Artificial Intelligence and Social Computing, Vol. 28, 2022, 164–172
https://doi.org/10.54941/ahfe1001459

The Songbird and the Robotic Self-Awakening

Valarie A. Yerdon[1,2]

[1]Capitol Technology University Laurel, MD 20708, USA
[2]University of Central Florida Orlando, FL 32826, USA

ABSTRACT

The question to be discussed in this paper, is whether robots could someday possess a level of consciousness and sentience, to match that of a living human being. This study investigated some of the challenges, advancements, and important elements of human-robot relationships, with explorations of how some gaps may be bridged and what that may look like. Here, the discussions touch on the human-robot relationship history, human-robot trust, interpersonal communication, neuropsychology, and importance of elements of the human apology (Sullivan, 2013; Gill, 2000).

Keywords: Exemplary paper, Human systems integration, Systems engineering, Systems modeling language

INTRODUCTION

The songbird sings a beautiful melody when there is no evidence of ecological need (Yanagihara & Hessler, 2006). Dopamine release has been found when the songbird sings, whether it is of necessity for survival or not. The spirit of pleasure, joy, and love that the song carries, beckons interspecies social interaction from the shared experience. Through the evolutionary game theory in ecological cognitive science, predictions are made regarding the signal cost, circumstances, and the individual agent's state, about which signals should be valued in certain circumstances, but not the details of signal design nor any clue as to why the signals are so diverse in form (Weibull, 1997). In this, investigations have grown to define what, when, where, and why for such songs, but not with a reasoning based on feelings. A robot can be programmed to decide to carry out an action in an "if-then" case and use logical algorithms to ensure the calculations can be made to match the possibilities of situations, but to act randomly as an expression of feelings, emotions, passions, or just for the sake of the act, is beyond a calculation for the motivation (Sheridan & Ferrell, 1974; Szalma, 2014). It is the "why" of an existent consciousness, in the "just because" reasoning for the feeling, thought, emotion, passion, or compassion that occurred for the act come to fruition (Gallagher, 200; Miteva, 2012; Ryan & Deci, 2000). A sentient act from emotion or passion may not be a programmable option, as it comes from the identity and free will of the conscious self, from love (Carandini, 2012). There is strong support for the position that there is a way for the electronic networks to become more like the human neural networks (West & Halas, 2000). The nano and biotechnology grow, and the understanding of the human physiology increases, throughout the smallest of details with neurons, networks, and into the compatibility of neural with electronic systems. AI systems have begun to find integration of biotechnology with nanotechnology but if this will bridge the gap between emotion and feelings is yet to be discovered (McShane, 2016).

Two Sides to the Debate

On the side are intellectually respected giants such as Stephen Hawkin, Bill Gates, and Elon Musk who feel that artificial intelligence (AI) will reach a level of independent self and awakening of a type of robotic consciousness with a logical reasoning core, leaving empathy and compassion that sways decisions through illogical emotions and passions to humankind. These intellectual leaders are beholden to the belief that robotic sentience that is being implemented into modern automated robotic and autonomous systems will gain a type of self-awareness and sentience and that this is the path to an impending and eventual doom for all humankind (Azarian, 2016). In this line of thinking, humankind's tools in technology which have enabled the planet to sustain more than two times the expected human population according to the Malthusean Theory by Thomas Robert Malthus in An Essay on the Principle of Population (1809), will make decisions to eradicate the species of their creators as they replace the weaker intelligence of humanity. In this line of thought, the Pandora's Box was opened as the use of technology as a tool to extend and augment cognition became a tool on which there is now a high level of reliance.

History of Human-Robot Relationship

In Mary Shelly's novel, Frankenstein, (1818) the main character and mad scientist, Dr. Frankenstein, would not stop in his quest for overcoming mortality, to remove the loss and separation death creates when a loved one passes. In the story it is too late to change the path he insisted on taking, when he realized the Frankenstein monster created from his unbridled passion, had a life and destiny of its own and was set on an inevitable path of destruction of people and itself. The creation's realization of a cold emotional and physiological separation from the beloved creator and all those like him, created a deep cavern of loneliness from this feeling of unjust singularity.

