

# Wanting the impossible

## The dilemma at the heart of intimate human-robot relationships

Dylan Evans

In a recent book entitled *Love and Sex with Robots*, the British scholar David Levy has argued that relationships with robot Companions might be more satisfying than relationships with humans, a claim which I call “*the greater satisfaction thesis*” (GST). The main reason Levy provides in support of GST is that people will be able to specify the features of robot Companions precisely in accordance with their wishes (which I call the total specification argument or TSA). In this paper, I argue that TSA is wrong. In particular, the argument breaks down when we consider certain behavioral characteristics that we desire in our partners. I illustrate my argument with a thought-experiment involving two kinds of robot – the FREEBOT, which is capable of rejecting its owner permanently, and the RELIABOT, which is not.

### Introduction

The idea of romantic relationships between humans and robots has long been a province of science fiction. In films like *Blade Runner* (1982) and *Bicentennial Man* (1999), for example, mutual love between a human and a robot is portrayed as both possible and beautiful.<sup>1</sup> Outside the realm of literary and film criticism, however, there was until recently very little serious discussion of the technological, social, legal and ethical questions that robots capable of such relationships might pose. The publication of *Love and Sex with Robots* has gone some way to

---

1. In *Blade Runner*, the protagonist Deckard falls in love with a robot called Rachael. There has been debate among film buffs as to whether Deckard himself is meant to be a robot, in which case this would be a robot-robot relationship rather than a human-robot relationship. The director, Ridley Scott, has gone on record stating that Deckard is a robot, but neither version of the film (the original version or the director’s cut) is explicit on this point.

changing this (Levy, 2007a). While the topic is not yet in the main stream of debate among roboticists, Levy's book has already inspired at least one international scientific conference<sup>2</sup> and attracted comment by researchers in a range of scientific fields.

At several points in his book, Levy suggests that romantic relationships with robot Companions will be more satisfying than romantic relationships with humans. I will refer to this claim as the greater satisfaction thesis (GST):

GST: It is likely that romantic relationships with robot Companions will be more satisfying than romantic relationships with human Companions.

Levy does not specify the scope of GST very precisely, but it is clear that he means it to refer to more than just a few exceptional people. Some of those interviewed for a British documentary about people who appear to have romantic relationships with sex dolls seem, on the face of it, to prefer these dolls to human partners (Holt, 2007). If a documentary maker can find some people who prefer inanimate sex dolls to human partners, then it seems very likely that, once the robot Companions envisaged by Levy become available, there will be a few people at least who prefer romantic relationships with these robots over romantic relationships with humans. If this was all that Levy meant, it would not be very surprising. But the fact that Levy clearly thinks his claim is a bold one implies that he intends the scope of GST to be quite wide. It is plausible to interpret Levy as claiming, then, that *many* people will find romantic relationships with robots more satisfying than romantic relationships with human beings. The main reason Levy provides in support of GST is that people will be able to specify the features of robot Companions precisely in accordance with their wishes. So, for example, people who are worried about infidelity will be able to request robot Companions programmed to be completely faithful. Alternatively, those who prefer the frisson of uncertainty may specify that their robot Companion be programmed to have some probability of being unfaithful. I will refer to this as the total specification argument (TSA):

TSA: One reason for supposing GST to be true is that we will be able to specify the features of robot Companions entirely in accordance with our wishes.

In this paper, I will argue that TSA is wrong. Robot Companions, at least those of the kind that Levy envisages, will not be more satisfying than relationships with humans. Before I proceed, however, I must first make it clear that the grounds for my objection have nothing to do with empirical facts about robot

---

2. The First International Conference on Human-Robot Personal Relationships was held at Maastricht University on 12–13 June 2008. This paper was first presented at the conference.

technology. The reason that robot relationships will not be more satisfying than relationships with humans has nothing to do with any technical limitations inherent in robots. There may or may not be such technical limitations, but I tend to agree with Levy when he dismisses objections of this kind, since we do not know enough about the limits of technology in general to be able to rule out, with any degree of confidence, any of the particular technical features that he discusses in his book. My objection to Levy's claim about the potentially greater satisfaction that we may obtain from relationships with robots is based, not on claims about robot technology, but on claims about the what people tend to want from an intimate relationship.

