Author's final version of a paper to appear in *Minding the Future: Artificial Intelligence, Philosophical Visions and Science Fiction* edited by Barry Dainton, Attila Tanyi, and Will Slocombe

All citations should be to the final published version.

Love in the Time of AI

Amy KIND

ABSTRACT: As we await the increasingly likely advent of genuinely intelligent artificial systems, a fair amount of consideration has been given to how we humans will interact with them. Less consideration has been given to how—indeed if—we humans will love them. What would human-AI romantic relationships look like? What do such relationships tell us about the nature of love? This chapter explores these questions via consideration of several works of science fiction, focusing especially on the Black Mirror *episode "Be Right Back" and the Spike Jonze's movie* Her*. As I suggest, there may well be cases where it is both possible and appropriate for a human to fall in love with a machine.*

Can a human love a machine? In the 1950 short story "EPICAC," Kurt Vonnegut suggested that the answer was no. EPICAC, a seven-ton machine that cost the government $776,434,927.54 to build, takes himself to have fallen in love with Pat, a mathematician who works with him on the night shift. After having several conversations about love with the story's narrator, also a mathematician, EPICAC ends up producing an epic love poem designed to win Pat over. Unfortunately for the machine, however, the narrator is also in love with Pat and passes off EPICAC's poetry as his own—so while the poem succeeds in sweeping her off her feet, EPICAC does not reap the benefits. Once Pat agrees to marry the narrator, he has to break the bad news to the machine. EPICAC is confused. He's smarter than humans. He writes better poetry than humans do. So why would Pat opt to marry the narrator rather than marry him?

The narrator has no real answer for EPICAC. After flailing about for a bit, he tries to set the matter to rest:

> "Women can't love machines, and that's that."

"Why not?"

"That's fate."

"Definition, please," said EPICAC.

"Noun, meaning predetermined and inevitable destiny." **[1, p. 282]**

EPICAC accepts the answer that he's given, having no access to material that might show otherwise. But more recent science fiction allows for a more nuanced treatment of this issue than the undefended declaration by the narrator of "EPICAC." This paper explores the possibility of romantic love between humans and machines, and in particular, what we can learn about the issue from the way that it's been tackled in two recent works of science fiction, both set in a not-too-distant future: "Be Right Back," an episode of the television series *Black Mirror*, and the film *Her* (2013), directed by Spike Jonze. As we'll see, there is more reason for optimism than EPICAC had been led to believe.

Some Preliminaries

The question of human-machine love has two parts: (1) Can a human love a machine? And (2) Can a machine love a human? Science fiction has had plenty to say about both parts. In addition to the affirmative answer to the second question that we've already seen from EPICAC, we also see numerous other cases throughout science fiction where machines form romantic bonds with humans. To give just one example: Vision, an extremely powerful Android, is in a romantic relationship with Wanda Maximoff, aka the Scarlet Witch, in films such as *Avengers: Infinity War.* Moreover, many roboticists think that reality is not too far behind science fiction on this score. In a seminal discussion of human-robot relationships, David Levy predicts that by the year 2050 robots will be developed that have the capacity to form romantic bonds with humans **[2, p. 22]**.

Exactly how plausible one will find this prediction to be depends in large part on one's definition of love. But consider, for example, views that take love to be an emotion. While it seems unlikely that love is a basic emotion like joy or anger, many philosophers and psychologists have theorized that love is a complex emotional attitude. Emotions are experiential in nature. Just like there is something it is like to feel joy or anger, there is something it is like to feel love. Thus, on this definition, the capacity to love requires one to be phenomenally conscious, to be sentient. And a similar requirement will be in place for

many other views of love that are prominent in the philosophical literature. With this requirement in place, Levy's prediction seems implausible. Though various techniques currently exist that allow robots and other machines to recognize and process emotional cues from human users and then mimic human emotions, machines have not yet developed sufficiently even to plan and carry out emotional reasoning, let alone to actually feel emotions (see, e.g., Scheutz **[3, p. 215]**; Sullins **[4, p. 398]**).

Might the requirement that a machine *feel* love be too strong? Wouldn't it be enough for the machine to produce loving behavior? Levy seems to make an argument of this sort:

> There are those who doubt that we can reasonably ascribe feelings to robots, but if a robot *behaves* as though it has feelings, can we reasonably argue that it does not? If a robot's artificial emotions prompt it to say things such as "I love you," surely we should be willing to accept these statements at face value, provided that the robot's other behavior patterns back them up. **[2, p. 11-12]**

While Levy makes an important point in this passage, it is also important not to take his argument as showing more than it does. It's true that if a machine were to produce exactly the same kind of behavior as a human being, behavior that is sufficient for us describe a human as being in love, it would seem like a kind of humancentric bias to deny that the machine can love just on the grounds that it is a machine. But that's not to say that behavior is all there is to being in love. As many philosophers have noted in response to Levy, his attempt to reduce love to the production of loving behavior should be rejected. Just as an especially proficient human actor might be able to produce loving behavior without being in love, so too might a machine. Love requires not just a certain kind of behavior but also a certain kind of mental state.[1]