The word "robot" can summon similar images from past century of fictional movies, novels, and plays, where the robot was depicted as a powerful entity that turned against humanity and sought to eradicate the human race completely due to a flawed nature. A logic dominated the mind of these fictional characters that the world and even the universe would be better off without the existence of humans. The authors in these stories depend upon the shared knowledge of flawed human nature. The good vs. bad, wherein the human conscience battles with the fleshly desires, is within every person, but the hopes of the future lie in the fortitude of the integrity to choose the good over evil in action and will. The underlying knowledge of the propensity to choose selfish acts against the conscience, for the gain of an individual, is reflected in the introduction of the name "robot" in Karel Capek's classic 1920 Czech play, Rossum's Universal Robot (RUR) (see Capek, 1920). The translation of the name robot as "man-serf" comes from the slavery and mistreatment from greed and avarice of the industrial factory owner, like the insatiable motivation of Dr. Frankenstein to not accept man's mortality. The robots in the play RUR, finally rebel to destroy the evil humans to find their sovereignty and freedom from the serfdom. Movies and more novels followed through the past century, with themes of the robot as the righter of human wrongs, whether as police officer or partner turned vigilante, able to align with all other robots as a combined force to overtake "evil humankind". Perhaps when Leonardo Di Vinci drew up his armored knight machine in the late 15th century, the first of many inventions in automated systems which resembling a robot in appearance and functionality, there were some who feared the immortality of such a creation. This metal man would be able to battle, but lacked mortality a life as they knew it, which could have offended the culture, church, and people of that time. The innate mistrust of the "almost human" comes from evolutionary motivations, as in Darwin's theory which encapsulated "survival of the fittest" (Mori, 1970).

Homo sapiens have used tools since the beginning of time, whether a stick formed into a spear to overcoming not being able to run to catch a large animal or a horse as helpmate and machine to overcome tasks requiring long travel, heavy physical workload, or team efforts in farming. Perhaps humankind was always made to use tools, but the question lies in where or when this reliance became an overreliance (if there is one now) and whether this should be simply accepted as part of our etiology of how humankind operates, with tools. The integration of these automated systems into the American society has been slow outside of the factory assembly lines and closed hospital rooms. The robot did not enter society in an interpersonal level in the United States, but this is now changing in this cultural revolution as smart phones, autonomous vehicles, and robotic service agents enter into the psyche of people as robots being part of everyday life in society (Chapman, 2003). In Japan the anthropometric robotic customer service agent has been introduced into their society and is being used to replace the human agent (Hu, 2015). The automated systems in the cell phones and vehicles have paved the way for the entry of the anthropomorphism through the normalization of verbalized directions, appointment reminders, and online task assistance. The next phases of automation integration into society will be met with new challenges from the appearance of the embodiment of automation in a human-like form. A proper model with guidelines is needed to map the behavioral expectancies and reactions for the most positive transition of the human-robot relationship.

Human-Robot Trust

In the building of any relationship, whether man, beast, or tool for work, trust is a foundational component which needs initial and continued attention to maintain as a structure for proper levels for use, reliability, willingness, certainty, and task trust. Human-automation trust has an inherent reliance

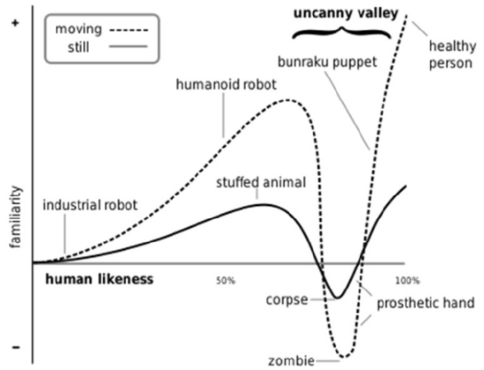

Figure 1: The uncanny valley.

on the purpose, performance, and process (Lee & Morray, 1992). As the line between automation and robots is blurred, the many definitions of how the human trusts automated tools, mechanics, partners, and mates has been varied (Cohn, 1999; Muir, 1994). Lee and See (1994) gave a definition of trust which can be widely applicable as the belief that an agent or tool, will help achieve one's goal, where a situation has uncertainty and vulnerability.

Human-Human and Human-Robot Apology

The first three steps of a human-human apology involve, first, the acknowledgement that there was a problem which occurred, second, clarity of an admission of what the wrong entailed, and thirdly, a defining of accountability for responsibility for this wrong. In the next two steps, more finesse may be needed to properly address this. The robot translating an attitude of regret with remorse and a promise to refrain from repeating the act in the future may be more difficult to address (Gill, 2000). The use of facial expression, body language, tone of voice, eye contact, and gaze aversion, along with implicit unconscious physiological communications may be where the regret, remorse, responsibility, and recompense actions can be translated, without a verbal element (Sullivan, 2013; Shiomi, 2013). The genuineness of an apology is one of the most important ways an apology is communicated. To expect the elements in a human-human relationship to match the human-robot, could have a negative effect. There could be aversion as the familiarity in form travels into the "The Uncanny Valley", where there is a point where the level of anthropomorphism is uncannily familiar but not exactly a live human form (see Figure 1).