It is worth spelling out here the kind of robots that Levy has in mind. For throughout his book, Levy only discusses a relatively narrow class of possible robot Companions. In particular, the robot Companions that Levy envisages are all:

- humanoid
- made in a factory according to specifications provided by the user
- consumer products designed to serve the needs and wishes of humans
- private property

Thus Levy does not consider, for example, robots that build themselves, or robots that exist for their own purposes, or robots that have any kind of right to leave their owner's service. Such characteristics may be too advanced to merit discussion in what Levy claims is an analysis of technology that is possible within a few decades, but the uncertainty surrounding the speed of technical developments is too great to make any firm pronouncements here. Besides, some of the other features that Levy thinks will be feasible within that time span, such as learning the preferences of a user by observing their behavior, might be just as complex as the ability to build copies of itself. So the fact that Levy restricts himself to considering just one kind of robot Companion, in just one of the various possible commercial and legal frameworks in which robot Companions might operate, can at times make the book seem more like one long advert for a product that doesn't exist yet. Nevertheless, since my aim in this paper is to criticise Levy's claim, I too will restrict myself to a discussion of the type of robot discussed in his book.

This paper is structured as follows. In the next section, I discuss what I mean by "technical objections" to GST in order to make it clear, by contrast, the kind of objection to GST that I advance in this paper. In the following sections I outline my own objection to GST, and illustrate it by proposing a thought-experiment involving a choice between two different kinds of partner robot – one which is capable of permanently rejecting its owner (a FREEBOT) and one which is not (a RELIABOT).

## Technical objections

By “technical objections”, I mean objections to Levy’s claims that are based on how difficult it would be to build robots with a particular feature. Take Levy’s discussion of robot fidelity, for example. A robot that was guaranteed to be completely faithful to its owner would require some means of identifying its owner with 100% accuracy. We might argue that such a capacity is unlikely on various empirical grounds:

- i. *No clear path from current technology*: One set of arguments is based on extrapolations from specific technologies that are currently in existence. For example, at least one current facial recognition system has been shown to be 100% accurate when tested with a set of still photographs taken from the Internet (Jenkins and Burton, 2008). However, when it comes to recognising faces from a moving video camera in real time under a wide range of lighting conditions (all conditions that would apply to face recognition in a humanoid robot), accuracy rates fall dramatically, and fall far short of the recognition capacities of humans. It could be argued that there is currently no clear path from the relatively poor accuracy rates of facial recognition possessed by mobile robots to the 100% accuracy rates that they would have to enjoy for robots to be capable of perfect fidelity.
- ii. *General engineering principles*: A second set of arguments might start, not with specific technologies, but with general engineering principles. For example, it could be argued that no technology can ever be 100% accurate because there are always multiple failure points in any machine.
- iii. *Hacking and deception*: A third set of arguments could appeal to the possibility that even if recognition systems could be built that were, in fact, 100% accurate under most circumstances, they would always be vulnerable to hacking and active deception. Let us imagine, for example, that a robot recognises its owner by means of some biometric system such as iris-recognition, or by some other means such as detecting an RFID tag with a unique user ID (perhaps embedded subcutaneously). Let us suppose, further, that this system never gave a false positive. Now, it would always be possible to fool such a robot by various means. For example, a malicious user could hack into the file which the robot used to associate the category ‘owner’ and the relevant iris scan or user ID. Alternatively, they could produce a fake retina that mimicked the pattern of the owner’s iris, or copy the owner’s RFID tag.

All these arguments have some merit. But to take these as the only arguments against Levy’s claim that robots will be potentially more satisfying than human Companions is to agree with Levy that, if such technical hurdles *are* overcome,

robots may indeed prove more satisfactory Companions than humans. Since it is this assumption that I wish to challenge, I will suppose for the sake of argument that all the technical limitations can be overcome. Let us suppose, for example, that it is possible to design a robot capable of 100% accuracy in recognising its owner, and therefore capable of being completely faithful if required. Would this give us good grounds for supposing that relationships with robots will be more satisfying than relationships with humans? I think not. The reason, as I explain in the following section, has to do with certain fundamental features of human emotions.

### Wanting to be wanted

Intimate relationships between humans are difficult and complex things, in which conflict and pain are pervasive features. If any evidence were required to support this claim, it could be found in abundance in the advice columns of agony aunts, in the statistics for divorce, and in the long tradition of tragic love stories in plays, novels and more recently, film. Even in the happiest and most long-lasting marriages, spouses acknowledge the persistence of stress and conflict (Ammons & Stinnett, 1980).

