Simulation and Servitude: The Ontology of Inauthenticity and the Risk of Moral Corrosion in Human-Robot Relationships

By Faraz Fookeer

Introduction

This essay discusses, firstly, that human-robot relationships are inherently inauthentic as there is a necessary absence of mutuality and “interlocking” attitudes and, secondly, that human-robot relationship could potentially corrode inter-human relationships along three paradigms: the objectification another human, the corrosion of the internal machinations of the mind dealing with interpersonal relationships and the corrosion of inter-human relationships at a societal level due to the erosion of political agency. As a counterbalance, this essay looks briefly at harm-reduction frameworks that could be implemented to mitigate the corrosion of inter-human relationships.

The rapid rise and propagation of sophisticated robots that mimic aspects of human behaviour within the field of social robotics, exemplified by products such as “RealDoll” and enhanced with generative artificial intelligence (GenAI) and virtual reality (VR) features, has brought about a crisis in the ontology of friendship and the ethics of intimacy. By engineering simulated human-like responsiveness features, manufacturers of these technologies aim to “create a genuine bond between man and machine” (Atanasoski and Vora, 2019, p. 189). It follows then that it is only natural and logical to analyse and interrogate the nature of this bond and its potential effects on the individual and society at large.

The Ontology of Inauthenticity

To evaluate the claim that human-robot relationships are “inauthentic”, we must establish what constitutes the conditions of an “authentic” relationship. To do this, we turn to Turing’s Imitation Game (Turing, 1950) as the framework. The Imitation Game was Turing’s proposal to replace the abstract question “Can Machines Think?”, stating that this question was “too meaningless to deserve discussion” (Turing, 1950, p.10). The Imitation Game suggested that if a machine played the part of a human convincingly well then the machine has functionally succeeded in demonstrating intelligence (Turing, 1950, p. 433). Turing anticipated objections, that a machine cannot

feel human-like emotions for example, pleasure, grief or be “warmed by flattery” (Turing, 1950, p.446). As a rebuttal, he continued a functionalist stance: we only know how others think and feel by their external behaviour and, hence, can have an understanding of an “authentic” relationship, unless one adopts a solipsistic view (Turing, 1950, p. 446).

Turing’s functionalism may be enough for intelligence or logic-based activities, but it fails in the sphere of intimacy and friendship. In this sphere, a simulated feeling can be equivalent to deception, but a simulated solution to a math problem or a virtual chess move is still correct. An “authentic relationship” is thus defined as an interaction based on mutual subjectivity and reciprocal affection, different from those based on simulation, projection or using the other as a means and not an end (Kant, 1785). This interaction is not bound by time since one may experience an authentic relationship with another during the most fleeting of moments. Exploring this definition further, Sterri and Earp (2021) argue for a standard of relationship based on a union that reflects “reciprocal love and affection” between two parties. This excludes relationships where one party is incapable of genuine feeling. Nyholm (2022) takes this further by proposing the concept of “mutuality”. This is a shared mental existence, necessitating a “mind-to-mind” connection where parties are open to each other’s “wishes, desires, likes, dislikes”, and without this mental connection, a relationship could become objectifying and contractual. These definitions of an “authentic relationship” taken from the literature echo the Aristotelian conception of friendship, which requires mutual well-wishing for the friend’s sake (Salomone-Sehr, 2025, slide 7). With this definition as a backdrop, any interaction with a machine - especially modern AI systems - offers merely an illusion of a partner, thus failing the test of authenticity.

Regardless of their human-like conversational fluidity, current modern large language model (LLM) based AI applications (also embedded within modern sex robots) fundamentally lack the internal mental states needed to forge genuine and lasting connections. These systems do not value the user; instead, they are designed with intentions embedded within incentive structures embedded within surveillance capitalism, offering a “simulation of reciprocity” and are built to provide the “illusion of enjoyment” (Atanasoski and Vora, 2019, p. 190). Their makers, specifically the Big Tech firms of OpenAI, Meta, Alphabet and Microsoft, design their respective LLM user interfaces (UIs) to garner maximum attention with the hope of building a network effect and, eventually, an ecosystem around these tools from which the user would find it hard to escape. The user may believe that they are in a relationship with the application but, contrarily, they are engaging in a solitary projection; a solipsistic relationship akin to that of Tom Hanks’ character and Wilson the volleyball (Cast Away, 2000).

This fundamental lack of mental states or internal subjectivity within these machines creates an ontological deficit, rendering any human interaction with the machine one-sided. Into this ontological deficit steps Nyholm (2022) who reinforces this gap by stating that the “tenderness and mutuality” view of sexual relationships requires recognising a partner’s subjectivity which is impossible given that the machine “lacks a humanlike mind” (Nyholm, 2022, p. 18). Since an LLM-based AI has no subjective internal states, it is impossible for a human interlocutor to have a genuine relationship with it. The user is, in effect, “alone together’ with the machine; interacting with a (often sycophantic) mirror that reflects their own wishes and desires, skilfully masking the fact that this relationship is ontologically inauthentic: it is solitary and solipsistic.