When the question of machine love is addressed in the philosophical literature, the objections that are raised to this possibility often stem from more general worries about the possibility of machine sentience. The question of whether a machine can love a human (or whether a machine can love at all) thus tends to be treated less as a specific question about love and more as a general question about machine sentience. Perhaps there might be machines that, despite being sentient, still could not experience love. This kind of possibility

[1] For related criticisms of Levy, see Nyholm and Frank **[5, p. 223-224]**; Hauskeller **[6, p. 205]**.

would be an interesting one to explore.[2] But because the issue of machine love seems so tightly interwoven with the question of machine sentience, issues specific to the notion of love tend to get lost. In contrast, these issues are front and center when we address the question about whether a human can love a machine. For this reason, it's this question that I will focus on in what follows.

Science fiction has presented us with a variety of cases in which humans have fallen in love with machines—or at least, have had romantic feelings for them. Not only is Vision in love with Wanda Maximoff, as mentioned above, but she is in love with him. Numerous characters throughout the various *Star Trek* series develop romantic attachments to holodeck characters. And in the film *Ex Machina*, the programmer Caleb Smith develops romantic feelings for the gynoid Ava.

For our purposes, it will be useful to sort these examples in terms of the kinds of machines involved as love objects. At one end of the spectrum, the high end, the machines are virtually indistinguishable from humans or distinguishable only by means of special scans or tests. Consider, for example, the humanoid Cylons of the reimagined *Battlestar Galactica* television series of the early 2000s. Though they possess some abilities that set them apart from humans, they generally pass as humans in everyday interactions; in fact, they often live among humans for years without their real nature being detected or even suspected. Throughout the series, we see several instances of humans falling in love with Cylons, perhaps most notably the loving relationship between Karl "Helo" Agathon and Sharon Valerii. William's love for the host Dolores in the first season of the HBO series *Westworld* and Deckard's love for Rachael, a Replicant, in *Blade Runner* (1982) provide other examples at this end of the spectrum.[3]

At the other end of the spectrum, the low end, the machines that humans seem to love are obviously non-sentient and lack any kind of emotional intelligence whatsoever. In some of these cases the machine outwardly resembles a human being. The machine may even

[2] One possible example is Data, the android from *Star Trek: The Next Generation*. Though it seems plausible that he should be considered to be sentient, the show does not come down firmly on this question. (In the episode "Measure of a Man," Data is said to meet two of three criteria for sentience (intelligence and self-awareness), but they leave it open whether he meets the third criterion (consciousness). But, until he is outfitted with a special "emotion chip," Data lacks the capacity to experience emotions.

[3] One might question the inclusion of the *Blade Runner* example here, since the Director's Cut raises the possibility that Deckard too is a Replicant. If he were, then this would be a case of machine-to-machine love rather than human-to-machine love.

outwardly appear to be physically identical to a human being. But despite its physical appearance, its behavior is clearly off—mechanical or in some other way clunky, such that on anything more than a quick or superficial interaction there can be no mistaking that it is really nothing more than a mechanical doll. For example, in the Futurama episode "I Dated a Robot," Fry uses a celebrity-download service to create a Lucy Liu robot. But though the robot looks just like Lucy Liu, its behavior and conversation show obvious limitations, for example, the repetition of pre-programmed messages, the implausibly sexualized behavior, and the use of a stilted recording for Fry's name whenever the robot needs to mention him: "I find your slack-jawed stare very attractive, PHILIP J. FRY."

Then there are the cases that fall somewhere in between these two ends of the spectrum. To my mind, this is where the most interesting philosophical questions arise. In cases where the machine is fully sentient and all but indistinguishable from a human being, it's hard to see why we would have any reason to deny that the purported love is a case of real love. Worries that humans can't genuinely love sentient beings who are non-biological are suspiciously reminiscent of worries that humans can't genuinely love sentient beings who are of a different race or of the same sex. In cases where the machine lacks sentience entirely and is nothing but a mechanical doll, it's hard to see why we would have any reason to accept that the purported love is a case of real love. When someone claims to have fallen in love with a new pair of shoes, we don't take the claim seriously. At best, it seems like a metaphorical invocation of the notion of romantic love. Things seem no different when someone claims to have fallen in love with a mechanical doll. Even if our definition of love were technically to allow for such cases, it seems likely that they will end up being characterized as mistaken or deficient in some way.

The interesting philosophical questions thus seem to lie in consideration of the intermediate cases. As the high-end cases show, the answer to the question of whether a human can love a machine is clearly yes. Were a machine to be just like a human, so much so that we can't even tell that it's not a person, then why couldn't we fall in love it? It's only in thinking about the intermediate cases that we are productively able to shift from the question of whether a human could love a machine to the question of what a machine would have to be like in order for a human to love it, and, just as importantly, what a machine would have to be like in order for such love to be natural and appropriate.