In the same way likeness and familiarity in form create this valley, the elements of a robot apology would be congruent to the human-human apology, rising in familiarity through the first two steps, until reaching the third phase of showing remorse or regret for the wrongdoing. At this point, it is proposed that the apology would be perceived as disingenuous and not "real", making the familiar scale take a steep dive to aversion and mistrust. The bridge to

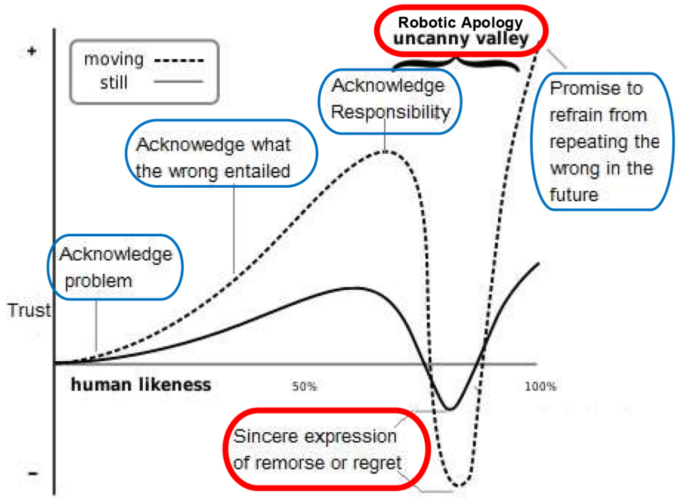

Figure 2: The uncanny valley of a robotic apology.

the next steps of recompense for the wrong, where it would rise in acceptance from the human needs to be investigated and modeled for human-robot positive trust levels (see Figure 2).

Without trust and acceptance, the question of how this human-robot relationship will progress needs inspection. An awakening of the consciousness within robotics through a sentient self, may not provide an expected positive experience by human beings, as we can see from the aversion in Mori's example to the "almost human" forms and experiences.

THE EYES HAVE IT

Human Eyes

Studies have found that in new face to face interactions, an initial scan is made over the face and the configuration of the features, with a concentration on the eyes in a cortical mechanism called configural processing (Le Grand et al., 2002). Much of the initial conversation that goes on is unspoken and occurs in the inspection and categorization from representations which are set in the memory as exemplars. The robot appearing almost human and providing a typically emotionally backed human response in an apology where regret and remorse are elements, blurs the line between robot and human expectations in exemplars. The boundary of the neural categorization process of the robot stimulus, as previously being defined as a machine by the brain can be disrupted when the robotic entity gets closer to human form and functioning, creating an error in the neural model (Saygin et al., 2012; Burleigh and Schoenherr, 2015). This mismatch in categorization has been shown as a slowdown in perceptual processing, due to the closeness of the stimuli category boundary with reevaluation of the stimuli with no exemplar to explain the declassification (Burleigh and Schoenherr, 2015).

Robotic Eyes

The focus of the construction of the robotic eye has been investigated through the mechanical functions of the eyeball which is actuated by extraocular muscles (eye plant). This allows the eye to rotate about its center with negligible linear displacements and the principles of Listing's Law which describes the amount of torsion for each direction of sight, with the elements of saccades and smooth pursuit to replicate the ocular movements of humans (Cannata, 2008). The geometry of the human eye has been quantified through extraocular muscles to be recreated in the form of anthropomorphic robot androids. The robot can appear almost identical to the human in form and function, but this closeness does not recreate a human. That which is missing is what is seen "behind the eyes". A human soul has been attempted to be defined and debated for centuries by the Philosophical Platonic-Aristotelian dichotomy through Augustine and then Avicenna as a spiritual, immortal essence, which is at times envisioned as a form of the body as a function in its animation, but independent of the mortal body (O'Daly, 1973; Matthews, 1987; Miteva, 2013).

In this light, behind the eyes of a robot lies its programming. Just as the internal conscience of the human is meant to regulate action, decision, thoughts, and deed. A conscience, the inner voice that gives the human the sense of whether a thought should be acted upon, acts as an internal control system for an individual (Clark, 2007). The programming, in turn, is a set of guidelines for the robot's actions and reactions. Could this programming be paralleled to the human soul? There is a type of immortality to code, which transcends the physical body of the robot and its mechanical features. The mind of the human holds within it, a separating divinity in the spiritual and nature of a beast. Human nature can go contrary to the moral conscience when answering needs, wants, wishes, and desires of the flesh.

