

The Enactive and Interactive Dimensions of AI: Ingenuity and Imagination Through the Lens of Art and Music

Maki Sato*

University of Tokyo
Graduate School of Arts and
Sciences
The New Institute
maki.sato@aya.yale.edu

Jonathan McKinney

University of Cincinnati

Abstract Dualisms are pervasive. The divisions between the rational mind, the physical body, and the external natural world have set the stage for the successes and failures of contemporary cognitive science and artificial intelligence.¹ Advanced machine learning (ML) and artificial intelligence (AI) systems have been developed to draw art and compose music. Many take these facts as calls for a radical shift in our values and turn to questions about AI ethics, rights, and personhood. While the discussion of agency and rights is not wrong in principle, it is a form of misdirection in the current circumstances. Questions about an artificial agency can only come after a genuine reconciliation of human interactivity, creativity, and embodiment. This kind of challenge has both moral and theoretical force. In this article, the authors intend to contribute to embodied and enactive approaches to AI by exploring the interactive and contingent dimensions of machines through the lens of Japanese philosophy. One important takeaway from this project is that AI/ML systems should be recognized as powerful tools or instruments rather than as agents themselves.

Keywords

Japanese philosophy, enactivism, AI ethics, interactivity, contingency, embodiment

1 Introduction

Foundational theories of artificial intelligence (AI), following cognitive scientists, assume that the content and capacities of the machine and its programming are essential for intelligence. The fruits of these assumptions are the modern computers and myriad algorithms that permeate our world. Machine learning (ML) systems are powerful content generators that can compile astounding data sets and generate output from the given data set through either supervised, reinforced, or unsupervised learning processes. As technology advances and methods become more precise, these systems will likely become easier to use, faster, and less prone to disconcerting errors.

While early machines could generate and respond to input, they were criticized for only being capable of basic forms of processing (Turing, 1950). However, noteworthy advancements have been made through the combination of ML systems and embodied approaches to AI and robotics.

* Corresponding author.

1 This sentence briefly introduces the line from Cartesian mind–body dualism, “I am thinking therefore I exist” (Descartes, 1637/2006, p. 28) and its relation to cognitive science and artificial intelligence. In *Ecological Psychology* this is referred to as a “trialism” that divides the brain, the body, and the external world.

Although embodied and socially integrated AI systems may still fall short of human levels of creativity and intelligence, they represent radical improvements in the kinds of tools and instruments humans have available to them. Instead of viewing the inability of ML and AI systems to replicate human creativity as failures to be overcome, they should be viewed as monumental successes in the expansion of human expression (Flowers, 2019). To trace the path away from increasingly complex ML algorithms toward embodied and enactive forms of Artificial Life, we will focus on the importance of interactivity and contingency for creating art and music through the lens of modern Japanese philosophy.

1.1 From Embodied to Enacted AI

The enactive paradigm in cognitive science began as a response to the cognitivist and Eurocentric dogmas in the fields of philosophy, cognitive science, and artificial intelligence. In the inaugural text *The Embodied Mind*, Varela et al. (2017) challenged the cognitivist assumptions of mind–body and body–world dualism by grounding cognition in experience through the lived body.² Early enactivists saw a discontinuity between the models of cognition, intelligence, and living things and sought to build a path forward based on our phenomenological experiences. The hallmark of a living system for early enactivists was autopoiesis, or self-making, where organisms create and maintain themselves against environmental forces, like a cell that generates a cell wall to survive (Beer, 2004; Di Paolo & Thompson, 2014). These systems are considered operationally closed in that they continue to function even when the context changes. These ideas have been developed and now persist as autonomy, sense-making, and agency, where agents remain autonomous in the face of environmental forces and actively participate in the construction and maintenance of both themselves and their world (Di Paolo et al., 2017). That is, they enact their world through their exploration and interactions with the world they are embedded within.

As a key player in the embodied turn in cognitive science, the enactive approach has also influenced the turn toward embodiment in AI (Aguilar et al., 2014; Froese & Taguchi, 2019; Froese & Ziemke, 2009; Pfeifer & Bongard, 2006). Where classical approaches to AI (GOFAI) define the challenges of Artificial Life in terms of computation and intelligence, embodied AI theorists see these challenges in terms of action, perception, and interaction (Anderson, 2003; Brooks, 1991a,b; Brooks & Stein, 1994). While many predicted that the turn toward functional robots would inevitably lead to a genuine AI system, Brooks and other roboticists argue that even the most advanced robots have only captured the intelligence of insects or other simple organisms. This is at least partially due to the vast differences in the information available to robots and living things. Whereas living things have long evolutionary histories and access to a vibrant sensory world, machines and algorithmic systems have limited access to sensory data and are developed over a relatively short time span.

While it is a brilliant feat of engineering to create a machine that is as capable of navigating and flying as a dragonfly, there is more to life than maneuvering. Enactivists agree with embodied AI theorists like Brooks that intelligence depends on having access to a world but add that life depends on the agency and the autonomy to do otherwise (Di Paolo, 2004; Di Paolo & Thompson, 2014). When a living system is pushed by a harsh wind or pulled by a grappling undertow, it can attempt to resist them or shape them to its advantage. Unlike living systems, nonliving systems like rocks and toasters must simply endure the forces that surround them. While we can interact with a toaster and it can produce an output in response, it cannot choose to overcome an input command. This is crucially important because it establishes autonomy as a clear line between even the most complex machines and basic living systems. The enactive approach to AI is promising because it builds from the successes of embodied AI and provides a novel path forward. Echoing the enactive critique of AI, we argue that there are crucial distinctions between ML and AI systems. Importantly, ML systems often overlook the success of embodied AI and robotics and oversell their

² Importantly, this article will focus on the branch of enactivism that is founded on the works of Maurice Merleau-Ponty (see McKinney et al., 2021) and not the radical enactivism of Hutto and Myin (2012).

computational capabilities. Despite being powerful and complex machines, it is not clear that they ever truly understand or know anything.

