Modeled After Life Forms: Embodiment and Ways of Being an Intelligent Robot

Asli Kemiksiz
Graduate School of Human Sciences, Osaka University

Abstract

Recently, the notion of "embodiment" has become a crucial concept in robotics and Artificial Intelligence (AI) research. Embodiment emphasizes that a physical body is a requirement for intelligent interaction with the physical world. Based mainly on my fieldwork in a cognitive robotics laboratory in the Kantō region, this paper examines the fluid thought styles of roboticists, who draw on a range of disciplines in order to operationalize embodied intelligence in robots. These thought styles shape computational assemblages and the making of widely different robots intended to mirror particular aspects of life. The paper shows "life" to be infused into a research area that primarily deals with machines, data, and complicated algorithms, despite the fact that the robots are not themselves seen to be alive.

Key words: Robotics, Embodiment, Japan, Science and Technology, Artificial Intelligence

Acknowledgements

An earlier version of this paper was presented to the East Asian Anthropological Association Annual Meeting 2017 with the support of the Japanese Society of Cultural Anthropology. I would also like to offer my thanks to my supervisor Atsuro Morita and to Casper Bruun Jensen for their support.

> "Anthropology holds up a great mirror to man and lets him look at himself in his infinite variety" (Kluckhohn 1949: 11)

In the mid 20th century, the American anthropologist Clyde Kluckhohn asserted that anthropology plays "a central role in the integration of the human sciences" (Kluckhohn 1949: 1). Lamenting the underdevelopment of human sciences, Kluckhohn assigned anthropology a special role in this nascent field. By exploring the differences and commonalities of humanity, anthropology could help formulate a comprehensive view of "human nature". Since then, both anthropology and the sciences in general have changed many times over. Traditional human sciences, including psychology, anthropology, and sociology, have become fully established disciplines and new fields such as cognitive science, artificial intelligence (AI) and robotics are now exploring human nature from angles that Kluckhohn could not possibly have imagined. These new disciplines too explore human nature but using different "mirrors".

Indeed, the metaphor of the mirror is widely used in these newer disciplines. For example, cognitive science in the 1970s saw the computer as a mirror for studying human minds. Rather than depicting human diversity, the computer reproduces (aspects of) human intelligence in a different medium and thus helps understand its essential nature.

The philosopher John Searle pondered the fact that metaphors for the brain have changed over time. The ancient Greeks compared it to a catapult, and in Searle's childhood it was likened to a telephone-switch: "Because we do not understand the brain very well we are constantly tempted to use the latest technology as a model for trying to understand it" (Searle 1984: 42). In the late 20th century, however, the robot has gradually replaced the computer, as roboticists have come to see the robot as a mirror for the human. This premise is embedded in their research practices.

Since its inception in the realm of fiction, the robot has always been associated with the question of what makes us human. Originally seen as an artificial person, the robot has been defined by its similarity to and difference from 'natural' humans. With the emergence of robotics as an academic field, the figure of the robot has moved from the realm of fiction and become the name for machines of various forms and functions. As part of this diversification, questions of what makes something alive have also become central to the field. Initially hailing from mechanical engineering and computer science, robotics might not seem to have much to do with life; yet over recent decades, roboticists have become increasingly interested in life as a point of reference.

A main factor in this emerging interest concerns what the roboticists Rolf Pfeifer and Josh Bongard (2007: 34) have called "the embodied turn". In short, this refers to the birth of a school of artificial intelligence (hereafter AI) research according to which intelligence requires a body. Traditional or classical approaches to AI conceived human intelligence to be similar to computer programs, as consisting of symbol processing in the brain (Pfeifer and Bongard 2007: 27). The metaphor of brain-as-computer also affected how cognitive neurosciences and psychology saw the human mind, particularly in the 1980s. In the same decade, artificial neural networks (ANNs)[1] also entered the scene. In the early 1990s, the "embodied turn" was ushered in by Rodney Brooks, whose notion of embodied intelligence viewed the body as a prerequisite for the development of intelligence, seen as the consequence of interactions with the world (Pfeifer and Bongard 2007: 29).

Briefly stated, the notion of embodied intelligence brought about a gradual shift in the design of artificial agents including algorithms and robots. It also led to the formation of new research areas shading into various disciplines and relying on novel combinations of theories and methodologies. One particular impact of the embodied turn was to open up new ways of thinking about intelligence, inspired by various life forms. Among other things, an increasing amount of robotics research, which previously used control engineering to reproduce human kinematic motions, began to find inspiration in cognitive or bio-inspired fields. In cognitive and developmental robotics, a new generation of scholars has thus developed what is known as the "synthetic approach," which, in a nutshell, aims to understand human embodied intelligence by making or recreating it in robots (Pfeifer and Bongard 2007: 21). This located the engineering efforts to create AI and robots in the context of broader explorations of human nature. The synthetic approach entails the assumption that making more intelligent agents deepens our understanding of intelligence —or some of its aspects.

