, Erving. 1974. *Frame Analysis: An Essay on the Organization of Experience*. New York: Harper & Row.

Grodal, Torben. 2003. "Stories for Eye, Ear, and Muscles: Video Games, Media, and Embodied experiences." In *The Video Game Theory Reader*, edited by Mark J.P. Wolf and Bernard Perron. London: Routledge.

Juul, Jesper. 2005. *Half-Real: Video Games between Real Rules and Fictional Worlds*. Cambridge, Massachusetts: MIT Press.

Juul, Jesper. 2010. *A casual revolution: reinventing video games and their players*. Cambridge, Mass.: MIT Press.

Järvinen, Aki. 2009. *Games without Frontiers: Methods for Game Studies and Design*. Saarbrücken: VDM Verlag.

Klevjer, Rune. 2012. "Enter the Avatar. The phenomenology of prosthetic telepresence in computer games." In *The Philosophy of Computer Games*, edited by Hallvard Fossheim, Tarjei Mandt Larsen and John Richard Sageng, 17-38. London & New York: Springer.

Merleau-Ponty, Maurice. 2002. *Phenomenology of Perception*. London: Routledge Classics. Original edition, 1962.

Sageng, John Richard. 2012. "In-Game Action." In *The Philosophy of Computer Games*, edited by Hallvard Fossheim, Tarjei Mandt Larsen and John Richard Sageng, 219-232. London & New York: Springer.

Salen, Katie, and Eric Zimmerman. 2004. *Rules of Play: Game Design Fundamentals*. Cambridge, Massachusetts: MIT Press.

Sudnow, David. 1983. *Pilgrim in the Microworld. Eye, Mind and the Essence of Video Skill*. New York: Warner Books.