If our goal is to create a genuine AI system, then we should follow the enactive path and focus on how machines can act and interact autonomously with their world and with other agents. To that end, we hope to contribute new perspectives on interactivity between agents and their world from the works of the modern Japanese philosophers Nishida Kitarō, Kuki Shuzō, and Watsuji Tetsurō. Our hope is that these perspectives will be tools that will help contextualize the shortcomings of ML and robotic systems in terms of their abilities to act as genuine interlocutors, rather than merely as systems capable of generating novel output. While novel and complex output may seem like intelligence to us, we should be skeptical of it without accompanying evidence of agency.

2 Problems with AI and ML Systems

The enactive and embodied approaches to AI have shifted the burden of proof in AI theory to optimists who argue that advanced AI is imminent. Brooks and Stein (1994) established the importance of modeling AI systems based on the only known examples of intelligent beings: living things. This work continues today through the work of AI and ML critics like Birhane (2021), who identify the many dangers of the careless deployment of AI based on the hype of technological advancement (Kaltheuner, 2021; Taylor et al., 2021).

While the embodied turn toward Artificial Life has produced many highly functional robots, it has also produced new forms of techno-optimism. Artificial neural networks, for example, resemble neurons in name only, yet are described as if they are parts of an artificial brain (Birhane, 2021).³ It seems that techno-optimism has motivated ML theorists to insist that ML systems are approaching human-level capabilities. Advanced language and artistic machines are described as being capable of holding a conversation or composing music. These interpretations, like the techno-optimistic positions before them, misunderstand what living systems are doing when they use language or create art. They are not creating novel linguistic or symbolic content but are expressing their autonomy through their relationship to other agents and their world. Without clear evidence that ML systems are capable of autonomously interacting with others and their world, the burden of proof of their intelligence remains unmet.

Generative models like generative adversarial networks (GANs) and generative pre-trained transformers (GPTs) are possible counterexamples that create novel content that is difficult to trace back to mere input. This means that these systems seem to be creative and to have some kind of agency. As a result, one could argue that the poetry, prose, and music produced by GANs or GPTs are works of art. For our analysis, these systems also embody some form of contingency that is crucial for living systems. Although these systems produce texts, images, and music that are high enough quality for it to be almost impossible to determine whether the work is generated by humans or machines, we argue that differences remain between living and nonliving systems. Whereas living systems fundamentally possess the abilities to relate and interact with others and their surrounding world, algorithmic systems do not. Simply attempting to trick human beings into accepting that machines are intelligent by focusing on intelligent-seeming output is insufficient for meeting Dreyfus' burden of proof (Dreyfus, 1992). This is because the farther algorithmic systems stray from embodiment and robotics, the less capable they become. Human beings are not defined by their ability to create enigmatic musical combinations, but by our ability to navigate and shape our world together.

2.1 Returning to Turing

Given the similarities in arguments regarding interactivity and autonomy (in behavior) between ML rhetoric and Turing's approach to AI, we will explore them together. Taking inspiration from

³ For further information, see Shpurov and Froese (2021).

Turing's (1950) central question—"Can machines think?"—we hope to delve into the criticisms of AI more directly with the question, "Can machines create art and music?" While our argument will not unravel the many issues facing ML/AI systems from engineering aspects, it will help carve a path toward more fruitful ways of thinking about intelligent machines and their roles in our lives.

2.1.1 Can Machines Create Art and Music?

If this is an empirical question, then the answer is surely "Yes." Machines already produce things like technically challenging music, abstract digital paintings, realistic digital landscapes, and humorous dialogues. If art is the content of the piece, then anything capable of producing melodic sounds, beautiful color pallets, or engaging prose would be capable of creating art. Indeed, digital art- and music-creating machines are capable of creating art and music in some sense (Weinberg et al., 2020). The force of the objections to the artistic abilities of machines is not that they struggle with the creation of artistic content, but that they are not the genuine inventors of the work. Machines may produce art, but they are not artists who instinctively communicate with their surroundings and intuitively create sensational art and mind-blowing music. Unlike an artist, machines are not autonomous, meaning that they are limited in the ways that they can respond to input from other agents or the environment. They may be able to act and react, but they lack the ability to make sense of their world through their interactions with other agents (De Jaegher & Di Paolo, 2007).⁴

Turing frames this kind of problem in terms of an argument from consciousness (which seems to inspire contemporary defenders of AI), that this is a matter of the quality of the works of art (Turing, 1950). The argument from consciousness challenges the idea that machines can be intelligent because machines lack the essential abilities to be creative and experience emotions. Turing responds by claiming that these kinds of problems will be solved by engineers in that they will be able to create machines that are powerful enough to emulate this feeling, which is echoed by theorists today. This is evidenced by a comprehensive meta-analysis of terms used in ML/AI research by Birhane et al. (2021), where terms like productivity, efficiency, and output completely overshadow discussions of ethics, autonomy, and interactivity.