The intermediate cases are themselves quite varied. In some, we have machines that seem to be sentient but are significantly different from human beings in other important ways. Consider L3, the droid from *Solo* (2018). Throughout the movie, viewers are strongly led to believe that Lando Calrissian loves her—and L3 herself clearly believes that he has feelings for her.[4] Though L3 exhibits human-like sentience, her robot-like body gives her a very different physical form from humans. In other kinds of intermediate cases, the machine has human-like intelligence, including emotional intelligence, but does not seem to be capable of experiencing emotions or of having phenomenally conscious experiences more generally. In some the machine provides some evidence of emotional capacity, but the evidence is equivocal. In some there simply isn't enough evidence to have a clear sense one way or the other.

Both of the examples that I will explore in this paper fall into this intermediate class—though for different reasons. In "Be Right Back," the machine in question is a humanoid robot who has been programmed with some of the memories and mannerisms of a recently deceased 20-something named Ash.[5] Though the robot looks just like a human being, the evidence for his sentience is ambiguous at best. In Spike Jonze's film *Her* (2013), the machine in question is Samantha, an artificially intelligent operating system. Though the movie strongly suggests that Samantha is sentient, she does not have a physical form. I will consider these examples in turn over the next two sections in an effort to determine what a machine has to be like in order to be the kind of being for whom a human could appropriately develop romantic feelings.

Lost Love

"Be Right Back" centers on the possibility that a machine could replace a lost love, a possibility that has long been explored by science fiction authors and filmmakers.[6] As early as 1927, Fritz

[4] Phoebe Waller-Bridge, the actress who played L3, and Donald Glover, the actor who plays Lando, believed so as well. As Waller-Bridge has said, "Both Donald and I had felt instinctively that there was a love between them, and that they were connected in a way that was romantic with a big 'R'." (See https://www.syfy.com/syfywire/phoebe-waller-bridge-on-l3-and-lando-the-first-romantic-human-droid-romance-in-star-wars)

[5] Interestingly, Ash/the Ash-Robot is played by Domhnall Gleeson, the same actor who plays the programmer Caleb in *Ex Machina*. Gleeson, then, has depicted characters on both sides of the human-machine romantic relationship.

[6] Science fiction has also often explored the possibility that a machine could substitute for an unattainable love. To give just one example, consider the Buffy Bot that was commissioned by Spike in Season Five of *Buffy the Vampire Slayer* after he confesses his love for Buffy and is harshly rejected by her.

Lang explored this idea in “Metropolis,” where the inventor Rotwang creates a humanoid robot in an effort to resurrect his lost love Hel. More recently, it’s been explored in several different media. The stage play (and subsequent movie) *Marjorie Prime* focuses on a relationship between Marjorie, an octogenarian with dementia, and the robot companion her family has hired to serve as a stand-in for her late husband. In the audio drama podcast *LifeAfter*, FBI clerk Ross Barnes begins to communicate obsessively with a digital resurrection of his wife, who has recently been killed in a car accident. And in the television series *Star Trek: The Next Generation*, after his beloved wife Juliana is seriously injured and on the brink of death, Dr. Noonian Soong creates a gynoid replica of her and transfers her memories into it.

But even if the basic premise behind “Be Right Back” is not a particularly new one, its take on the issue is fresh, thought-provoking, and slightly disturbing. Martha and Ash are a young couple in love. When the episode starts, they’ve just moved back into Ash’s childhood home, an isolated fixer-upper in the countryside. As we watch their interactions, it becomes clear how much they thoroughly enjoy each other’s company, even if Martha is sometimes frustrated by Ash’s preoccupation with social media. But then Ash is killed in a car accident, and shortly thereafter, Martha discovers that she is pregnant. Alone in her grief, and wanting nothing more than to share her news with Ash, she decides to make use of a service that a well-meaning friend had signed her up for, a service that allows individuals to stay in touch with dead loved ones via chat bots based on the deceased person’s social media posts. Though Martha is initially horrified by the idea, she ends up finding comfort in communicating with the chat-bot. Since Ash had been a heavy user of social media, the bot does a particularly good job of replicating his conversational style.

Communicating via text messaging quickly leads Martha to an upgraded service, chatting via phone, and then ultimately to an experimental service the company has just begun offering. Soon a life-sized robot, designed to look exactly like Ash and programmed with his personality, arrives on her doorstop. Though their initial interactions provide her with both company and comfort, she ultimately becomes frustrated and dissatisfied with the limitations of the robot. The episode ends with a scene that takes place several years later. In the final plot twist, a moment suffused with typical *Black Mirror* creepiness, we learn that the still-activated Ash-robot is now kept in the attic, entirely alone except for weekend visits from Martha’s daughter.