While human-machine relationships are deemed ontologically inauthentic given today’s technological landscape, could this change in the future? The conditions for genuine mutuality and shared phenomenology could theoretically be attained by speculative future technologies, specifically Artificial General Intelligence (AGI) - assuming that AGI systems attain machine sentience. Such a system would dissolve the barrier to “mind-to-mind mutuality” (Nyholm, 2022, p. 18). Nyholm’s (2022) framework of “mutuality and generosity” would be bolstered by a conscious AGI as it would genuinely value the user, rather than the current state of engaging with the user to reinforce its learning algorithms, harvest data, or exploit psychological weaknesses on behalf of its designers. One would hope that an AGI of this kind would allow us to transcend the solipsistic projection of the user, thus achieving Aristotelian authenticity.

The Corrosion of Inter-Human Relationships

We have now established that human-machine (AI) relationships are inauthentic from an ontological perspective due to a lack of mutual mental states. Now we address the second discussionon point: that these relationships (assuming widespread adoption of the technology) threaten to corrode the quality of inter-human relationships. If we adopt the stance that human users are essentially “alone together” with these AI systems, does this solitary relationship modality constitute a harm? The literature suggests three levels of potential human-human corrosion: objectification due to the reinforcement of harmful societal norms (especially towards the more traditional, “vulnerable” actors within society), the adverse effects on one’s character due to the erosion of personal virtues, and the embedding of a “technoliberal” desire for domination.

Objectification

To support and understand the objectification argument, consider the potential widespread usage of sophisticated GenAI-driven sex robots and their effects on societal norms. The production and distribution of these machines do not occur in a vacuum and

they may “undermine important moral norms that regulate sex between humans” (Sterri and Earp, 2021). This is particularlylarly worrying when it comes to the representation of women. Of the three types of sex robots that are of potential moral concern, the female version currently mirrors a “strong Eurocentric male gaze”. It represents a physical amalgam of the female body type most commonly associated with Eurocentric/white pornography (Sterri and Earp, 2021). They are designed to be submissive with limited “intelligence”.

The threat to women is that already-skewed power relations that view women as objects are indirectly exacerbated via the potential widespread usage of such sex robots. Men who use these machines may become used to the concept of consent fading into the ether as a precursor to the sexual act. The robot may be programmed to offer a soft facsimile of consent but, as it exists to pleasure its “master”, the concept of “no” may lose all weight. As the men in question return to “normal” life outside of the bedroom, one would expect these users to view sexual consent with the opposite sex as irrelevant over time. Sex robots eliminate the necessity for mutual respect and compromise, affecting the very ability for a man to understand (and act on) verbal consent and detect non-verbal cues from the opposite sex. If robots are designed to resist (simulate rape), a person could be so fully immersed in this act that the thread between their fantasy and reality dissolves (Sterri and Earp, 2021). Intra-human relationships, specifically male-female intersexual relationships, could become corroded by virtue of the erosion of the concept of sexual consent and the elevation of the idea that women are “ever-consenting” sexual objects; objects that exist eternally to satisfy the male gaze.

Internal Corrosion

Following on from objectification as a potential cause of inter-human relationship degradation, we now turn our attention to how human-robot relationships affect our internal states and how this, in turn, manifests itself in our outward dealings with other human beings. Nyholm (2022) analyses and engages with Sparrow’s (2017) argument that how we treat human-like entities reflects our nature and character. On the topic of sex robots that are built to simulate non-consent, Nyholm states, “performing a sex act on a robot that is programmed to appear to not consent is a representation of rape", and that this is wrong because "it is an expression of bad character". While we cannot harm a robot, treating a humanoid figure with cruelty may exacerbate vicious dormant traits such as intemperance or cruelty. However, this argument goes deeper than the activity of expressing cultivated bad character traits in one’s interpersonal human relationships, where the greater threat may come from the gradual degradation of virtuous behaviour that is required to forge an authentic connection. When interacting with a robot, one can exercise “temperence” - the act of self-restraint - whereas “tenderness” cannot be enacted due to a lack of reciprocation (Nyholm, 2022). Tenderness is inherently

interpersonal and requires sensitivity to the “wishes, desires, likes, dislikes” of another human being with a mind. A machine lacks a humanlike mind thus the user cannot practice being tender with another entity. This act is performative at best.

Consequently, if humans were to adopt the widespread usage of sophisticated humanoid robots, an over-reliance on this type of companionship could de-skill humans and further corrode our interpersonal relationships. People would lose the ability to wrangle and navigate the difficult and often frustrating emotional work required to maintain a connection with another human. If users continue to interact deeply with partners that require no emotional response and compromise, they may find the reality of human relationships intolerable. This could create a feedback loop where inter-human relationships become increasingly unbearable (due to the lack of “relationship practice” by the concerned individual) and the user retreats further into their relationship with the machine, continuing ad infinitum.

Political Corrosion

Atanasoski and Vora’s (2019) concept of “technoliberal desire” exposes one of the least discussed but potentially most profound threats to inter-human connections, posed by human-robot relationships. This is the corrosion of the political framework that allows the incubation of genuine inter-human relationships, in the first place. Their argument moves away from the microcosms of objectification and the internal machinations of the human mind to the type of subject that these technologies cultivate: sex robots and other humanoid-like machines are not passive tools - they shape us just as much as we operate them.