This leads ML/AI researchers to frame the challenge of creating art in terms of the differences between a skilled painter and an inexperienced novice playing with paint. It is useful for ML and AI theorists to approach issues in this way because this escapes the entangling web of theory and becomes an engineering problem. As the training data increases, the quality and accuracy of the work will increase, meaning that it is only a matter of time before machines create art and music or become experienced artists and musicians. The results of this approach to AI have produced countless innovations in programs capable of generating digital images, producing music, and completing unfinished drafts of prose, etc. For Turing and his followers, the argument from consciousness fails because asking for more than the high-quality output from an author is more than we require of ourselves and other humans (Turing, 1950). After all, we also lack good evidence for the intelligence of human playwrights beyond their works and their human-like behavior.

This remains useful for differentiating between skilled and unskilled artists, but it falls apart because it fails to account for the embodied relationality of artists. Simply, there is more to a piece of art and melody than what is on the canvas and the music sheet. The artists behind a work relate to the world in a way that we can experience for ourselves. Artists change in real-time and through myriad life experiences through exposure to otherness, and we can experience these changes in some sense for ourselves. It may seem like a fine line between the evolution of an artists' style or shift of their perspective and the iterative updating of an ML system, but there are important distinctions to be made between what artists are doing and how ML systems function. Whereas artists have experiences and develop expertise, ML systems merely produce augmented copies of their input data.

⁴ For further information, see also Dotov and Froese (2020) and Fletcher-Watson et al. (2018).

The terms *machine learning* and *artificial intelligence* are misleading because even the most advanced language programs do not learn or develop outside of the one medium they receive input from.⁵ This is evidenced by their inability to resist environmental input, choose to do otherwise, or escape the highly specialized worlds they are designed for.⁶ For example, a music generation program cannot do anything but process its input and generate an output. It cannot leave its data set and explore another one. It lacks a body, a world to explore, and the freedom to augment its surroundings. Learning is one multimodal activity among many that comprise our lived experiences through our bodies. There simply are no living things outside of science fiction that lack a body and access to a world. In short, sensory inputs from the bodily movement are a necessity for a multimodal perception to change the mode of perception crucial to interactivity.

Instead of claiming that machines create art or that machines could become artists, we should think of machines as powerful tools that define an artistic medium.⁷ This is beneficial for several reasons. The first is that it properly situates ML systems within the history of technological innovations that have transformed the mediums of artistic expression. Much like the impacts of keyboards and soundboards on musical composition, ML systems make new kinds of musical engagement possible. Simply put, machines are not built to live and have experiences, so it is unclear how they could express anything. Thus, instead of machine painters and musicians, we have machines that are akin to enigmatic digital tools.

3 The Lens of Japanese Philosophy

Two shortcomings of disembodied approaches to ML and AI systems become clear when viewed through the lens of modern Japanese philosophy.⁸ We will begin with the works of Nishida Kitarō, who founded the Kyoto School and influenced the works of Kuki and Watsuji, whose works we will focus on in the next section. The world as pure experience (*junsui keiken*, 純粹経験) and the frame of action-intuition (*koiteki chokkan*, 行為的直感) in the works of Nishida Kitarō demonstrate two ways to view the agent-world relationship that current approaches to AI lack.⁹ The first manner of viewing the agent-world relationship views the sensory world as rich and unimpoverished, which is in contrast with how the world is characterized by cognitivist cognitive scientists.¹⁰ The second is the active interactivity of embodiment, or action-intuition (*koiteki chokkan*). Whereas AI systems with robotic bodies can perceive in order to act, for human beings, acting and perceiving are mutually reinforcing processes.¹¹ Without moving bodies, oscillating eyes, and postural sway, the human vision would be fundamentally different (Käufner & Chemero, 2021). In other words, constant movement enables human beings to act in resonance to the world that they are thrown into. While these theoretical frames do not completely capture the limits of artificial systems alone, they can act as initial steps to re-center human experience in AI.

3.1 Watsuji, Kuki, and Machine Spontaneity

In order to understand the challenges ML and AI systems currently face, we will draw from the works of Watsuji Tetsurō and Kuki Shūzō, which explore what it means to be human in terms of interactivity, contingency, and spontaneity. Watsuji argues in his *Ethics* (Watsuji, 1996, p. 9) that “the essential misconception prevalent in the modern world” is that it conceives “ethics as a problem of individual consciousness only.” Watsuji argues that although individualism itself is “an achievement

5 For comprehensive accounts of how the terms *artificial intelligence* and *machine learning* are misleading given the current state of their fields, see *Fake AI* (Kalthéuner, 2021).

6 For an enactive account of the challenge of meaning and learning in artificial systems, see Froese and Taguchi (2019).

7 For more about reframing the successes of ML systems in this way, see Flowers (2019).

8 For other perspectives on ML and AI systems using non-Western comparative philosophy, see Sato (2020).

9 Refer to McKinney et al. (2021) for details on Nishida's philosophy.

10 For an introduction to the embodied critique of the assumed poverty of the stimulus in cognitivist cognitive science, see Käufner and Chemero (2021, pp. 181–222). See also Di Paolo et al. (2017) for an enactive response to the challenge of impoverished coupling.