Though there are moments when Martha allows herself to think of the Ash-robot as Ash, she mostly seems to see him as an inadequate substitute. In one moment of reflection on the issue, she describes her take on the situation to him, “You aren’t you, are you? … You’re just a few ripples of you. There’s no history to you. You’re just a performance of stuff that he performed without thinking, and it’s not enough.” Viewers are inclined to agree with her assessment, and reviewers of the episode did as well. As Morgan Jeffery put the point in a piece published in *DigitalSpy*: “it's not really Ash - the replicant is hollow, without a soul - and so much of what made Ash the man he was, and the intricacies of his and Martha's life together, is lost in translation.” **[7]**

But even though it’s clear that the Ash-robot isn’t Ash, it’s considerably less clear what we are meant to think about the machine’s sentience and emotional intelligence, thus giving us the kind of intermediate case where interesting issues arise. Many of the things about the Ash-robot that bother Martha don’t seem to bear on the issue of sentience—they seem either to go towards showing that he isn’t Ash (e.g., he doesn’t remember something that Ash would have remembered) or towards showing that he isn’t human (e.g., he doesn’t need to eat or sleep or breathe). But none of this goes towards showing whether he should count as sentient. So what other evidence is there?

On the one hand, the Ash-robot does not seem to be able to feel pain, as evidenced by his lack of reaction when a shard of glass pierces his palm. He also doesn’t seem to be bothered by the slights and insults that a sentient being would be bothered by. On the other hand, he can smile and laugh and cry, and he is able to read Martha’s emotional states. And as a general matter, he responds as a human would (even if not always exactly as Ash would) in conversational interactions. That said, however, his ability to switch seamlessly from one reaction to another when the first is deemed inappropriate by Martha makes his behavior seem more a matter of algorithm than of choice.

At times this last point seems decisive—so much so that one might begin to wonder why this case falls into the intermediate range rather than at the “clearly not sentient” low end of the spectrum. But here we have to think about the end scene of the episode and, in particular, our reaction to it. We wouldn’t be creeped out to learn that Martha had consigned

her Roomba or her iPhone to the attic.[7] But we are creeped out to learn that Martha has consigned the Ash-robot to the attic. To my mind, the creepiness at the end of the episode derives, at least in part, from a worry about the Ash-robot himself, and this shows that we are thinking of him very differently from the kinds of robots at the low end of the spectrum.

Does Martha have romantic feelings for the Ash-robot? Clearly she can't bring herself to deactivate him. Perhaps this is just a kind of sentimentality over Ash. Perhaps her treatment of the robot has more to do with her feelings for Ash than with any feelings she has for the robot. But I'm inclined to think that she feels something for the robot himself. That said, we're given no reason to believe that her feelings for the Ash-robot amount to love and, perhaps more importantly, no reason to think that they should. If Martha's feelings for the Ash-robot had been different, if they had deepened and developed in such a way that these feelings started to seem more like love, we would be troubled.[8] In asking the question, "Can a human love a machine?", then, we are not just asking a question about possibility but about appropriateness. What's of interest to us is not simply whether machines could be objects of human love but whether they could be *suitable* objects of human love.

So let's think a bit about why the Ash-robot isn't a suitable object for Martha's love. Unfortunately, the issues are muddied here by the fact that the Ash-robot is designed to be a substitute for the actual Ash. In this regard he clearly fails. So we need to separate two things: the ways in which the Ash-robot fails to be sufficiently Ash-like and the ways that the Ash-robot fails to be sufficiently person-like.

Of course, these two things are not entirely distinct. Some of the ways that the Ash-robot fails to be sufficiently Ash-like arise precisely because he fails to be sufficiently person-like. When we think about these kinds of failures, we're led to see that for machines to be suitable candidates for love, it's not enough for them to be decent conversationalists and amiable companions. They need to have more fully fleshed-out personalities across a multitude of dimensions. They need to bring something to the relationship as well.

---

[7] That said, many people do turn out to be pretty attached to their Roombas—naming them and ascribing distinct personalities to them. One recent study even refers to the attachments that people have formed with their Roombas as "intimate relationships" **[8]**. As quickly becomes clear, however, the notion of intimacy in this context does not come close to rising to the level of romantic love.

[8] As Alexis Elder notes in an insightful discussion of the ethical dimensions of chat-bots that draws extensively on "Be Right Back," machines like the Ash-robot "are not people and yet they look and feel enough like them that, like artificial sweeteners, they might trick us into thinking we have something valuable that we in fact lack." **[9, p. 4]**

But this brings us to a further important point. When Martha signed up for the service that created the Ash-robot, she was not just looking for love. She was looking for Ash. And even if the Ash-robot were more fully fleshed-out, even if he brought something to the relationship as well, he still wouldn't be Ash. So even if the Ash-robot were sufficiently person-like to be a suitable candidate for love, that wouldn't be enough for Martha.