GST suggests that some of the unpleasant features of human-human relationships will be absent from, or at least diminished in, human-robot relationships, providing the robots are appropriately designed. This implies that the problem lies in some contingent feature(s) of human Companions that prevents them from fully satisfying our desires, and that appealing robots can be designed without this feature or set of features. According to TSA, the contingent features of human Companions are their failure to fully match our preferences. Presumably, the line of reasoning is as follows:

People do not come made to order, and so we must settle for some compromise between what we really want and what we can get. This difference between what we really want and what we have to settle for is at least partly responsible for the less pleasant aspects of human relationships. Robots, on the other hand, are made to order. Unlike people, they can be designed to match our preferences perfectly. Therefore, those problematic aspects of human-human relationships that are caused by the gap between desire and reality will be absent from human-robot relationships.

This argument may have some degree of plausibility when we consider only the physical characteristics that we desire in our partner. Someone may want a boyfriend who is very tall and muscular, or a girlfriend with large breasts, but may not be able to find someone with all the physical characteristics they desire (or

they may find someone with these characteristics, but that person does not reciprocate the affection). In this case, it would presumably be easy to specify a robot partner with all the right proportions. However, the argument breaks down when we consider certain behavioral characteristics. For among the various desires that most people have regarding relationships are certain second-order desires (desires about desires), including the desire that one's partner has freely chosen to be one's partner, and has not been coerced into this decision.

With the kind of robots that Levy envisages – robots that are ordered from a factory, and purchased as consumer goods – it is clear that this desire, at least as stated above, cannot be fulfilled. These robots do not choose their owners, at least initially. Nevertheless, if the robot had the capacity to reject its owner, robots that did not exercise this capacity might be said to choose their owners in some ongoing sense, even if there was no choice initially. In the absence of an initial choice on the robot's part, therefore, the robot's capacity to reject its owner would be crucial for satisfying the owner's desire that he or she has been freely chosen (if the owner does have this desire).

Levy anticipates this idea when he states that an important part of loving relationships is "the possibility of failure or denial." Indeed. It does seem that the feeling of being loved romantically depends crucially on the belief that this love is not unconditional, i.e. that it is contingent on the way you behave, and might disappear. This may seem counter-intuitive; after all, people often *say* they want unconditional love from their partner, or say they love their partners unconditionally. Nevertheless, if one's partner really did love unconditionally, and would put up with any treatment, no matter how badly one behaved, I suspect that the charm would rapidly fade. This point is crucial to my argument in this paper, so let me elaborate on it.

People often say they want their partners to be reliable, faithful, always there for them, never to leave them, and so on. But they want these qualities to be the fruit of an active and ongoing choice. The most effective way to signal that there is a real choice involved here is for the partner to drop hints that there is a genuine possibility that they could leave, if they ever wanted to. So, paradoxically, for people to feel secure that their partners freely choose to be with them, and not with anyone else, they must occasionally be made aware of the partner's freedom by occasional rejections (huffs, moods, and so on), and by the occasional sign that one's partner finds other people attractive too. It can be very painful when one's partner is grumpy, or seems attracted to someone else, but it is also strangely compelling. "Treat them mean to keep them keen", as the saying goes.

The paradoxical dynamics that this complex attitude can engender may play some role in explaining the mixed blessings of marriage. Traditionally, marriage represented the renunciation of any further choice. This promise has perhaps

been honoured more in the breach than in the observance, but it appears that many people continue to believe that their partner is more fully ‘theirs’ when married. If the desire for fidelity in one’s partner were a simple matter, the renunciation of choice signalled by marriage would only serve to make people happier. If, however, the perverse dynamics of mate choice just described apply, marriage should be a more ambiguous benefit, for the gain in reliability would be partly offset by the reduction in the interest that such reliability brings. Which of these two alternatives is more in line with the facts? I know of no scientific research on this matter, but folk wisdom and popular culture provide some anecdotal support for the latter prediction. There is an old saying that “if love is a disease, then marriage is the cure”.<sup>3</sup> In the film *The Witches of Eastwick*, for example, the diabolic figure of Darryl Van Horne points to the same idea in more graphic terms when he exclaims: “Don’t believe in marriage. Good for the man, bad for the woman. She dies. Then the man runs around complaining that he’s fucking a dead woman, and he’s the one who killed her!” It is an old story; a man finds a woman attractive because she is dynamic and independent, and works hard to win her affection. Eventually, he persuades her to marry him. Once they are married however, the dynamism and independence that he previously found attractive are now perceived also as a threat, and he does his best to domesticate her. If he succeeds, however, he complains that his wife is no longer so interesting. This effect is no doubt caused by a whole range of factors, from psychological factors such as habituation to biological mechanisms such as the Coolidge effect, but one of the ingredients in this heady mix is probably the sense that there is less chance that the person will leave, and therefore less to fight for.