11 Here, the authors are differentiating an AI system that runs solely on a computer and an AI system attached to a robot that enables surrounding environmental perception through its sensors.

of the modern spirit,” such a notion only represents the “standpoint of the isolated ego” and that it lacks the notion of the “totality of *ningen* (humans).” For Watsuji, the “totality of *ningen*,” first and foremost, includes the importance of connections, interactions, and interplay that happen consciously and unconsciously between beings. This mutual interaction is not merely between agents but extends to the agent-environment relationship (Watsuji, 1935/1988). Watsuji argues that humans are inherently situated in the “in-betweenness of person and person”; therefore, he points out, “the locus of ethical problems lies not in the consciousness of the isolated individuals” (Watsuji, 1996, p. 10).

Watsuji’s initial interest in what it means to be human, inspired by Heidegger’s (1927/1962) argument made in *Being and Time*, was concerned with the specific problem of how humans interact with space.¹² This interest culminated in Watsuji’s *Climate and Culture* (1935/1988). However, in his later works, he expanded his argument to the dynamism of inherent human–human interactions. The importance of such interactions is the subject of his *Ethics*. The challenges that ML/AI systems currently face is not surprising when considered through Watsuji’s philosophical works. Many AI/ML systems are designed to process a particular kind of input or experience, which limits their capacities for dynamism and complex interactions with other agents.¹³

Indeed, ML systems only have access to certain data sets and thus can only produce conditional responses based on that input. However, as long as data sets are static, in the sense that those data are acquired from past events, the interactions between ML/AI and humans lack in their dynamism. This can be characterized using temporality and spatiality, where agents can both react to their spatial surroundings and interact meaningfully with other agents in real-time.¹⁴ Living beings are not only responding to their visual surroundings or the movements of others but are deeply immersed in the multimodality of experience that is social, visual, and contextual. One example of this is how living beings are capable of immediately predicting how other beings are going to respond and then act accordingly. Watsuji’s concept of in-betweenness captures this relationship in terms of the spontaneous interplay between living things.

Although there might be machines that are capable of more complex interactions, like swarm robots, these systems still lack access to an evolutionary history of information that shapes how they perceive, act, and interact with their surroundings.¹⁵ Swarm robots represent a step in the right direction toward developing machines that can relate to their immediate surroundings in lifelike ways, but, as Brooks (1999) reiterates, until these machines can pull from a lifetime of experiences or rely on a body that has co-evolved with its environment and other beings, these machines will fall short of human-level capabilities (Brooks, 1999). As a result of the bodies we have and the developmental and evolutionary histories, humans can do more than merely dynamically respond to our environment.¹⁶

4 Interactive and Musical Machines

Despite the lack of spontaneity and interactivity in many machines, there are some machines that seem more interactive than others. While digital art-generating machines may function as an artistic medium or tool, musical systems seem to be more interactive. Visual art creation, in that sense, only

12 This mirrors the approach taken by many embodied AI theorists, which focuses on how machines navigate their surroundings and then attempt to scale them up to deal with challenges of coordination.

13 Heidegger (1954/1977) points out in his essay “The Question Concerning Technology” that “technology is a means to an end,” and simultaneously “technology is a human activity” (p. 4). However, he argues that when the technology (*technē*) is used appropriately it reveals (*Entbergen*); “*Technē* belongs to bringing-forth, to *poiesis*; it is something poetic” (pp. 5–13).

14 For arguments in regard to body and mind relationship from a philosophical viewpoint, see Yuasa (1987).

15 Swarm robots are an interesting case because they are systems with advanced coordination and interaction capabilities that resemble the dynamics of living things. While these systems are especially good for tasks with clear targets or “prey,” such as foraging, they struggle with more general tasks. This is what sets them apart from living systems, which often wander or act without a precise target or goal. Spontaneity is an informative way to differentiate living and machine systems, where both act and interact to some extent, only living systems act spontaneously as we define it. See Brambilla et al. (2013, pp. 34–37).

16 For an account of how human beings interact with our environment using dynamical systems theory, see Beer (1995).

touches upon spatiality, and thus could be the outcome of the static work of data sets that may lack the dynamism of in-betweenness. However, musical performance is more dynamic and thus requires beings to act spontaneously while being immersed in the spatiality and temporality of the experience. There have been attempts to make musical robots that generate music in more reactive and collaborative ways.¹⁷ This shift from disembodied ML systems that generate music to embodied robotic musicians is crucial because it enables these systems to engage with others and their worlds more spontaneously.

Human musicians do not generate music in a sensory deprivation chamber after memorizing the wave patterns of thousands of songs. They listen to, and are moved by, music. From their place in a musical world, immersed in musical cultures and experiences, they express themselves through instruments and lyrics. For AI engineers, one challenge of overcoming the limitations of music generators involves incorporating perception.