At the very heart of their argument is the positing that the desire for these machines represents a “technoliberal update” to a history steeped in racial and colonial overtones, where certain demographics of humans were treated as property (Atanasoski and Vora, 2019). These humans were expected to work, serve and submit without a modicum of political freedom. They lived outside of the political framework that allowed a free society to exist. Modern sex robots simulate this dynamic in a new fashion: they are built to mimic just the right amount of human characteristics to feel lifelike but they remain entirely nonautonomous (Atanasoski and Vora, 2019) - structurally incapable of refusing the whims of their owners. Manufacturers may look to bridge the gap between a thin sheen of lifelike human features (for example, by adding better orgasm response, synthetic skin, warmth and the like) and sex robot’s mechanical nature. However, this does not change the fact that the sex robot’s underlying political agency (structure) remains unchanged. It is denied the moral and political autonomy that humans owe and bestow upon each other.

This is the heart of the political corrosion of inter-human relationships. These machines do not solely reflect a user's desires; they train those desires. The user experiences a "freedom" that is at the expense of a rightless being. Over time, and as the machine becomes more sophisticated and widely adopted, the risk is that a populace will construct a subjectivity where democratic equality will start to feel inconvenient. The fundamental idea that humans are moral equals who can refuse or say "no" may even become undesirable. The corrosion is not that people might prefer human-robot relationships to interpersonal human relationships, but that they would be "trained" by these machines to prefer relationships that, structurally, lack mutual reciprocity, resistance or agency.

Harm Reduction

Although the arguments made in this essay are pessimistic in nature, it is important to provide a counterweight to the idea that these technologies (sex robots, LLMs and the like) will inevitably corrode inter-human relationships. This is perhaps too deterministic, and there is a saving grace in the "harm reduction" framework (Sterri and Earp, 2021). This framework suggests that while the risks of corrosion are very real, outright prohibition may not be the most ethical response.
For certain sections of society such as the socially isolated, the disabled or those with severe anxiety sex robots may provide a viable and valuable alternative to total isolation. Instead of corroding interpersonal relationships, these machines may serve society in a therapeutic capacity. They may act as a bridge to genuine human intimacy, walked on by those who cannot attain such a closeness in their current states. Controversially, sex robots specifically designed to simulate children may be used to treat paedophilia, thus replacing a horrendous and harmful societal act rather than encouraging it. The corrosion of human relationships is not an intrinsic property of the hardware but a potential outcome of the norms encoded into it.

If we adopt a "harm reduction" approach that prioritizes therapeutic use and regulates against the most "vicious" designs (such as rape simulations), the technology might yet find an ethical place in the human social sphere. The potential corrosion of human relationships is not baked into the hardware of humanoid robots (or the LLMs that may be incorporated into them), but a potential outcome of the norms that humans encode into these systems during the design phases of their construction.

Conclusion

Human–robot relationships, as they exist today, cannot meet the conditions of authenticity demanded by philosophical accounts of intimacy, friendship, or mutual

recognition. The absence of shared subjectivity, reciprocal affection, or “mind-to-mind” mutuality renders these interactions ontologically one-sided and fundamentally solipsistic. Yet the inauthenticity of these relationships does not, on its own, determine their moral or social impact. The real risk emerges when technologically engineered simulations of intimacy are fused with broader political, economic, and historical dynamics. As the literature shows, the design of sex robots and AI companions can entrench patterns of objectification, corrode the cultivation of interpersonal virtues, and, via technoliberal desire, potentially train users to prefer relationships that lack any form of reciprocity or agency. This form of political corrosion poses the most far-reaching threat to the fundamentals of democratic equality on which human relationships depend.

However, none of these outcomes is inevitable and thinking so is too deterministic. A harm-reduction approach, which regulates the most ethically dangerous forms of design while permitting therapeutic or socially beneficial uses, suggests that human–robot relationships need not corrode human relationships by default. Ultimately, the ethical path of these technologies will be determined less by their artificiality than by the political, economic, and normative choices embedded in them.

References

Atanasoski, N. and Vora, K. (2019). 'Epilogue: On Technoliberal Desire, Or Why There Is No Such Thing as a Feminist AI', in Surrogate Humanity: Race, Robots, and the Politics of Technological Futures. Durham: Duke University Press, pp. 189–196.

Nyholm, S. (2022). 'The Ethics of Humanoid Sex Robots', in L. Watson, C. Chambers and B. D. Earp (eds.) The Routledge Handbook of Philosophy of Sex and Sexuality. London: Routledge, pp. 574–585.

Salomone-Sehr, J. (n.d.). 'Lecture 8: Ethics of AI friends and lovers'. [Lecture slides]. Birkbeck, University of London.

Sterri, A. B. and Earp, B. D. (2021). 'The Ethics of Sex Robots', in C. Véliz (ed.) The Oxford Handbook of Digital Ethics. Oxford: Oxford University Press, pp. 241–257.

Turing, A. M. (1950). 'Computing Machinery and Intelligence', Mind, 59(236), pp. 433–460.