Given these changing contexts, research areas such as cognitive robotics can be placed among wider transformations of the notion of life and the scientific practices that explore it. An example of such a

---

[1] These are computational models designed to resemble human brains. They were first developed in the 1940s. Today they are used in many computational technologies.

point of contact is provided by the anthropologist of science and technology Stefan Helmreich's ethnographic work on the "limit biologies" of artificial lifers [2] (ALifers), marine biologists, and astrobiologists, which particularly points to the instability of the concept of "life" (Helmreich 2011: 693). As biology expands through new channels in technoscience, Helmreich showed, understandings of "life" become contested, unstable and transformable. In this analysis, he distinguished *life forms* seen as "embodied bits of vitality" from *forms of life* understood as "social, symbolic, and pragmatic ways of thinking and acting that organize human communities" (Helmreich 2009: 6). On this basis, he argued that life forms and forms of life "inform, transform and deform each other" (Helmreich 2011: 676). Life-inspired robotics, too, joins this traffic of life forms and forms of life. In a similar vein, Richard Doyle has analyzed the transformation of "life" and life sciences through hermeneutic readings of molecular biology and artificial life (Doyle 1997, 2003). In this context, various life forms inspire the designs of robots, which, in turn, inform and transform our understanding of human intelligence, as a certain generalized form of life.

In terms of my ethnography, life does not appear very universal, even within robotics research. During my fieldwork over the last few years, I have encountered many divergent views on life and practices trying to deal with it. As if corresponding to this diversity, the robots I have observed are also extremely variable both in form and function.

To analyze this divergence, I have been inspired by studies of the

---

[2] In this paper ALife refers to the domain of research in sciences and engineering that aims to understand—and sometimes to create—life through computer simulations and machines.

genesis and development of scientific facts by the German microbiologist and sociologist of science Ludwik Fleck (1935). His major work developed the notion of thought styles *(Denkstilen)*, defined as the "readiness for directed perception and the assimilation of what has been perceived" (Fleck 1935: 142). While thought styles generate knowledge, they do so from within particular "thought collectives *(Denkkollektiven)* in which practitioners interact and exchange ideas (Fleck 1935: 39). Through this process, different thought collectives gradually construct their own ways of perceiving and churning data or ideas, which in turn shapes the facts they make. The difference of thought styles in various subfields of robotics is evident in how they see grand concepts like life, intelligence, or embodiment.

This paper is about roboticists who are inspired by life forms, in Helmreich's sense, yet who do not see their robots *as* alive. Instead, their research practices engage with bits and pieces of different live entities. They attempt, that is, to reproduce certain humanlike or lifelike features (e.g. behavior, movement) in the medium of a robot by modeling the artifact after fragments of other life forms. The metaphorical and pragmatic relation they establish between the organic and the technological can be seen in their ways of thinking and acting within particular "computational assemblages" (Johnston 2008). As I discuss further on, modeling *after life forms* is a crucial dimension of this research practice.

For this research, I have been conducting multi-sited fieldwork focusing in particular on a cognitive robotics laboratory in the Kantō Region. This is a university laboratory with around 40 students, one associate professor, one part-time lecturer, one post-doc, two secretaries and Professor Hideaki Matsumoto[3] as their head. Like many other roboticists, Matsumoto originally comes from a mechanical engineering background. This was followed by many years of work on computer science and cognitive science. Currently, he is a well-known figure in both AI and robotics circles.

Matsumoto's laboratory reflects his particular vision of robotics: a cognitive approach aiming to develop learning computational systems that interact with the real, physical world. Specifically, the laboratory members use ANNs to teach robots certain behaviors. Students take introductory level courses in cognitive science, and later on familiarize themselves with the relevant academic literature. Even though this is a robotics laboratory, it is not a hardware-oriented laboratory that develops its own robotic platform. Instead, it uses existing research-oriented robotic platforms such as Softbank's NAO to enable tool manipulation and interaction with the environment using machine-learning techniques.

The structure of the Matsumoto laboratory is based on groups, which are coordinated according to research themes, and smaller collaborative teams spanning other research institutions and private companies. If they are working with ANNs, students from other robotics laboratories of the university are also located in the rooms of the Matsumoto laboratory. The groups and teams change dynamically according to the flow of students and changing collaborative arrangements, but the focus on utilizing both machine learning and robotics is the common denominator, and, as we shall see, this makes embodiment a central concern.

In the following sections, I examine important aspects of the thought styles that guide the research conducted in Matsumoto

---

[3] I use pseudonyms for my informants throughout the paper.

laboratory. As I show, these thought styles shape the "computational assemblages" (Johnston 2018) through which research unfolds. In particular, this research experiments with modeling robots *after life forms*, without actually ascribing the robots themselves with *life*.

A Story of Robots and How Anthropologists Came to Care About Them

The history of robots follows a meandering path that cuts across the realms of fiction, engineering and various scientific questions about humanity. Famously, the term "robot" was coined in Karel Čapek's *Rossum's Universal Robots (R.U.R.),* published in 1921. In this play, robots were laborers made of organic material, made to work instead of human beings. Of course, the attribution of human qualities to things is an ancient phenomenon, and there had been humanlike or lifelike machines both in fiction and in real life prior to R.U.R. *Automata,* machines that mimic life, gathered widespread attention in Europe in the 18th century. There are still earlier examples in the Islamic Middle East[4] and of *karakuri ningyō (*mechanical dolls) in Japan (e.g. Thornbury 1992), which were fascinating and complex machines made mainly for entertainment.

In the hands of Karel Čapek, the "robot" gained a distinctive "labor" dimension. Indeed, his brother Josef had derived the word itself from *"robota,"* which means "forced labor" in Czech (Stableford 2006, 442). In *R.U.R.*, the robot symbolized the working class under capitalism. Yet, the concept of a humanlike machine working on

---

[4] For instance, Ismail ibn al-Razzaz al-Jazari depicts his designs for automata and similar devices that resemble life forms in *Book of Ingenious Devices* (1207) (Masood 2009: 163–164).

behalf of humans quickly came to be seen as desirable, and real-life machines dubbed “robots” started appearing during the same decade (Kubo 2016: 49). Ever since, one part of the semantic field of robots has been the idea that, for better or worse, they might replace humans. Today, “factory robots are the ‘species’ that mostly closely conform to this idea” (Pfeifer and Bongard 2007: 10). Even so, the term “robot” came to be applied to a much broader range of machines with many different functions.