While these considerations are certainly far from being decisive, they do lend some support to the intuition that the sense of being freely chosen by one’s partner is an important determinant of the satisfaction that one derives from a relationship. Although people typically want commitment and fidelity from their partners, they want these things to be the fruit of an ongoing choice, rather than inflexible and unreflexive behavior patterns. An obvious objection to Levy’s GST, then, is that robots will not be able to provide this crucial sense of ongoing choice, and that relationships with robots will therefore be less satisfactory than relationships with humans with respect to this feature. Let us call this “*the free-will objection*” (FWO).

---

3. I cannot find the original source of this saying. The earliest variant of it I can find is in the works of the Earl of Chesterfield: “Marriage is the cure of love, and friendship the cure of marriage.” (Chesterfield, 2005: 361).

## A tale of two robots

As already noted, Levy partly anticipates the free-will objection when he states that an important part of loving relationships is “the possibility of failure or denial”. His response to FWO is that partner robots will need to be able to mimic human “capriciousness.” However, this is too quick, and it begs a fundamental question – namely, how far should the mimicry go?

The various answers to this question may be classified in various ways. One way is to divide them into two groups, depending on whether or not the mimicry of capriciousness extends to the point at which the robot can reject its owner permanently. To make these two alternatives more vivid, picture them as two types of partner robot, the RELIABOT and the FREEBOT:

1. With the RELIABOT there is a point at which the mimicry of capriciousness stops; the robot might appear to reject you for a while, but it always relents in the end.
2. With the FREEBOT the mimicry of capriciousness is not bounded; there is always a chance that the robot might reject you permanently.

To see why Levy’s proposed solution to FWO is inadequate, consider the following question: if you were going to buy a partner robot, which model would you choose – the RELIABOT or the FREEBOT?

This dilemma has already been anticipated by connoisseurs of robot sexual fantasies. There is, for example, a website in which users can specify the characteristics of an imaginary fembot that they wish to order. After specifying the ‘type of body’ and the ‘*personality*’ required, the user is asked to specify the ‘control programming’ desired by choosing from the following options (Anonymous, 2008):

1. Total Overriding ( Functions separate from main AI, will force robot to obey to the letter a command regardless of will of primary AI. Not recommended as will cause inefficiency in Unit.)
2. Integrated (Unit will have control programing integrated into primary AI. Unit will see it’s sole purpose is to serve and please you.)
3. Subliminal (Unit will like you, but will not see it’s sole purpose is to serve you, it will however obey any suggestion or command you give it, viewing it as “a good idea”.)
4. None (No control program. WARNING: Not recommended. Androtech will not be held accountable for the outcome if this is selected

The choice here differs slightly from the choice I am considering under the RELIABOT versus FREEBOT headings, but it hints at a similar dilemma. The first three

options correspond (roughly) to types of RELIABOT, while the fourth option is a FREEBOT. Users of the website are in a similar position to that of the consumer who wishes to purchase one of Levy's imaginary partner robots – he or she must choose the level of capriciousness. And here lies the rub. For if we choose the RELIABOT, the considerations advanced in the last section of this paper suggest that we will soon tire of the robot, and come to take it for granted, since it will lose the frisson associated with the contingency of a human Companion. On the other hand, if we choose the FREEBOT we spend our money on an unreliable product. Indeed, the FREEBOT might be a difficult product to sell.