What if the Ash-robot were much better at replicating Ash? What if he were a near-perfect, or even perfect replication? Would this be enough for Martha? I'm inclined to think not. For even in this case, he still wouldn't be Ash. He wouldn't be the person with whom she originally fell in love. He wouldn't be the person with whom she shared a history of experiences, both the silly, trivial ones and the deeper, formative ones. And he wouldn't be the person who fathered her daughter.

Reflection on this case helps us to see that there's something inherently troublesome about the very project of trying to replicate a lost love. On this score, the fact that the replication is a robotic one is almost irrelevant. The same problem would arise from biological cloning. As a general matter, we tend to think of our loved ones as irreplaceable. When we lose someone we love, we cannot simply substitute someone else in their place, no matter how similar the second individual is to them. As Robert Nozick has noted, even though someone may come to love another person because of the other person's characteristics, "it is the other person and not the characteristics, that is loved. The love is not transferable to someone else with the same characteristics, even to one who 'scores' higher for these characteristics. And the love endures through changes of the characteristics that give rise to it." **[10, p. 168]**[9]

No matter how enamored science fiction is with the idea that machines can replace lost loves, then, this is one area where fiction seems to be far removed from reality. The problem is not due to technological limitations but due to psychological ones. Love is not the kind of attitude that is indifferent about where it is directed, and we cannot simply replace one love object with another, no matter how similar. If humans are to love machines, then, we have to be able to love them for who they are.

[9]For a discussion and defense of this irreplaceability, see Grau **[11]**.

Disembodied Love

Unlike the Ash-robot in "Be Right Back," Samantha is not specifically designed to substitute for a lost love. In fact, she's not specifically designed to be a love object at all. Consideration of the romance presented in *Her* thus allows us to avoid the worries encountered in the previous section.

As the movie begins, it's clear that Theodore is lonely, unfulfilled, and somewhat at loose ends after the dissolution of his marriage to his longtime partner Catherine. But he isn't specifically looking for love when he goes to purchase the OS1, advertised as the first artificially intelligent operating system. After the system is installed, Theodore is presented with Samantha, an OS who has been customized just for him. Her voice is personable and friendly, but Theodore is not sure what to make of her at first. She describes how she's been programmed, but she also notes that she's constantly evolving: "What makes me me is my ability to grow through my experiences." Theodore expresses some puzzlement: "You seem like a person but you're just a voice on my computer." But Samantha dismisses his worry, noting that it only seems that way because of the limited perspective of his unartificial mind.

Over the next few days, as Theodore and Samantha continue to interact, his limited perspective begins to broaden, and he discovers how much he enjoys their interactions. Soon thereafter they become involved in a romantic relationship. Though the relationship goes well for a while, ultimately things start to change. Samantha seems distracted when Theodore talks to her, and there are times when she's unavailable. Eventually, when they have it out, he learns that she's talking to thousands of other people—8,316 other people, to be exact—at the same time that she's talking to him. Even worse, he discovers that of the more than 8,000 people with whom she's talking, there are 641 with whom she is in love. Though she tries to tell him that it doesn't affect the way that she feels about him, he has trouble making sense of it. It's when they next talk that she tells him that she's leaving.

For the moment, in thinking about the relationship between Theodore and Samantha, let's set aside what happens at the end of the movie. After all, the fact that one partner changes in such a way that they no longer find their romantic relationship satisfying, or that something else becomes more important, does not mean that the relationship was not a genuine one before that. And indeed, for much of the film, the relationship between Theodore and Samantha seems to be a mutually fulfilling one. They play video games and go on walks. They double-date with another couple, both of whom are human. They talk for

hours and share with one another their innermost thoughts. They aim to make one another happy. And, as often happens in romantic relationships, they have sexual interactions as well.

Our focus in this essay is love, not sex, it will be worth our pausing for a few moments on this topic—partly because the notion of sex is closely intertwined with the notion of romantic love, but partly because considerations of this topic will lead us to some broader questions about the significance that Samantha's disembodiment has for her ability to be a suitable love object for Theodore. Perhaps unsurprisingly, it's not uncommon for science fiction depictions of human-machine romance to take up the question of human-machine sex as well as the question of human-machine love. One striking example comes in "The Naked Now," an episode of *Star Trek: The Next Generation*, when the android Data, about to embark on an interaction with crew member Tasha Yar, reassures her that he is "fully functional" and programmed with many sexual techniques.

Martha has sex with the Ash-robot in "Be Right Back," and the robot Ash turns out to be better able to provide her with sexual stimulation than the human Ash was, even if his sexual performance is strikingly mechanical in execution. But unlike the Ash-robot, Samantha does not have a body, so the sex between Theodore and Samantha is quite different from the sex between Martha and the Ash-robot. At first their sexual interactions are presented as something more akin to phone sex. Later, wanting to take things to the next level, Samantha hires a body surrogate to stand in for her in person.[10] For Theodore, however, the encounter is strange and uncomfortable, and he puts a stop to it before things progress very far at all.