The very diverse contexts in which the term robot has been adopted makes it resistant to unequivocal definition. Dictionaries tend to emphasize “programmability” by a computer and the automated execution of tasks. Yet, there are robots that do not fully embody those notions, and non-programmable or human-operated machines can also be referred to as robots. Both in fiction and in real life, robots have thus become a mixed category of machines of various shapes, sizes, and functions. Furthermore, as robots have become part of more and more contexts, the literature and research on them have also branched out, and in consequence, robots are now being shaped by increasingly diverse and entangled thought collectives with quite divergent thought styles, some of which I return to below.

Eventually, robots also began to pique the curiosity of some anthropologists. Considering that they started to appear in the real world soon after the term was coined by Čapek, it could be said that the anthropological interest was somewhat belated. By the late 20th century, automation with industrial robots had transformed factories (e.g. Zuboff 1988). Yet what mainly attracted anthropologists was not so much these prosaic working robots as different kinds that shared the intimate social spheres of people.

Within the anthropology of (Japanese) robots, Jennifer Robertson’s work is exemplary in this regard. She has closely examined the socio-economic dynamics of Japanese robotics and their role in the supposedly post-human transformation of Japanese society, connected with problems such as the ageing society, gender issues and labor shortage, and often tied to the claim that the Japanese belief system is relatively more welcoming to robots than those of Western countries (Robertson 2007, 2010a, 2010b, 2011, 2014). For Robertson, robots are thus mirrors that reflect Japanese society and its political economy. As we shall see, this emphasis on cultural particularity is not necessarily shared with roboticists, who see the robot as a synthetic means to explore an imagined universal human nature.

In fact, the practices of robotics are not totally absent in Robertson's studies. For example, she has commented on embodied intelligence and its importance for Japanese roboticists (e.g. Robertson 2007, 2010, 2014). It is fair to say, however, that the cultural parameters of her analytical framework have remained stable. Yet, while keywords like "ageing society" or "labor shortage also show up with some frequency in the discourse of roboticists, their main function is to signal the value of their research to outsiders. Because it is common and often necessary to tie research to general societal concerns, roboticists working in Japan do indeed allude to the demographic and economic issues of the country. For roboticists engaged in research on embodied intelligence, however, these terms are rather marginal. For my informants, at least, what matters is to understand human cognition in order to make better robots, while any societal benefit this may generate is a secondary added value. In order to understand the practices through which robots are actually imagined, built and transformed, it is thus necessary to probe deeper into the thought styles and practices of the roboticists themselves.

The anthropologist Lucy Suchman pioneered discussions on computational artifacts, focusing particularly on their interaction with humans. The second edition of her *Human-Machine Reconfigurations: Plans and Situated Actions* (2007) was expanded with analyses of developments in robotics and AI since the mid-1990s – the same general developments that shaped my field site. Suchman's comprehensive and critical account of computational artifacts across disciplines, studied through the lens of feminist science studies, has been particularly inspirational for my own thinking about the material-semiotics of the robot.

An anthropological orientation towards the practice of robotics is exemplified by Akinori Kubo's *Robotto no Jinruigaku* (*The Anthropology of the Robot,* 2015), which provides a detailed history of robots in Japanese popular fiction and technology. Based on fieldwork with communities focused on the robot dog AIBO, the book offers a wide examination of the robot in mainstream culture, drawing attention also to the importance of paradigms by Japanese *anime* and *manga* critics. Rather than treating robots as a mirror of Japanese society, Kubo depicts robots as hybrids emerging from transdisciplinary movements of ideas and imaginations across popular culture, engineering and science. Kubo's work is interesting due to its multidimensionality and because it pays close attention to the situated concerns of divergent practitioners.

In certain ways, Kubo's approach is similar to John Johnston's *The Allure of Machinic Life* (2008), a cultural study of robots and AI. In this work, Johnston developed the idea of "computational assemblages" to analyze the history of cybernetics and artificial intelligence in the United States. A computational assemblage designates "a material computational device set up or programmed to process information in specific ways together with a specific discourse that explains and evaluates its function, purpose, and significance"

(Johnston 2008: 8). In a sense, the notion of computational assemblages formalizes Kubo's historical description of the transboundary travels of ideas around robotics. That is, the concept captures how bits and pieces of discourse, imagination and artifacts are related to particular devices or algorithms that process information in specific ways. Johnston further argues that these assemblages are designed to produce what he calls "machinic life," characterized by "mirroring in purposeful action the behavior associated with organic life while also suggesting an altogether different form of 'life'" (Johnston 2008: 1). Along these lines, Johnston shows how various computational assemblages, including some relating to robots and neural networks, were created to produce diverse forms of machinic life, thereby contributing to further transformations and complications of the concept of life.

In this analysis, Johnston shows the importance of being equally attentive to the discursive and material aspects of robotics research. For instance, he shows that connectionist models based on the immune system and neural networks are in some ways discursively similar. At the same time, however, he makes it clear that the "respective discourses are directed toward different ends, making each one part of a distinctly different computational assemblage, to be analyzed and explored as such" (Johnston 2008: 8). Consider, for example, the difference between a companion robot and a factory robot. The first is associated with sociality and made to interpret and express emotions whereas the properties emphasized in the second are safety and accuracy of operations. In both cases, the physical bodies and technical capacities of each robot are fully entangled with the discourses and imaginations used to make sense of them.