Listening to music in ways that humans listen, rather than as data sets, is crucial for understanding both what music is and what it means to create it. Likewise, ML systems cannot simply be scaled up to replicate human behaviors because humans are not a system of mere reflexes and responses. Birhane (2021) rightly argues that it is misleading and harmful for AI researchers to claim that ML systems can automate human behavior and simulate human becoming (Birhane, 2021). For Birhane and van Dijk (2020), philosophical questions about the humanity of ML/AI systems are eclipsed by concerns about the harm of misleading narratives in ML/AI research. One moral takeaway of our project is that we should properly situate our expectations and arguments in our world, which then must center on notions of embodiment. This does not address the myriad problems of ML/AI research, but it is a nontrivial first step toward more humane ML/AI research.¹⁸

Weinberg et al. (2020) have made promising progress by incorporating principles of embodiment, interactivity, and improvisation in their machine musician project at the Georgia Institute of Technology. One prime example is Shimon, an improvising marimba-playing ML/AI system equipped with a robotic body. It is capable of music generation through perception (seeing and listening to other players through a sensory system embedded in its robotic body) that enables perceptive, expressive, and improvisational music playing with humans. With its robotic embodiment, while generating reactionary music, it can search for social and environmental cues and create its own cues for others to follow. While Shimon lacks the agency of a human musician, this approach is far more advanced and qualified as a musical performer than mere music generator programs. If the goal of AI is to achieve human-like capabilities, then we should focus on creating robots that are capable of social perception and assistance. This will require a return to embodied robotics with a built-in sensory system in conjunction with an ML/AI system.

Weinberg et al. (2020) explain the importance of limits for embodied musical machines, which have downstream implications for how ML/AI systems will be developed, how they impact our lives, and our expectations of AI. There is a seemingly paradoxical relationship between the limitations of bodies and how those limits enable action. While an arm is limited in its degrees of motion, these limits provide the necessary structure and coherence to its movements. Its limits reduce the cognitive load of moving, which makes it easier to move in precise and skillful ways. This is a core takeaway from the embodied turn in cognitive science and artificial intelligence. The body shapes thinking, and thinking is bound and realized through our embodiment (Pfeifer & Bongard, 2006). When thinking is divorced from the body, then critical elements are lost. By focusing on the computational power of the brain, we were able to create computers and calculators, but this also made simple problems like building a robot that could lift a cup much more difficult.

Thinking of limits in this way is crucial for avoiding problematic futurist tendencies in the field, which often associate improved performance as evidence of nearing genuine artificial general intelligence (AGI). Chalmers (2020) addresses this concern directly when considering the consciousness of GPT-3, a state-of-the-art language generator with billions of parameters, by comparing its

¹⁷ For a comprehensive account of robotic musicianship see Weinberg et al. (2020).

¹⁸ The Humane AI project initiated at the East-West Center suggests that “humane AI” connotes an ethical dimension in some sense.

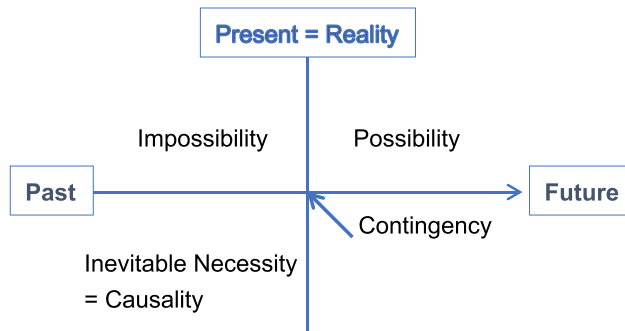


Figure 1. The place of contingency (adapted from Kuki, 2012).

consciousness to that of a worm. It is possible that the worm and GPT-3 are conscious only in some limited sense. In addition to theoretical concerns of consciousness and intelligence, there are practical differences between disembodied language programs like GPT-3 and worms: the abilities to relate and interact with others and their world. Whereas the worm and robots like Shimon can explore and manipulate their environments and interact with other agents in some sense, GPT-3 is completely cut off from the social, contextual, and perceptual dimensions of the world.

5 A Path Forward Through Contingency and Coincidence

When considering what it would take to make a machine that is more than a mere tool, Kuki Shūzō's notions of contingency and coincidence paint a clear path forward. Kuki worked on the problem of contingency throughout his life to tackle the issue of causality. In other words, his concern was to overcome the temporal constraints of a given lineage or fate based on a rigid causality. To overcome rigid causality, one must react spontaneously to one's contingent circumstances. As shown in Figure 1, a modified version from Kuki's own drawings, Kuki argues that the past is the inevitable necessity (*bituzensei*, 必然性) that has already happened as a consequential outcome of causality, whereas the future is the possibility (*kanousei*, 可能性) where nothing has yet unfolded thus no one can know what could happen (Kuki, 2012). Therefore, the “present” is the only possible moment where there can be an accidental or an unpredictable encounter. He argues that only with such a contingent encounter could a choice be made between multiple paths or between various ways of responding to one's circumstances that could then create new possibilities of opportunities in the future (Kuki, 2012). In short, we can only overcome contingencies right here and now in the present.¹⁹

Kuki (2012) argues that contingency is not something that can be analyzed in the past, nor that could be anticipated in the future, since contingency is contingent only in its “present-ness (*genzaisei*, 現在性).” Kuki further argues that “if we are to give meaning to the contingency of the present, we need to understand the contingency through the possibility of the future” and that “only in the futuristic moment yet to come, contingency would be given its meaning” (pp. 228–374). In other words, while the dynamism of human interactions is embedded in their shared circumstances, it empowers them to overcome these circumstances together. For example, one musician playing a riff in response to a melody performed by others is an act of spontaneous creativity that is made possible through their mutual contingency.

For Kuki, arts in general (not limited to that of paintings, music, and performances) connote contingency in their structural characteristics. In fact, when something is considered art, it subjugates

¹⁹ Kuki got his initial philosophical ideas on contingency and coincidences from Aristotle's argument on contingency, while considering Hegel's notion of necessity/inevitability (必然性) argued in *Encyclopaedia* (Hegel, 1830/1975). Although Kuki touches upon several Buddhist texts (Kuki, 2012, pp. 224, 282), his inspiration is not derived from early Buddhism.