Does the fact that Theodore cannot have sex with Samantha present an obstacle towards our understanding his relationship with her as one involving romantic love? What is the relationship between sex and love? Note first that, as a general matter, we seem to accept the conceptual possibility of sex without love. We accept it in the case of human-human interactions, and we also accept it in the case of human-machine interactions. In addition to the sex dolls that are already on the market, roboticists are at work developing machines that could serve as more interactive sex partners for human beings. In the typical cases, however, these robots that are being designed for the purpose of serving as sexual companions are not

[10] Something similar happens in *Blade Runner 2049*, when the hologram Joi arranges to merge with a replicant prostitute so that she can have sex with her boyfriend, the protagonist K.

meant also to serve as loving companions.[11] For our purposes here, however, what's more important than the possibility of sex without love is the possibility of love—romantic love—without sex. This too, as a general matter, seems to be a conceptual possibility that we accept. Perhaps because of physical distance, perhaps because of physical disability, or perhaps for some other reason entirely, some people who are in romantic love with one another do not, even cannot, have sex with one another. And just as we'd accept that two individuals who do not have sex for one of these reasons (or for a different reason entirely) can still have a loving relationship, we should accept that two individuals who do not have sex because one of them is a machine (either without a physical body or with a physical body ill-suited for sex with humans) can still have a loving relationship.

The argument just given depends in part on drawing an analogy between Theodore and Samantha's relationship and relationships between people who are physically distanced from one another. In an interesting discussion of *Her*, Troy Jollimore calls into question the appropriateness of this analogy. Importantly, his concern is not that romantic love requires sex; he grants that an inability to have sex with the person one loves, or even to be in physical contact with the person one loves, is possibly "a frustration one can learn to live with, where the love is deep enough." **[12, p. 131]** Rather, his concern is that we, as humans, cannot have romantic love for disembodied minds. Though the interactions that Theodore has with Samantha bear some resemblance to the kinds of interactions that someone might have with a lover who is physically distant, Jollimore takes this resemblance to be merely superficial. Samantha is not physically distant but rather physically non-existent. In Jollimore's view, when Theodore relates to Samantha, he mistakenly imagines her as having some physical presence: "Although Samantha has no body, he still imagines himself as relating to her body, and to her mind via her body." **[12, p. 133]** Insofar as romantic feelings for Samantha would have to be based on this kind of confusion or delusion, Jollimore suggests that we should not see her as a suitable candidate for love.

To my mind, this concern of Jollimore's should not be given much weight. It's not clear why Theodore's feelings for Samantha would have to be based on this kind of confusion, i.e., it's not clear why such confusion is a necessary feature for anyone in love with an artificially

[11] This distinction is often blurred by use of the term "robot lover." This phrase is sometimes used to refer to a robot with whom one has sex and sometimes to a robot whom one loves. Indeed, sometimes it is used indiscriminately to refer to both kinds of robots.

intelligent OS. Consider someone who is perfectly clearheaded, perfectly clear that their OS is disembodied, and who never imagines their OS as having a body. Why couldn't the same sort of relationship develop as the one that develops between Theodore and Samantha? And just as importantly, even if Theodore did imagine her this way, why must this mean he is subject to a delusion? People engage in all sorts of imaginings about their romantic interests. Someone might imagine that their loved one is taller, shorter, more considerate, more adventurous. It's not at all clear why such imaginings threaten the idea that the relationship is a healthy one, based on love.

The concern that Jollimore raises about Samantha's disembodiment is not his only concern about the appropriateness of Theodore's love for her. In his view, worries also arise from the fact that we can't really know whether she is conscious and whether she is capable of experiencing genuine emotions. His argument then goes one step further. Even if we grant that she is conscious, that may still not be enough. As many philosophers have argued, love requires forming a *we*. In making this point, Jollimore draws extensively on work from Robert Nozick, e.g., Nozick's claim that when two people form a *we*, "the people *share* an identity and do not simply each have identities that are enlarged" (**[13, p. 82]**; quoted in **[12, p. 138]**. Given the nature of Samantha's disembodied existence, her consciousness "is presumably so different from Theodore's that it will be quite impossible for them to understand each other" **[12, p. 138]**. Absent an ability to genuinely understand one another, it seems impossible that two individuals could genuinely share an identity. Thus, whatever Theodore may think that he feels for Samantha, she is not a suitable target for his love.