Building on and contributing to the anthropological literature on robotics and AI, I examine how the scientific imagination embedded in roboticists' thought styles—in particular their views on life, embodiment, and intelligence—are integral to the computational assemblages they create, the models they make, and the robots they build. In particular, I emphasize that models that try to reproduce certain aspects of organic life forms are crucial to the computational assemblages of robotics research.

Embodiment in Robotics

Among roboticists, a commonly used argument for embodied intelligence emphasizes that all forms of intelligence that we know of are, in fact, embodied. Or, our understanding of intelligence is invariably in reference to embodied agents. According to Professor Matsumoto, my main informant, even mathematical concepts and equations, such as "1+1=2," which are ostensibly not dependent on a body, are either grounded in a body, or can be traced back to one. Nevertheless, even if originally created by embodied agents, aspects of intelligence that deal with symbols are in some sense "detached from reality." Matsumoto illustrates this point with reference to languages: a computer program that translates from one language to another does not need a human body. In contrast, movement requires a body, and movements in terms of meaningful interactions with the real world are thus indicative of intelligence. In Matsumoto's view, therefore, even though traditional AI cannot be repudiated because it focuses on symbol processing, embodied intelligence "makes sense."

As mentioned, Matsumoto is a leading figure in Japanese robotics and AI societies. He often points to the different perspectives on embodiment held by researchers in these disciplines in Japan, which he describes as a "gap between the worlds". This "gap" can be seen as a significant indication of the divergence between thought collectives involved in AI and robotics in Japan. Matsumoto exemplifies the gap by noting that embodiment is a core concept for roboticists who work on intelligence, whereas for AI researchers it is merely one application among others. There is also a wide area within robotics that is simply not concerned with (embodied) intelligence. At the risk of oversimplification, the development of robots for automation that only requires the agent to produce the same movement repetitively does not depend on "intelligence" in the sense in which many roboticists refer to it. Instead, intelligence is about being able to sense changing circumstances in the environment and change movements accordingly. In my fieldwork, intelligence is thus almost always akin to adaptability, and—in Matsumoto laboratory—to learning.

What does the "body" invoked by "embodiment" refer to in this context? Robotics textbooks often skip the definition and even Pfeifer and Bongard's thorough discussion is not explicit on this point. Following my informants' insistence on the importance of changing movements in response to changing circumstances, however, we can at least venture that a body is a physical entity capable of flexible adjustment to its environment.

During my fieldwork, I was referred to a definition provided by Yasuo Kuniyoshi, one of the most famed Japanese roboticists of our time, for whom the body is "the invariant factor that universally restricts the interactions with the outside world such as ways of movement and reaction." The body, Kuniyoshi further explained, is therefore "the factor that is the slowest to change in comparison to the change in interaction"[5] (Kuniyoshi 2008: 284). Noticeably, this definition emphasizes limitation; the body is a configuration that only allows certain types of movement and reaction under certain

---

[5] My translation.

circumstances. In other words, the limits imposed by the body determine how an agent interacts with the environment. You only have to imagine an elephant with a fly on its trunk to grasp how wildly different bodily interactions and their attendant "intelligent behaviors" can be. The elephant might reach for a tree branch with its trunk, thereby making the fly take off, only to land again seconds later. The elephant and the fly both act intelligently and yet it is not possible to replicate the fly's behavior with an elephant's body, or vice versa. This is why the morphology and materiality of a given body is an integral part of the agent's intelligence.

Even though embodied intelligence assumes that intelligence requires a body, it does not prescribe what *type* of body is necessary. All organic bodies have their own sets of possibilities and constraints, but there is no need for robotic bodies to be configured in the same way as organic ones. For instance, integrated sensors such as GPS and ultrasound provide a set of inputs for the perception of the world that no human possesses, while wheels or rotary wings are body parts that do not exist in nature. In practice, however, research is overwhelmingly influenced by the workings of existing, organic bodies, sampled from nature. Furthermore, even wheeled robots use operating systems that are modeled after the perception of human bodies. In these ways, models from life inspire robotics.

Multiple groups in Matsumoto laboratory do research on or with humanoid robots such as NAO, Nextage, and Baxter.[6] In general

---

[6] NAO, as mentioned before, is a humanoid robotic platform developed by Aldebaran Robotics of Softbank (Softbank 2018). Nextage is a humanoid designed by Kawada robotics, mainly for factory use (Kawada 2018). Baxter, which is also a humanoid industrial robot is designed by Rodney Brooks' Rethink Robotics (Rethink 2018). Nextage and Baxter are not traditional industrial robots as they can be trained with ANNs.

terms, they teach the robotic systems to act in certain ways in response to certain circumstances. In most cases, the robot learns to recognize, categorize and manipulate objects with the help of neural networks. There are also projects related to symbol grounding, which, in this context, translates to coupling language with actions.

In the laboratory, Matsumoto's notion of embodiment is connected with other concepts and practices. Consider, for example, the "intelligence" part of "embodied intelligence." In an interview, Matsumoto casually told me that: "arms and hands are about intelligence, and vision [is also about intelligence]" while "legs are about movement." I was puzzled, since this explanation was quite different from anything I had read or heard before. For Matsumoto, apparently, intelligence is about manipulating objects and tools, which is a very particular way of interacting with the environment. At that point, I realized that most of the humanoid robots in his laboratory consist only of an upper-half body.