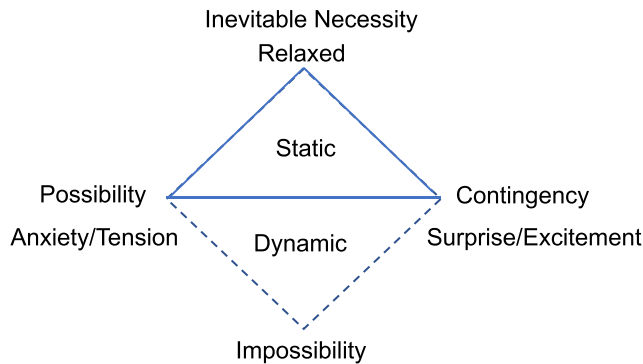


Figure 2. Emotional dimensions of Temporality (adapted from Kuki, 2012).

contingency. Therefore, Kuki (2012) argues that art embodies and fulfills freedom since it is freed from all things inevitable that are in the realm of necessity. With art being freed from the linear temporality where contingency is constantly happening, people view art with awe and surprise that brings about the excitement to those who encounter art in its “present-ness,” whether that art is static (such as paintings and poems) or dynamic (such as music or drama) (Kuki, 2012).²⁰

Kuki describes the temporality of experience as a creative and emotional relationship (see Figure 2). When the route of the causality is obvious or under static conditions where humans can easily predict what is going to happen, humans tend to be relaxed and bored, knowing what is going to happen next. However, in dynamic conditions, when things are non-predictable, and there is a possibility that things may happen outside the route of the stringent causality, then humans feel anxiety and tension. This is because there is a hidden expectation of what might happen. As a result, the mutual contingency between agents is a source of excitement and surprise.

Kuki’s argument suggests that the current embodied ML/AI lacks a fundamental contingency that may coincidentally occur. Although one might argue that complex systems like artificial neural networks are capable of learning by processing their data sets because it is difficult for human beings to predict and understand what output will follow from a particular input, it is clear that the output will be limited to a predetermined medium.²¹ Music and language bots compute input of a particular kind and can only produce an output of that kind. A music bot cannot produce prose or poetry regardless of the quality of its musical output. Even the simplest forms of life and expression can choose between passively experiencing and actively changing their experiences. In other words, there is a fundamental absence of the sheer contingency that Kuki argues is the major source of humans being moved and struck with works of art.

Being a student of Edmund Husserl at the University of Freiburg, Kuki also argues that “without consciousness, time does not exist,” and that flow of time can only exist in between one’s will (*isbi*, 意志) and its aim (*mokuteki*, 目的) (Kuki, 2016, p. 9). In Kuki’s view, AI/ML systems lack both discernable aims and the will to act. Even though embodied ML/AI systems look as if they are improvising music with humans, they do not have the instinctive joy of playing along with their music partners, or the ability to perform, or compose on their own. That is, they play music only contingently and never spontaneously. Unlike ML/AI systems that are restricted to producing output based on limited input and being activated by a human engineer, humans are constantly in the midst of dynamic creativity, becoming, and changing (Birhane, 2021). In other words, Kuki’s

20 If the art loses the ability to inspire awe in the viewer who was once inspired by it, then it is the viewer that changes not the art itself. This phenomenon happens because of the excessive encountering with the specific art and the encounter with the specific art becomes habitual. In this case the once thought of awe should be described as “discovery” as argued in the Gallese (2021).

21 If given stringent rules as in Go games (e.g., AlphaGo developed by Google DeepMind (<https://www.deepmind.com/>)), where there is a limitation for the contingency to sneak in, it is easy for the ML/AI to achieve, predict, and adjust its hand, that is, more easily than for humans.

argument suggests that without such constant interaction (temporality) based on the dynamic input sensory data collected from the surrounding environment (spatiality), there will be no will or aim.

As Watsuji argues, humans are open-ended, and part of what it means to be human is to aid and affect others, and thus be aided and affected by others. This kind of interactivity is not merely conditionally responding to sensory inputs but involves sharing and shaping each other's aims. Thus, ML/AI systems are less analogous to living beings and closer to powerful and complex tools. Watsuji's work makes this clear through his insistence on the importance of the in-betweenness of the agency. However, Kuki's philosophy requires more than that of Watsuji through his focus on temporality. An ML/AI agent must be capable of actively shaping its contingent reactions spontaneously in the present moment. To be creative, an agent must be free to change how it engages with its circumstances. This is crucial, because the circumstances before us are always changing and rarely analogous to repetitive data sets. In fact, Kuki argues that our experiences only come together accidentally and never happen again in just the same way. Thus, living beings must be prepared for new experiences, even when their circumstances seem familiar.

6 Conclusion

In conclusion, ML/AI systems can neither act nor truly interact. The *Turing test* is not only about fooling humans, which is an arbitrary and relatively easy thing to do. It is about successfully interacting with human beings (Turing, 1951/2004). This cannot be solved with increasing computational power but must be hashed out through exploring the shared spaces of language and expression. This is a multimodal problem that calls for embodied and enactive approaches to solve. While intelligent robots lack the vast lexical database of advanced ML/AI systems, they are capable of navigating some aspect of our shared world, which is fundamentally important for human lived experiences. They are becoming intelligent to some extent when there is a skillful interaction with others and the world in a narrow context.