The question of whether the charms of a RELIABOT will rapidly fade is an empirical question, and in the absence of good data about relationships with such robots we can only speculate. Nevertheless, if the observations above about the tendency of people to tire of relationships with people who become too predictable are correct, it is plausible that the same will be true of relationships with robots. There is also the possibility that RELIABOTS could create favourable conditions for the emergence of cruelty. The urge to vent one's frustration on people and objects in the close vicinity is often curtailed by the thought that it may lead someone to retaliate, or (which may be worse) to cease interaction. If RELIABOTS are not allowed to retaliate (a capacity that would raise all sorts of ethical dilemmas that I will not discuss here), there will be no effective sanction open to them for punishing cruelty (since by definition they cannot make credible threats to walk away forever). This is not a positive incentive to cruelty, but the mere absence of a disincentive may be enough to foster cruelty. Again, whether it does in fact have this effect, and how frequently, is an empirical question, but it also has moral implications. For example, if the way that people with robot partners treat those partners influences the way they treat other people, then those with RELIABOT partners might tend to be more cruel (and, more generally, more instrumental) in their relationships with other people. It is easy to see how this might happen. First, some behavior that is unacceptable when directed at a human being might be considered more acceptable when directed at a robot. The similarity of humans and robots might then lead this behavior to become more common in human-human relationships. If so, then perhaps the rather unsettling cartoons by Itsuki Takashi involving "Amputee Robot Girl Bondage" are an ominous sign of things to come (Takashi, 1986).

If the RELIABOT has all these difficulties, then how does the FREEBOT alternative fare? The owner's knowledge that the FREEBOT might reject him or her permanently might reduce or eliminate the risks of boredom and cruelty just discussed, but it is not clear whether this advantage would be worth the price. Imagine if the FREEBOT came with the following warning label:

Rossum's Universal Robots does not guarantee the ongoing compliance of this fembot with your wishes. Bad treatment may cause her to enter permanent non-responsive mode. Resetting her after this will erase all memory of previous interactions.

Would you buy a product that carried such a label? Perhaps you would, if you valued the feeling of being actively chosen that the knowledge of your robot's capacity for permanent rejection might lead to. But consider the following scenario. You have invited a friend for dinner. The friend arrives at your flat and spots your fembot lying on the couch. "What's that?" he asks. "Oh, that's my fembot," you reply. "Why is she just sitting there?" asks your friend. "I treated her badly," you reply; "she's entered permanent nonresponsive mode now." Your friend pauses for moment, and then asks: "How much did you pay for her?" You look a bit sheepish. "About ten thousand bucks," you murmur. Your friend raises his eyebrows. "Wow, that's a lot of money for a useless hunk of metal!" Once again, it is an empirical question as to how many people would buy a product for which the *lack* of a guarantee was supposed to be one its selling points, but my guess is that this would be a hard sell.

Manufacturers may be deterred from producing FREEBOTS for other reasons too. They may, for example, fear being sued on a variety of grounds. Let's say you get attached to your FREEBOT, but one day you lose your temper. As a result, the robot goes into permanent nonresponsive mode. You are afflicted by grief. Can you then sue the robot company for emotional distress?

Other possible disadvantages of FREEBOTS are that they may learn to use the threat of rejection to manipulate their owners. That is, they may come to use emotional blackmail in the same way that people often do.

From this brief (and admittedly speculative) discussion of the pitfalls of both FREEBOTS and RELIABOTS, it seems that FREEBOTS have many of the potential downsides of real human relationships, so may lose much of their potential advantage over human Companions, while RELIABOTS have fewer disadvantages, but thereby lose some of their attractiveness. The problem at the heart of this dilemma does not lie with any technical limitations of the robots themselves, just as it does not lie with some contingent feature of human partners. It lies with our very desires themselves. We want contradictory things: a romantic partner who is both free and who will never leave us. TSA is therefore not correct, for it implies that the reason why human-human relationships often fail to live up to our expectations is that the humans we encounter are so often flawed, when in fact a more fundamental reason is that our expectations themselves are inconsistent. It may be easy to avoid the uncomfortable realisation that our desires themselves are to blame for our frustration, and blame our partners, when we do not have a way

of correcting the supposed flaws we find in our partners. But if robots of the kind that Levy imagines were available, and those who sought solace in robot partners were forced to make a deliberate choice between FREEBOTS and RELIABOTS, it might be harder for them to persist in the illusion, and they might be forced to accept that the source of their dissatisfaction lay partially inside them. Levy presents robot Companions as the best of both worlds. But in fact, they merely make the dilemma inherent in all relationships more visible.