It became increasingly obvious to me that this seemingly minor aspect of the laboratory's thought style has major consequences in terms of the lab's computational assemblage and research activities: There is a particular way Matsumoto and his lab sees intelligence and this vision is shared by some (but not all) in the Japanese robotics thought collective. The research conducted in the laboratory and the materiality of the research practices are both compatible with *that* vision, and the computational assemblages created in the Matsumoto laboratory therefore embody Matsumoto's vision, both physically (e.g. the legless NAOs) and computationally.

To illustrate how embodied intelligence can be interpreted differently by different thought styles within robotics and how that creates different computational assemblages, I can draw a comparison. In the summer of 2016, I visited Professor Howard Williams' laboratory in central Europe, which focuses on micro- and nano-scale robotics, mainly used for biomedical applications. Williams, an American professor who has worked for over a decade in Europe, is a prominent figure in this research area, nearing retirement. He gave me a laboratory tour, which was in itself extraordinary because it was the only robotics laboratory in which I did not *see* any robots –their robots cannot be seen except under a microscope. He also agreed to sit for an interview in which I got to listen to his story.

Williams told me that he wanted to "make robots under a microscope" that can be injected into a person to perform tasks in the human body. His research hit a wall mainly because developing a minuscule robot is very different from making bigger robots since "the physics of the micro- and nano-scale world operates differently." He also explained that his laboratory looked into how "nature" solved this problem. And indeed, it turns out that nature has many solutions to this engineering problem, provided by various organic bodies. Eventually, they modeled their robots after certain bacteria. Williams' bacteria-inspired robots and Matsumoto's cognitively inspired humanoids are illustrative of quite different computational assemblages that reflect different understandings of intelligence, and differences in research practices including modeling.

Williams' robots replicate the way a certain type of bacteria propels itself in the bloodstream. However, while the bacteria that inspired these roboticists cause diseases in animals and humans, the robot is developed instead to deliver drugs to specified places in human bodies. In contrast, rather than accurately replicating human movement, Matsumoto's humanoids are developed to recognize objects and learn what to do to them: for instance, to recognize a coffee mug, learn how to hold it, and apply that skill to mugs of various sizes and shapes. These humanoids cannot sense the environment as human beings do, and their movements are fairly limited in comparison to an able-bodied human. It is clear, however, that cognitive robotics puts an emphasis on object recognition, decision-making, and similar higher cognitive functions, which are not a concern for micro- and nano- scale robotics. As this shows, the robots of both labs are modeled and designed differently, and their envisioned purposes—working *in* human bodies and working *with* human bodies, respectively—make them as dissimilar as the life forms that they partly represent.

During our conversation, I also asked Williams where he draws the line when it comes to intelligence. He replied that even "viruses have intelligence even though they do not have a single neuron." Compared to Matsumoto's hands-and-vision-centered understanding of intelligence, Williams' take is evidently broader. Centrally, it indicates how greatly understandings of intelligence vary within robotics. The robots they develop as computational assemblages differ not only in terms of form and function; they also embed different views of the traits they attempt to replicate (i.e. intelligent behavior).

Despite their significant differences, Matsumoto and Williams both highlight the importance of embodiment. This suggests that, across subfields of robotics, embodiment in reference to interaction with the real world is widely perceived as a requirement for intelligence. Yet the particularities of interaction are understood differently across thought collectives, even down to individual laboratories. At the same time, while it is true that newly emerging thought collectives, such as cognitive robotics or nano-robotics, draw ideas from other established ones (e.g. neuroscience, biology, developmental psychology), they do not adopt such thought styles wholesale. For example, even though Williams' laboratory frequently collaborates with biologists, and indeed employs some of them, their relation to biology could be described as an *à la carte* adoption of knowledge and methods. And because the mingling of significantly different thought collectives with different thought styles is often project-based, it shapes particular computational assemblages, such as the creation of bacteria-sized robots used to deliver medicine inside a patient's body in Williams' laboratory. In recent years, these kinds of multidisciplinary ventures and collaborations have become increasingly ordinary, generating a corresponding increase in the diversity of thought styles and computational assemblages. In the following section, I examine how these forms of research deal with, or refrain from dealing with, life.

Lifeless Robots

Given the recurrence of the notion of "embodiment" and the prevalence of life-as-inspiration during my fieldwork, in laboratories or at conferences, I was regularly struck by what might be called *the absence of robotic life.* Due to the previously discussed common conception of robots as an analogue for life forms, I had anticipated the robots in my fieldwork to be considered to have at least the potential of being alive. Yet while Matsumoto, Williams and others were clearly fascinated by existing life forms, they were hardly interested in trying to create alternative ones. Indeed, during my months at a laboratory desk, and while participating in every meeting and seminar, I encountered absolutely no talk of "life," —unless, that is, I specifically brought it up.

Matsumoto often described his views by comparing them with those of others, and our conversations about life took a similar form. He briefly described the major positions, first noting that if the definition of life is based on DNA and proteins, then machines cannot be considered alive. If, however, the definition is based on differentiating between self and other—on autopoiesis—or on evolution, then life can be ascribed to machines. This, Matsumoto added, is the stance of ALifers. He then told an anecdote from his graduate student years, a time when the research field of human-robot communication was still considered unusual, but which they were trying out nevertheless. At the time, they conducted surveys about human attitudes to robots alongside their experiments, and found the results were divided into two broad groups: Some people consider robots to be lifelike and attribute emotions to them, while others see them as machinelike. Matsumoto and other researchers then tried to find common denominators that accounted for this difference, yet neither gender nor age worked. *"The criteria for people to accept life as life[like] is quite ambiguous"*, Matsumoto asserted; "people" here referring both to his survey audience and to his academic peers. In other words, since they did not know what is lifelike in a life form, they could not make their robots more lifelike. In this story, Matsumoto is distancing himself from ALife narratives. He is well aware of the popular association of robots with life, and the implications thereof. Yet Matsumoto finds "life" to be much more ambiguous than "intelligence," and thus he is keen to distinguish his own work from those who actively study the former.