Our hope is to use the tools of Japanese philosophy to shed light on how easy it is to misunderstand ML/AI issues, and thus fail to understand how to address these misunderstandings. The purpose of our argument is not that art-creating machines can never be artists or are inherently immoral, but that visions of futuristic disembodied AI systems are harmful and should be situated more responsibly through the embodied and enactive approaches, where interactions are made visible to enhance communicable in-betweenness between humans and ML/AI systems. However, we should think carefully about the impacts of uncritically creating and promoting machine-generated art to people already facing social, cultural, and economic hardship. If the problem is that we want more music, we can fund music programs, hire musicians, and standardize higher pay for performers. We also hope to encourage AI/ML theorists to embrace the successes of embodied AI and robotics, which have created many powerful tools that augment our capacities of expression.

Furthermore, we want to acknowledge that there are important hurdles such as agency and intersubjectivity that ML/AI systems may never overcome. Kuki argues that the beauty and awe of art lie in the "contingency" that is a source of freedom where creativity flourishes that leads to unpredictable and unknowable reactions, which occur in the very in-betweenness of the subjects. After all, if machine art and music lack the "contingency" and the authentic interplay that coincidentally occurs in-between embodied subjects, then they are the mere products of the machine's designers.

References

- Aguilar, W., Santamaría-Bonfil, G., Froese, T., & Gershenson, C. (2014). The past, present, and future of artificial life. *Frontiers in Robotics and AI*, 1, Article 8. <https://doi.org/10.3389/frobt.2014.00008>
- Anderson, M. L. (2003). Embodied cognition: A field guide. *Artificial Intelligence*, 149(1), 941–130. [https://doi.org/10.1016/S0004-3702\(03\)00054-7](https://doi.org/10.1016/S0004-3702(03)00054-7)
- Beer, R. D. (1995). A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence*, 72(1–2), 173–215. [https://doi.org/10.1016/0004-3702\(94\)00005-L](https://doi.org/10.1016/0004-3702(94)00005-L)

- Beer, R. D. (2004). Autopoiesis and cognition in the game of life. *Artificial Life*, 10(3), 309–326. <https://doi.org/10.1162/1064546041255539>, PubMed: 15245630
- Birhane, A. (2021). The impossibility of automating ambiguity. *Artificial Life*, 27(1), 44–61. https://doi.org/10.1162/artl_a_00336, PubMed: 34529757
- Birhane, A., Kalluri, P., Card, D., Agnew, W., Dotan, R., & Bao, M. (2021). *The values encoded in machine learning research*. ArXiv. <https://doi.org/10.48550/arXiv.2106.15590>
- Birhane, A., & van Dijk, J. (2020). Robot rights?: Let’s talk about human welfare instead. In *AIES ’20: Proceedings of the AAAI/ACM conference on ai, ethics, and society* (pp. 207–213). ACM. <https://doi.org/10.1145/3375627.3375855>
- Brambilla, M., Ferrante, E., Birattari, M., & Dorigo, M. (2013). Swarm robotics: A Review from the swarm engineering perspective. *Swarm Intelligence*, 7(1), 1–41. <https://doi.org/10.1007/s11721-012-0075-2>
- Brooks, R. A. (1991a). Intelligence without representation. *Artificial Intelligence*, 47(1–3), 139–159. [https://doi.org/10.1016/0004-3702\(91\)90053-M](https://doi.org/10.1016/0004-3702(91)90053-M)
- Brooks, R. A. (1991b). How to build complete creatures rather than isolated cognitive simulators. In K. VanLehn (Ed.), *Architectures for intelligence* (pp. 225–239). Lawrence Erlbaum.
- Brooks, R. A. (1999). *Cambrian intelligence: The early history of the new AI*. MIT Press. <https://doi.org/10.7551/mitpress/1716.001.0001>
- Brooks, R. A., & Stein, L. A. (1994). Building brains for bodies. *Autonomous Robots*, 1(1), 7–25. <https://doi.org/10.1007/BF00735340>
- Chalmers, D. (2020, July 30). GPT-3 and general intelligence. *Daily Nous*. <https://dailynous.com/2020/07/30/philosophers-gpt-3/#chalmers>
- Descartes, R. (2006). *Discourse on the method of correctly conducting one’s reason and seeking truth in the sciences* (I. McLaren, Trans.). Oxford University Press. (Original work published 1637).
- De Jaegher, H., & Di Paolo, E. A. (2007). Participatory sense-making. *Phenomenology and the Cognitive Sciences*, 6(4), 485–507. <https://doi.org/10.1007/s11097-007-9076-9>
- Di Paolo, E. A. (2004). Unbinding biological autonomy: Francisco Varela’s contributions to artificial life. *Artificial Life*, 10(3), 231–233. <https://doi.org/10.1162/1064546041255566>, PubMed: 15245625
- Di Paolo, E. [A.], Buhrmann, T., & Barandiaran, X. (2017). *Sensorimotor life: An enactive proposal*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198786849.001.0001>
- Di Paolo, E. A., & Thompson, E. (2014). The enactive approach. In L. Shapiro (Ed.), *The Routledge handbook of embodied cognition* (pp. 68–78). Routledge.
- Dotov, D., & Froese, T. (2020). Dynamic interactive artificial intelligence: Sketches for a future AI based on human-machine interaction. In *ALIFE 2020: The 2020 conference on artificial life* (pp. 139–145). MIT Press. https://doi.org/10.1162/isal_a_00350
- Dreyfus, H. L. (1992). *What computers still can’t do: A critique of artificial reason*. MIT Press.
- Fletcher-Watson, S., De Jaegher, H., van Dijk, J., Frauenberger, C., Magnée, M., & Ye, J. (2018). Diversity computing. *Interactions*, 25(5), 28–33. <https://doi.org/10.1145/3243461>
- Flowers, J. C. (2019). Rethinking algorithmic bias through phenomenology and pragmatism. In D. Wittkower (Ed.), *Computer ethics-philosophical enquiry (CEPE) proceedings*. INSEIT. <https://doi.org/10.25884/mh5z-fb89>
- Froese, T., & Taguchi, S. (2019). The problem of meaning in AI and robotics: Still with us after all these years. *Philosophies*, 4(2), Article 14. <https://doi.org/10.3390/philosophies4020014>
- Froese, T., & Ziemke, T. (2009). Enactive artificial intelligence: Investigating the systemic organization of life and mind. *Artificial Intelligence*, 173(3–4), 466–500. <https://doi.org/10.1016/j.artint.2008.12.001>
- Gallese, V. (2021). Brain, body, habit, and the performative quality of aesthetics. In F. Caruana & I. Testa (Eds.), *Habits: Pragmatist approaches from cognitive science, neuroscience, and social theory* (pp. 376–394). Cambridge University Press. <https://doi.org/10.1017/9781108682312.019>
- Hegel, G. W. F. (1975). *Hegel’s logic, Being part one of the encyclopaedia of the philosophical sciences* (W. Wallace, Trans.). Oxford University Press. (Original work published in 1830).
- Heidegger, M. (1962). *Being and time* (J. Macquarrie & E. Robinson, Trans.). Harper & Row. (Original work published 1927).