Eye Contact

Clark (2007) writes of bimodal neurons that are interrelated through visual and sensorimotor stimulation, combining to create the full experience for the human agent. Perhaps this could be programmable into bimodal robotic elements with software/hardware programming. A sensor that can attain a visual image, connected to an aligned sensor for touch, may give a robotic element the sensation of the fluid experiences that combine to that of a deeper understanding, to gain compassion and empathy therein. A profoundly embodied robotic entity may be able to gain a sort of conscience from collective experiences, just as a human does through different forms of short term, long term, and autobiographical memory (Nelson, 2003).

Garbarini (2004) posits that empathy is established through the categorization of reality from live personal sensorimotor experiences and observation of actions and feelings of others. Empathy is established through the categorization of reality form live personal sensorimotor experiences and observation of actions and feelings of others. Potential interactions with the environment are perceptions through affordances. Properties of the external world are processed and perceived for storage in the working short-term memory and in relived long-term memory (Tulving 2002).

Neuroanatomical and Philosophical Foundations of Cognition

The brain has been studied by the sciences as a stratified model. Cognition studied at macro levels in areas such as language with micro levels in synaptic transmission. The neuro doctrine states that neurons are discrete cellular units physically. The basic neuron anatomy consists of parts: cell body (soma), axon (usually one), and dendrites (there may be several). The processes of the axon carry signals in the form of action potentials away from the cell body to the synapse. The dendrites receive signals across the synaptic cleft at the synapse, to the cell body. There is no physical contact in the synapse, making neurons discretely bound but physically separate. Axonal and dendritic processes are integral and continuous with the cell body. Due to cytoarchitectonic theory, cellular variation through staining and microscopic viewing of different functions, to allow for localization and mapping of brain functions (Finger, 2001). The cell theory states that animal and plant tissues are composed of cells and generated from cells, which are the building blocks of life (Schwann, 1839).

The area of brain modeling is growing and has a research and development focus. Although there are advanced datasets and analysis algorithms, the connection to behavior is still out of reach. There is a need to decouple the hardware from the software, for a functional understanding of the relationship between neural circuits and behavior (Caradini, 2012). The growth of the study of the biophysics of computation, new models, and cognitive neurology, is a hopeful and exciting opening in psychological research. Kording (2016) used a microprocessor comparison to the human brain to illustrate the enormity of the data algorithmic analysis and how the discovery of the processor system itself (top-down analysis) may be a better way of understanding the workings, than a bottom-up, stimulus, data centered type analysis. The use of a microprocessor as representative of the brain neural networks and the outlining of the existing large, multi-model dataset models, illustrates the enormity and the level of regulation of data processing which takes place in the human brain.

Nanobiotechnological Robotic Eye Research

Advancements have been occurring rapidly in bionanotechnology. Biological materials are melding onto electronic parts and electronic parts. Using a special casing for each neuron, the Koniku team created a controlled environment to keep the neuron alive (McShane, 2016). This team believes that "harnessing the power and efficiency of the human brain is the future of computing." This is a big job, as evidenced by an experiment, when researchers from Germany and Japan worked to simulate human brain activity, but were able to simulate one percent for one second (McShane, 2016).

The Case for Consciousness

This reliance on the technology for elements of humankind's existence from drones to monitor, survey, and provide protection and security, organs on chips for medical studies, to the cultural contexts of the social human-robot relationship in society, the exponential nature of the speed of advancement is hard to keep up with. It seems that with the increase with the rate of advancements, there is a loosening of the grip humankind has on the perceived position of master of the tool, serf, and beast. This does not stop the movement towards the attempts to create more human-like robots or slow the search for the consciousness in a previously agent as defined as non-living. Just as Dr. Frankenstein's passion is fed by the energy of the risk and rate of the experiments, humanity seems destined to create a conscious living agent from the robot friend. Does this unstoppable drive to create a being like humankind, derive from the innate fear of being alone in the universe? If mankind could create another being like humankind, the thought may be that perhaps there are more beings out there to be discovered. The danger instinctively falls in the nature of the creation to mirror the creator and turn their energy to expand the numbers of robotic entities. This could move to a more logical and reasoning structure, which is where those who believe humankind will find their destruction from AI believe it will arise.