Whereas Matsumoto wants to make his robots smarter, he is not concerned about whether they become lifelike *in consequence*. He has no interest in judging whether or not lifelike traits (cognitive, for example) really occur in the robots. What he *is* interested in is how certain mechanisms evolve in living beings, and for that, he uses robots as a "reference." In other words, he uses his robots as mirrors to understand cognition. Furthermore, he does not think robots need to be wholly lifelike –or human; for him "it is alright that robots have limits."

This perspective is prevalent. Indeed, most people in Matsumoto laboratory see their robots or neural networks as mere tools. Robots are registered as laboratory equipment and lumped in together with other equipment such as computers, printers, etc.; and they *are*, in one sense, simply tools that allow graduate students to master machine learning techniques. At the same time, however, it is also fairly common for laboratory members to anthropomorphize the robots, for example referring to one of them as *"koitsu"* (roughly translating as "this guy/fellow"), *"konoko"* (this kid), or *"kare"* (he); but hardly ever as *"kore"* (this thing).

When my informants talk about their ANNs or their robots they may say that it's trying hard *(ganbatteiru)*; and the robots are considered clumsy *(heta)* when they fail. On the one hand, such anthropomorphic language is strikingly common in daily interactions in Japan; yet on the other hand, it is significant that only the robots and the ANNs are anthropomorphized in the laboratory. The robots and ANNs may have qualities that invoke lifelikeness, but not so much as to be taken as a potential life form. Matsumoto himself says that the robots and ANNs sometimes surprise him during experiments. In a sense this is unsurprising, precisely because they are tools meant to create surprises[7]. Nevertheless, when asked directly, my informants are clear that robots are primarily laboratory equipment for making movements.

Indeed, my informants seemed to *particularly* avoid talking about "life" when talking about their tools. Even so, they get their inspiration from, and build their models on the basis of, life forms.

---

[7] ANNs often produce surprising results because of their complexity. Their surprises are very useful feedback for the researchers (See Holland 2014).

Thus, *fragments of life* are rather tenuously embedded into the research in the laboratory and, as I show below, this happens most prominently through modeling practices.

Professor Williams told me that he, as an engineer, likes to break the world into little boxes and write programs in them. This divide-and-conquer strategy is visible in the practices of many roboticists, and it conforms to how my informants are constantly digesting knowledge and methods from other disciplines, including biology and cognitive science, partially integrating it into their thought style and enfolding it in their computational assemblages. "Life" seeps into my informants' research, despite the fact that they are not interested in creating new life forms. In consequence, their research, intentionally or not, also contributes to proliferating understandings of life.

This eclectic integration of different elements into laboratory thought styles relates to the history of robotics and AI research. From the start, robotics was an interdisciplinary research area, and new research areas and the computational assemblages they generate feed on a variety of already existing discourses. Thus, for example, the discourses of cognitive robotics and developmental robotics are entangled, respectively, with those of cognitive science and developmental psychology. Yet, to repeat, the conceptions of those sciences are not fully adopted, but only adapted piecemeal to the extent that particular insights seem relevant and can be computationally modeled. As the whole science is still emerging, the boundaries remain quite fluid. In turn, this means that there is no generally accepted roadmap to designing intelligent robots by using machine learning techniques. The consequence is that even though the robotic platform is shared among researchers, the specifics of computational assemblages continue to diverge between laboratories.

As I discuss in the following section, modeling is a particularly important way in which this differentiation occurs. The variability of modeling after life forms thus allows me to highlight that eclectic adoption of ideas is an important characteristic of these rather *fluid* thought styles, which generate equally fluid and eclectic computational assemblages.

Modeling Within the Computational Assemblage

In a robotics laboratory that focuses on software rather than hardware, most of the daily practices involve sitting in front of a computer: coding, running simulations, etc. When members of the Matsumoto laboratory discuss work, it is almost always in terms of a "model", which is embedded in their codes and simulations and serves as the core of their projects. I had read and been told by informants that in research areas such as bio- or cognitively-inspired robotics, the robots are modeled after organisms or their cognitive traits. Yet, the question of *how to replicate such traits of living beings within a computational assemblage* turned out to be quite complex. Modeling is where my informants' computational assemblages get tied to life forms and it is simultaneously where life-as-a-grand-concept disappears.

Models vary greatly in how they are related to the phenomena to which they are linked, as well as in how they are used. The philosopher of science Ian Hacking describes a triad of phenomena, theory, and models, in which the models play an intermediary role, sitting in between a phenomenon and the theory that "aims at the truth" about it (Hacking 1983: 217). Due to this position, he asserts that "models are doubly models" (Hacking 1983: 216) because they must necessarily at once capture important aspects of the phenomenon-under-investigation and the theory that tries to explain it. Previously, the philosopher Max Black had similarly highlighted that “there is no such thing as a perfectly faithful model; only by being unfaithful in some respect can a model represent its original” (Black 1962: 220). Among the models examined by Black, analogue models, which aim to capture “the structure or the web of relations” of a phenomenon in a different medium (Black 1962: 222) are closest to the models that the roboticists in the lab construct.