Let's consider both steps of this argument. First, should we share Jollimore's concern that we cannot really know whether Samantha has the capacity for consciousness and emotion? To my mind, the answer is a qualified no. One important set of considerations arises from the fact that the movie seems to intend that viewers take her to be conscious. Advertisements for OS1 describe it not just as intelligent, as noted earlier, but also as conscious: "it's not just an operating system, it's a consciousness." Samantha's conversational responses do not seem to be pre-programmed set-pieces. They are sophisticated and varied. Her responses suggest a strong understanding of human emotions. She reacts appropriately to Theodore's expressions of emotions, and she herself also evidences emotional responses that are appropriate to the situations that she encounters. She composes music that has emotional resonance. She anticipates his needs and desires and arranges thoughtful surprises

for him that suggest she is able to understand what's important to him even without its ever having been told to her directly. Based on all this behavior, it seems reasonable for Theodore not only to think that she's conscious but also that her consciousness is not radically unlike that of humans.

Could this all be an act? Might all of this just be the result of extremely clever programming? Yes, that does remain a possibility, and it's for this reason that my answer is somewhat qualified. But I take it that this remains a possibility in the way that it also remains a possibility for the Replicants of *Blade Runner* and the Cylons of *Battlestar Galactica*. As we've noted earlier, love requires not just a certain kind of behavior but also a certain kind of mental state—and there is no way for us to be absolutely certain that these machines with which we've been presented are actually in the relevant mental states. There's also no way for us to be absolutely certain about this even for other people. Granted, with other people we do have some evidence over and above behavior. Given that we can each know that we ourselves are conscious, the fact that other people are the same kinds of biological organism that we are gives us some reason to believe that they are conscious too. Ultimately, though, skepticism of the sort that drives worries about the consciousness of highly sophisticated machines—machines like Samantha and Replicants and Cylons—seems to lead one to a lonely existence in which the only consciousness one can really recognize is one's own. As Alan Turing made this point in his discussion of machine intelligence back in 1950:

> According to the most extreme form of this view the only way by which one could be sure that a machine thinks is to be the machine and to feel oneself thinking. One could then describe these feelings to the world, but of course no one would be justified in taking any notice. Likewise according to this view the only way to know that a man thinks is to be that particular man. It is in fact the solipsist point of view. **[14, p. 446]**

Let's now turn to the second step of Jollimore's argument, namely, that Theodore and Samantha are incapable of forming a *we*. Earlier I suggested that we temporarily set aside what happens at the end of the movie, when we discover that Theodore is not the only person with whom Samantha has been carrying on a romantic relationship. But as this fact plays a key role in Jollimore's defense of this argumentative step, it's now time to think more about it. Jollimore raises various concerns that stem from the lack of exclusivity on Samantha's part

– see **[12, p. 135-139]**. Because she has so many different romantic partners, Samantha's well-being cannot be especially tied up with Theodore's in the way that we would expect when two people are in love. Because so much of her life—so much of her interactions with other people—remains invisible to Theodore, it turns out he actually knows considerably less about her than he might have thought. And for the same reason, it turns out that they share considerably less than he might have thought. Normally when two people are in love we think of their having some sort of special connection with one another, but it it's hard to see how Samantha's connection with Theodore could be special given that she has a similar connection with over 600 other individuals. In short, given that so much of Samantha's attention has been directed elsewhere, it does not seem like she has fully given herself over to her relationship with him in the way necessary to becoming a *we*. Of course, the fact that she is a remarkably sophisticated AI may give her capabilities for forming special bonds that humans lack, but unfortunately the movie doesn't really help us to see how that could be the case.

Though Jollimore is right that many philosophers have built a requirement of exclusivity into their conceptions of love, this requirement has been persuasively questioned in recent defenses of polyamory (see, e.g., **[15]**). I will not attempt to settle that question here, or even to enter the debate. I do not think it is necessary to do so for our purposes. For even if we reject a strict exclusivity requirement on love, we might still be troubled by the extremely large number of romantic partners with whom Samantha is involved. Despite Samantha's insistence that her feelings for her 600+ other partners do not take away from her feelings for Theodore, despite her insistence that she's never loved anyone the way that she's loved Theodore, that she's madly in love with him, he finds it difficult to process what he's learned. And even if the number of Samantha's romantic partners were not to give us pause, there's a further important fact that we should find troubling. Samantha was not honest with Theodore about the kind of relationship that they were in. Though he was clearly thinking of the relationship as exclusive, she did nothing to correct this impression and kept him entirely in the dark about her other relationships (and even about her other conversations). This kind of deception does not seem consistent with a healthy, loving relationship.

So Jollimore is right to worry about Samantha and Theodore's relationship. But it's important to be clear about what exactly this shows, or perhaps better, what it doesn't show. Even if it turns out that Samantha's behavior detracts from her suitability as an object of

Theodore's love, it doesn't really show that she's *in principle* an unsuitable object for his love. Based on what we know about Samantha, or about the OS1 more generally, there doesn't seem to be any reason in principle that she couldn't direct her romantic attention exclusively towards a single individual. It's only in light of the lack of exclusivity of her romantic attention (and her deception about it) that we're disinclined to view Samantha as having formed a *we* with Theodore. Though these details prove essential to the plot of the movie, they don't seem to be essential to the kind of relationship that Samantha and Theodore could theoretically have. These details aside, we have not seen reason to believe that Samantha's disembodiment serves as an obstacle to her forming the kind of bond that constitutes a *we*.