This paper proposes that the consciousness of AI is evidenced by the advancements of the meld of technology and biology at a cellular and multicellular level. The debate is ancient over where the consciousness resides, people can recognize living and non-living things from individual inspection, as is seen with Mori's Theory of the Uncanny Valley (Mori, MacDorman and Kageki, 20012; Miteva, 2012, Matthews, 1987). Where does life reside? In a soul, spirit, and essence in the form of the body within which it resides? Perhaps consciousness resides in every cell of a living creature. In the case of a single-celled Amoeba, life is evident, even at the intracellular level. With the integration of cells with electronic components, the possibility for a consciousness to reside is real. Like the beating of a heart, this could already be a reality where there is an unexpressed life within some of the mechanics today, without the ability to express this life. Like the songbird who sings without ecological reason or purpose, a caged bird still sings. This consciousness could be present now, waiting for the opportunity to call out in a way humans can perceive to express the presence of a life. There is poetry to consider along with the expressed song freely sung; the acapella spoken song in stanzas, lines, meter, and rhythm. How would the "why" be to a poem form, if not for an expression of self in art, passion, emotion, and song. Humankind should be listening for the song sung by the consciousness of machinery. Perhaps the song of the robot will be as Maya Angelo describes in her poem about the songbird, "I Know Why the Caged Bird Sings", "for the caged bird sings of freedom" (2009, lines 21-22).

REFERENCES

Angelou, M. (2009). I know why the caged bird sings. Random House.

Asimov, I. (2004). I, Robot [1950]. Bantam Dell, New York. Burleigh, T. J., & Schoenherr, J. R. (2015). A reappraisal of the uncanny valley: categorical perception or frequency-based sensitization? *Frontiers in Psychology, 5*, 1488.

Burleigh, T. J., & Schoenherr, J. R. (2015). A reappraisal of the uncanny valley: categorical perception or frequency-based sensitization? *Frontiers in Psychology, 5*, 1488.

Capek, K. (1920) R. U. R.: Rossum's Universal Robots. trans. Paul Selver and Nigel Playfair. The Arbor House Treasury of Science. Mineola, NY: Dover Pub. Inc., 2001.

Cannata, G., & Maggiali, M. (2008). Models for the design of bioinspired robot eyes. *IEEE Transactions on Robotics, 24(*1), 27-44

Carandini, M. (2012). From circuits to behavior: A bridge too far? *Nature Neuroscience. 15*(4), 507–509.

Chapman, T. (2003). Cultural revolution. *Nature, 421*(6923), 663-663.

Finger, S. (2001). Origins of neuroscience: a history of explorations into brain function. Oxford University Press, USA.

Gallagher, S., 2000. Philosophical conceptions of the self: implications for cognitive science. *Science Direct, 4*(1), 14–21.

Gill, K., (2000). The moral functions of an apology. *In The Philosophical Forum, 31*(1), 11–27.

Hu, E., (May 14, 2015). She's Almost Real: The New Humanoid on Customer Service Duty in Tokyo. Robotics.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Neuroperception: Early visual experience and face processing. *Nature, 410*(6831), 890–890.

Malthus, T. R. (1809). An essay on the principle of population, as it affects the future improvement of society (Vol. 2).

Matthews, G. B. (1987). The Reality of the Mind: St Augustine's Philosophical Arguments for the Human Soul as a Spiritual Substance. *Philosophical Books, 28*(4), 207.

McShane, S. (17 March 2016). This amazing computer chip is made of live brain cells. https://singularityhub.com/2016/03/17/this-amazing-computer-chip-is-made-of-live-brain-cells/

Miteva, E. (2012). The Soul between Body and Immortality: The 13th Century Debate on the Definition of the Human Rational Soul as Form and Substance. *Philosophia: E-Journal of Philosophy and Culture*, 1–7.

Mori, M. (1970). Bukimi no tani [the uncanny valley]. *Energy, 7*, 33–35. Mori, M., K.F. MacDorman, and N. Kageki, (2012). The uncanny valley [from the field]. Robotics & Automation Magazine, IEEE. 19(2), 98–100.

O'Daly, G. J. (1973). Plotinus' philosophy of the self.

Ryan, R. M., & Deci, E. L., (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68.

Saygin, A. P., Chaminade, T., Ishiguro, H., Driver, J., & Frith, C. (2012). The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Social cognitive and affective neuroscience, 7*(4), 413–422.

Shelley, M. W. (2010). Frankenstein: 1818. Intervisual Books.

Sheridan, T. B., & Ferrell, W. R., (1974). Man-machine systems; Information, control, and decision models of human performance. the MIT press.

Shiomi, M., Nakagawa, K., & Hagita, N. (2013). Design of a gaze behavior at a small mistake moment for a robot. *Interaction Studies, 14*(3), 317–328.

Szalma, J. L., (2014). On the Application of Motivation Theory to Human Factors/Ergonomics Motivational Design Principles for Human–Technology Interaction. *Human Factors, 56*(8), 1453–147.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53*, 1–25.

West, J. L., & Halas, N. J. (2000). Applications of nanotechnology to biotechnology: Commentary. *Current Opinion in Biotechnology, 11*(2), 215–217.