To understand the function of models, Williams’s previously quoted description of “breaking the world into little boxes and writ[ing] programs in them” is quite telling. A model of a given life form represents the target system but only in certain specified aspects. For instance, the ambition to explain bird movements can yield numerous models, relating to the flight of a single bird, or that of a flock of birds, or their migratory patterns, nesting behaviors, or mating behaviors—the list goes on. Such models can be combined, but in practice they are usually focused on representing a certain aspect of, say, being a bird. Modeling can thus be seen as an act of breaking apart and reshaping a form of life in order to create a partial reflection of the source material. Since cognition, behavior, or movement have all been understood in diverse ways since long before the emergence of robots, a multitude of models are available. The robots are modeled *after* life forms but more accurately after fragments of life forms.

As mentioned above, I constantly encountered the term “model” during my fieldwork. Just as the design of an embodied agent capable of learning and adapting depends on a particular computational assemblage, it clearly also requires an assemblage of different models and modeling practices. When my informants in the Matsumoto laboratory refer to “*their* model”, it often means their “(machine) learning model" *(gakushū moderu), i.e.* the structure of the neural network they are using. Yet that is not the only model involved in making an embodied agent move. Since the robot moves about in the physical world, a physical model is also necessary. This type of model, sometimes referred to as a kinematic or body model, governs the hardware (the robot's body) and its interaction with the environment (the experiment setting). The humanoids are evidently modeled after human bodies. The resemblance may differ in terms of form and function; for instance, NAOs are smaller in scale than humans[8] and have very limited bodily capabilities. Yet the limitations imposed on them by their humanoid bodies, as asserted by Kuniyoshi, allow them to move in a humanlike manner which is regulated by the body model.

Furthermore, in terms of software, the neural network is itself modeled crudely after the human brain with nodes representing neurons. But not only are there several types of ANNs, many tasks also require combining multiple ANNs, which means that every architecture is different. According to my informants, the construction of a model is often a process of trial and error that goes on until desirable results, such as correctly sensing the experimental setting and producing the appropriate action, are obtained. In addition, the "cognitive" aspect of robotics sometimes requires the input of a computational model of a form of human behavior or a cognitive trait, which is imported from cognitive or behavioral sciences. But since the physical and cognitive models are largely implicit, the notion of the "model" in Matsumoto laboratory is more or less synonymous with the ANN designs, which to some degree are inspired by understandings of human cognition.

---

[8] A full body NAO is 574 millimeters tall.

Movement, perception, recognition, and learning are phenomena that have long been modeled by different research fields. What this means, *pace* Hacking's argument, is that the models that populate the computational assemblages of Matsumoto laboratory also implicitly theorize aspects of embodiment, learning or cognition in a range of different, and sometimes incompatible, ways.

Yet, while models proliferate in the laboratory and thus create complexity at the level of the computational assemblage, when it comes to individual models, Matsumoto advocates simplicity. That is, he tends to prefer the simplest models that behave like the target system with *no explicit mechanism* for the generation of the behavior. He emphasizes that he does not want the robot to be ordered to do things; rather, the behaviors should appear "as a result" – i.e. in an emergent manner. The principle is that even if the phenomena—or the behavior—are complex, the mechanism that generates them should be simple.

The reason for this preference relates to Matsumoto's awareness that phenomena are often explained differently in different fields. Yet no field has a perfect understanding of human cognition. According to Matsumoto, for example, physics and mathematics are not flexible enough to capture important dimensions of cognition. He rather favors neuroscience; however, this field has changed a lot over the last couple of decades and consensus is still lacking. The same can be said about fields including sociology, psychology and philosophy. Thus, Matsumoto's preference for simplicity can be understood in light of the fact that he is trying to design an intelligent agent without a definite roadmap, having to rely instead on many patchy and incompatible ones that furthermore keep changing.

Matsumoto is thus very well aware that practically all the keywords embedded in the thought style and computational assemblages of his laboratory originated in different disciplines and come with their own theoretical baggage. Within the thought style of the laboratory, there are traces of many different disciplines and ways of understanding. As suggested by Fleck, these entanglements of knowledge within a thought style are not the result of rational planning and selection. Matsumoto explained that he simply takes knowledge or models from different disciplines according to the particular phenomenon he is studying. If a certain behavior is explained in simpler terms by neuroscience, for instance, then he will choose a neuroscience-based model, but in another case, he might opt for a model based on behavioral sciences.

In sum, any robotic experiment to replicate a human behavior draws on models that represent selected aspects of the phenomena. We might say that Matsumoto is pragmatic and eclectic both in terms of how he organizes his laboratory and research and in his grab-bag approach to foreign concepts. The robots emerge from the bits and pieces of life, machines, data, models, theories, and practices that make up Matsumoto's laboratory.

Conclusion

Robots today are technologies marked by their likeness to us. Although they all resemble us—humans and non-human life forms—to a degree, some robots are specifically developed to mirror us, to show us what we are like. In this paper, I have examined how roboticists who work to achieve such mirroring conceive of and operationalize embodiment, intelligence, and life. Doing so, I have described how their thought styles draw upon a range of ideas from adjacent disciplines, which are combined in a piecemeal manner for particular research purposes such as replicating certain cognitive traits or behaviors in robots. The fluidity and variability of roboticists' thought styles are a consequence of partial adoptions and eclectic mixture of ideas from fields including neuroscience, behavioral sciences, and biology. In turn, these thought styles shape widely differing computational assemblages; made of silicon, metal, scientific facts, theories, models, and discourses, as well as lots of 1s and 0s. It is in these assemblages that the process of mirroring occurs. As mirrors, they reflect everything that is put in to make them: particular fragments of life forms as well as working practices and knowledges, temporarily patched together by a thought style.