### Love and/or sex

Levy's book is titled *Love and Sex with Robots*. This paper is primarily concerned with the former – that is, with ongoing human-robot relationships involving mutual affection and cohabitation. Sex may or may not be a part of this relationship. Sex with robots outside of such a relationship would be akin to one night stands between humans, or sex between human prostitutes and their (human) clients. *Prima facie*, a version of TSA in support of GST that was restricted to such sexual encounters (outside of long-term romantic relationships) might be more plausible than the version of TSA/GST advanced by Levy, since the purely physical aspects of such encounters may dominate, or completely replace any second-order desires. In other words, the claim that non-romantic sex (“just sex”) with a robot may be more satisfying than non-romantic sex with a human might be more plausible than the claim that romantic relationships with robots may be more satisfying than romantic relationships with humans, given sufficiently advanced robots. TSA may be more applicable with this form of GST because, in this case, TSA can be restricted to purely physical features. To take the most extreme case, the client of a prostitute may have preferences about the prostitute's physical characteristics, but surely he does not have any desire that the prostitute freely chooses him.

Apparently, even this cannot be stated with confidence. There is some evidence that male use of prostitution depends in part on the client entertaining the belief (or fantasy) that the prostitute has freely chosen to have sex with them (indeed, prostitutes try to mimic this) or at least that the prostitute has freely chosen her profession. When regular clients are forced to attend “re-education classes” that explain how pimps recruit, control, and exploit women and girls for profit, and learn about the links between local street prostitution and larger systems of human trafficking, these men tend to reduce their use of prostitutes (Shively et al., 2008). This suggests that many clients not only believe that prostitutes choose to have sex with them, but also want this to be the case. So it seems that even here, where it is plain to most people that the prostitute does not choose

her clients out of a desire to be with them, the clients prefer to fantasise that they have been chosen, and find that their enjoyment of sex with prostitutes diminishes if they cannot maintain this fantasy.

It is hard to say how this observation bears on the '*just sex*' form of GST. It could point either way. On the one hand, it might suggest that the objection to TSA outlined above would apply equally here, since people's enjoyment of sex, like their enjoyment of romance, involves second-order desires as well as first-order desires. On the other hand, it might suggest that it is remarkably easy for people to fool themselves into thinking that they have been freely chosen by their partner when in fact it should be clear that the partner has had very little, if any, choice in the matter. If this is true, then perhaps the dilemma outlined earlier is vastly overstated, and people will be very satisfied with RELIABOTS after all, since they will easily fool themselves into thinking that RELIABOTS have freely chosen to be with their owners. In this way, people really would get the best of both worlds – romantic partners who (appear to) freely choose them, and yet will never leave. People would gain this delight, of course, only at the price of deluding themselves.

## Conclusion

The 'greater satisfaction thesis' (GST) states that romantic relationships with robot Companions might be more satisfying than relationships with humans. One argument that has been advanced in support of GST by David Levy is that people will be able to specify the features of robot Companions precisely in accordance with their wishes (which I call the "total specification argument", or TSA). In this paper, I have argued that TSA is wrong, because it breaks down when we consider certain behavioral characteristics that we desire in our partners. We typically want our partners to choose us freely and therefore to be capable of rejecting us permanently. I call this the "free will objection" (FWO) to TSA.

FWO is not decisive against TSA – it suggests difficulties with it, no more. Ultimately, GST is an empirical hypothesis which may become more or less likely as robots become more advanced, and we gather more evidence about human-robot relationships.

Even if my argument against FWO turns out to be right, this would not show that GST is wrong. It would merely demolish one particular argument for GST. TSA could be wrong, and GST might still be correct. In particular, the version of TSA advanced by David Levy is intertwined with a series of other assumptions about the social context of partner robots (they are made in a factory according to specifications provided by the user; they are consumer products designed to serve

the needs and wishes of humans; and are private property). In a different social context, TSA might not even apply. For example, if robots made copies of themselves, for their own purposes, and were not the property of anyone, and were sufficiently advanced, they might enter into relationships with humans of their own free will. For all we know, it might well be the case that many humans would find relationships with these robots more satisfying than relationships with other people. However, this is not the kind of situation envisaged by Levy.

This consideration suggests that predictions about human-robot relationships always need to be considered as more-or-less complete scenarios, in which a whole set of hypotheses is advanced jointly. These hypotheses always include auxiliary assumptions about the social and legal status of the robots in question, assumptions about the level of technological advancement, and other matters. The plausibility of any prediction about human-robot relationships can only be assessed in the context of the general scenario within which it is advanced. This suggests that the scholarly and academic treatments of this subject (such as Levy's) have much in common with fictional treatments (such as those of Philip K. Dick and Isaac Asimov). Indeed, the fictional treatments may often be superior, since they often take more care to make their assumptions about the social context explicit, and take more care to make the story internally consistent, and externally consistent with what we know about human nature

Uncorrected proofs - © John Baughman/Philosophy.com