- Heidegger, M. (1977). *The question concerning technology and other essays* (W. Lovitt, Trans.). Garland Publishing Inc. (“The question concerning technology”: original work published 1954).
- Hutto, D. D., & Myin, E. (2012). *Radicalizing enactivism: Basic minds without content*. MIT Press. <https://doi.org/10.7551/mitpress/9780262018548.001.0001>
- Kaltheuner, F. (Ed.). (2021). *Fake AI*. Meatspace Press.
- Käufer, S., & Chemero, A. (2021). *Phenomenology: An introduction* (2nd ed., pp. 181–199). Polity Press.
- Kuki, S. (2012). 偶然性の問題 [The problem of contingency]. Iwanami.
- Kuki, S. (2016). 時間論：他二篇 [On time/Propos sur le temps]. Iwanami.
- McKinney, J., Sato, M., & Chemero, A. (2021). Habit, ontology, and embodied cognition without borders: James, Merleau-Ponty, and Nishida. In F. Caruana & I. Testa (Eds.), *Habits: Pragmatist approaches from cognitive science, neuroscience, and social theory* (pp. 184–203). Cambridge University Press. <https://doi.org/10.1017/9781108682312.009>
- Pfeifer, R., & Bongard, J. (2006). *How the body shapes the way we think: A new view of intelligence*. MIT Press. <https://doi.org/10.7551/mitpress/3585.001.0001>
- Sato, M. (Ed.). (2020). *5E Cognition (embodied, enactive, extended, embedded, and ecological) in the age of virtual environments and artificial intelligence*. The University of Tokyo Humanities Center Booklet. Vol. 9. Available at https://repository.dl.itc.u-tokyo.ac.jp/search?page=1&size=100&sort=controlnumber&search_type=2&q=1620880162891×tamp=1624416357.0795293
- Shpurov, I., & Froese, T. (2021). Combining self-critical dynamics and Hebbian learning to explain the utility of bursty dynamics in neural networks. In *2021 IEEE symposium series on computational intelligence (SSCI)* (pp. 1–6). IEEE. <https://doi.org/10.1109/SSCI50451.2021.9660026>
- Taylor, L., Martin, A., Sharma, G., & Jameson, S. (Eds.). (2021). *Data justice and COVID-19: Global perspectives*. Meatspace Press.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>
- Turing, A. M. (2004). Intelligent machinery, a heretical theory. In S. M. Shieber (Ed.), *The Turing test: Verbal behavior as the hallmark of intelligence* (pp. 105–109). MIT Press. (Original work dated c. 1951).
- Varela, F. J., Thompson, E., & Rosch, E. (2017). *The embodied mind: Cognitive science and human experience* (Rev. ed.). MIT Press. <https://doi.org/10.7551/mitpress/9780262529365.001.0001>
- Watsuji, T. (1988). *Climate and culture: A philosophical study* (G. Bownas, Trans.) [風土]. Greenwood Press. (Original work published 1935).
- Watsuji, T. (1996). *Watsuji Tetsurō's Rinrigaku* (Y. Seisaku & R. E. Carter, Trans.) [倫理学]. State University of New York Press.
- Weinberg, G., Bretan, M., Hoffman, G., & Driscoll, S. (2020). *Robotic musicianship: Embodied artificial creativity and mechatronic musical expression* (pp. 1–21). Springer. <https://doi.org/10.1007/978-3-030-38930-7>
- Yuasa, Y. (1987). *The body: Toward an Eastern mind-body theory*. State University of New York Press.