Since researchers borrow inspiration and practices from various disciplines, whether life sciences or behavioral sciences, what they replicate in their robots can be completely different. Even in the case of a specific behavior, there are multiple ways to recreate it in a robotic body. On the one hand, then, thought styles locally determine how roboticists perceive the target organism and how they replicate them. On the other hand, the lack of a generally accepted roadmap for design means that both thought styles and computational assemblages proliferate and diverge.

Existing forms of life not only provide insight into concepts such as intelligence and embodiment, but also offer practical solutions to engineering problems. The end result—a robot—is expected to behave somewhat like the organism that it is made after. While roboticists thus integrate fragments of life forms into their robots, they themselves also add to the multitude of existing understandings of life.

Yet, rather than seeing robots as either alternative or artificial life forms, my informants view them as offering *depictions* of life. As for life itself, or even humanlike intelligence, these goals seem too elusive to pursue. Indeed, Matsumoto, who has been working on replicating human cognition for over 25 years, claims that he will not see human-like artificial intelligence in his lifetime, adding: “Do not underestimate human [intelligence]”.

References

Black, Max.
1962 *Models and Metaphors: Studies in Language and Philosophy.* Ithaca, NY: Cornell Univ. Press.

Čapek, Karel.
1921 *Rossum's Universal Robots*. Translated from the Czech language by David Wyllie. http://ebooks.adelaide.edu.au/c/capek/karel/rur/. Accessed Feb 22, 2018.

Doyle, Richard.
1997 *On Beyond Living: Rhetorical Transformations of the Life Sciences.* Stanford, CA: Stanford University Press.
2003 *Wetwares: Experiments in Postvital Living.* Minnesota, London: University of Minnesota Press.

Fleck, Ludwik.
1935 (1981) *Genesis and Development of a Scientific Fact.* Chicago and London: University of Chicago Press.

Hacking, Ian.
1983 *Representing and Intervening: Introductory Topics in the Philosophy of the Natural Sciences*. Cambridge: Cambridge University Press.

Helmreich, Stefan.
2009 *Alien Ocean: Anthropological Voyages in Microbial Seas.* Berkeley, Los Angeles and London: University of California Press.
2011 “What Was Life? Answers from Three Limit Biologies”, *Critical Inquiry* 37(4): 671-696.

Holland, John H.
2014 *Complexity: A Very Short Introduction,* Oxford: Oxford University Press.

Johnston, John.
2008 *The Allure of Machinic Life: Cybernetics, Artificial Life and the New AI.* Cambridge, Massachusetts and London, England: The MIT Press.

Kawada Industries Inc.

2018 "Nextage Works Side by Side with People" http://nextage.kawada.jp/en/ Accessed Jan 29, 2018.

Kluckhohn, Clyde.

1949 *Mirror for Man: The Relation of Anthropology to Modern Life.* New York: Mc Graw-Hill.

Kubo Akinori 久保明教

2015 『ロボットの人類学——20 世紀日本の機械と人間』(*Robotto no Jinruigaku: Nijusseiki Nihon no Kikai to Ningen*). 京都：世界思想社 (Kyoto: Sekaishisousha).

Kuniyoshi Yasuo 國吉康夫

2008 「知的行動の発生原理」(*Chiteki Kōdō no Hassei Genri*), 『人工知能』(*Journal of Japanese Society for Artificial Intelligence*) 23(2): 283-293.

Masood, Ehsan.

2009 *Science and Islam: A History,* London: Icon Books.

Pfeifer Rolf and Josh Bongard.

2007 *How the Body Shapes the Way We Think: A New View of Intelligence.* Cambridge, Massachusetts and London, England: MIT Press.

Rethink Robotics.

2018 "With our Smart, Collaborative Robot Pioneer" http://www.rethinkrobotics.com/baxter/ Accessed Jan 29, 2018.

Robertson, Jennifer.

2007 "Robo Sapiens Japanicus: Humanoid Robots and the Posthuman Family", *Critical Asian Studies* 39(3): 369-398.

2010a "Gendering Humanoid Robots: Robo-Sexism in Japan", *Body and Society* 16(2): 1-36.

2010b "Robots of the Rising Sun" *The American Interest* 4(1): 60-72.

2011 "Rubble, Radiation and Robots" *The American Interest* 7(1): 122-126.

2014 "Human Rights vs. Robot Rights: Forecasts from Japan" *Critical Asian Studies* 46(4): 571-598.

Searle, John.

1984 *Minds, Brains and Science: 1984 Reith Lectures.* Cambridge, Mass: Harvard University Press.

Softbank Robotics.

2018 "Who is NAO?" https://www.ald.softbankrobotics.com/en/robots/nao

Accessed Jan 29, 2018.

Stableford, Brian.

2006. *Science Fact and Science Fiction: An Encyclopedia.* New York, NY: Routledge.

Suchman, Luch.

2007 *Human-Machine Reconfigurations: Plans and Situated Actions, 2nd Edition.* New York: Cambridge University Press.

Thornbury, Barbara E.

1992 "Puppets on Strings and Actors on Floats: Japan's Traditional Performing Arts in a Festival Setting", *The Journal of the Association of Teachers of Japanese*, 26(2): 181-192.

Zuboff, Shoshana.

1988 *In the Age of the Smart Machine: The Future of Work and Power.* New York: Basic Books Inc. Publishers.