One of our reasons for considering *Her* in this chapter was that it presented us with an interesting intermediate case—a case where the machine, even if sentient, was importantly different from humans. Though our discussion has raised some concerns about the relationship between Theodore and Samantha, it has also given us some reason for optimism about this kind of human-machine love more generally.

Concluding Remarks

This paper began with the question of whether a human could love a machine. Our consideration of these two science fiction examples has enabled us to see when and how this kind of love would be both possible and appropriate. In considering this question, we separated it from the parallel question of whether a machine could love a human. As a result, we've largely operated under the assumption that these questions were independent of one another. Before we close, however, it's worth noting that our discussion has given us some reason to question that assumption. When we think about the limitations of the Ash-Robot and the problems that were presented by Samantha's behavior towards Theodore, we see that romantic love lends itself towards a certain kind of reciprocity. If this is right, then the question of whether a human can love a machine depends at least in part on the question of whether a machine can love a human (and, of course, vice versa).

This is not to say that unrequited love is psychologically impossible. It's not even to say that unrequited love is unsuitable or unhealthy. But it nonetheless seems that when we try to determine whether a machine could be an appropriate kind of object for human affection, it matters whether the machine is at least in principle capable of feeling affection itself. So finally, returning to the Vonnegut story with which we began, it seems that the narrator really

did EPICAC wrong. If we assume that EPICAC really did feel love for the mathematician Pat, then there's no reason to think, in principle, that she couldn't have reciprocated that love. After all, his poetry really did sweep her off her feet.[12]

References

[1]. Vonnegut, K.: EPICAC. *Collier's Weekly*. Reprinted in *Welcome to the Monkey House.* (1950)

[2]. Levy, D.: Love + Sex with Robots. New York: Harper (2007)

[3]. Scheutz, M.: The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots. In: Abney, K., Lin, P., Kerkey, G.A. (eds.) Robot Ethics: The Ethical and Social Implications of Robotics, pp. 205-223. Cambridge, MA: MIT Press (2012)

[4]. Sullins, J.P.: Robots, love, and sex: the ethics of building a love machine." *IEEE Transactions on Affective Computing* 3: 398-409 (2012). https://doi.org/10.1109/T-AFFC.2012.31

[5]. Nyholm, S. and Frank, L.E.: From Sex Robots to Love Robots: Is Mutual Love with a Robot Possible? In: Danaher, J., McArthur, N. (eds.) Robot Sex: Social and Ethical Implications, pp. 219-243. Cambridge, MA: MIT Press (2017)

[6]. Hauskeller, M.: "Automatic Sweethearts for Transhumanists." In: Danaher, J., McArthur, N. (eds.) Robot Sex: Social and Ethical Implications, pp. 203-218. Cambridge, MA: MIT Press (2017)

[7]. Jeffery, M.: 'Black Mirror' series two 'Be Right Back' review: 'Creepy and Moving'. (2013) Available at https://www.digitalspy.com/tv/a457935/black-mirror-series-two-be-right-back-review-creepy-and-moving/

[8]. Sung , J. Y. , Guo, L., Grinter, R.E., Christensen, H.I.: 'My Roomba is Rambo': Intimate home appliances. In: Krumm, J., Abowd, G.D., Seneviratne, A., Strang, T. (eds) Proceedings of UbiComp 2007, 145—162. Innsbruck, Austria: Springer (2007) https://doi.org/10.1007/978-3-540-74853-3_9

[9] Elder, A.: Conversations from Beyond the Grace? A Neo-Confucian Ethics of Chatbots of the Dead. J. App. Phil. (2019) https://doi.org/10.1111/japp.12369

[10] Nozick, R.: Anarchy, State and Utopia. New York: Basic Books (1974)

[12] Thanks to Barry Dainton, Attila Tanyi, and Frank Menetrez for comments on previous draft of this paper.

[11] Grau, C.: Irreplaceability and Unique Value. Phil. Topics 32(1/2), 111-129. (2004) https://doi.org/10.5840/philtopics2004321/219

[12] Jollimore, T. 'This Endless Space between the Worlds': The Limits of Love in Spike Jonze's *Her*." Midwest Stud. Phil. 39, 120-143. (2015) https://doi.org/10.1111/misp.12039

[13] Nozick, R.: The Examined Life. New York: Simon and Schuster (1989)

[14] Turing, A.: Computing Machinery and Intelligence. Mind 59, 433-460 (1950) https://doi.org/10.1093/mind/LIX.236.433

[15] Jenkins, C.: *What Love is and What it Could Be.* Basic Books (2017)