IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 1 December 2022, accepted 15 February 2023, date of publication 20 February 2023, date of current version 23 February 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3246840

THEORY

Developmental Autonomous Behavior: An Ethological Perspective to Understanding Machines

SATORU ISAKA, (Member, IEEE)
Vision Del Mar, LLC., San Jose, CA 95120, USA
e-mail: sisaka@visiondelmar.com

ABSTRACT Developmental autonomous behavior refers to the general ability of a machine to acquire new skills and behavior from its birth to maturity on its own without human intervention. This article describes the principles of behavior development in machines, providing a practical framework to analyze and synthesize machines with developmental capabilities. Inspired by biological views of behavioral causation, the work emphasizes principled explanations to, not only the "how" question on mechanisms but also the "why" question on causation of behavior development. This ethology-oriented perspective offers a renewed opportunity to construct a theoretical framework from the ground up, overcoming the age-old problems of intrinsic motivation and symbol emergence in autonomous machines. One of the key contributions of this article is the logical explanation of why and how value systems drive successive development of memory functions, resulting in progressive changes in behavior from innate reflexive to episodic, procedural, and autonomic behavior. Another notable contribution is the logical and plausible explanation of why and how a physical sensorimotor system becomes a symbol processor, fostering conceptual and social behavior development. This article provides an extensive review of prior research, followed by detailed descriptions of the causality and mechanisms of behavior development, and concludes with discussions on criticism, future work, ethics, and system architecture.

INDEX TERMS Autonomous machines, developmental behavior, intrinsic motivation, machine learning, robotics, symbol emergence, value systems.

I. INTRODUCTION

" Human *beings, viewed as behaving systems, are quite simple. The apparent complexity of our behavior over time is largely a reflection of the complexity of the environment in which we find ourselves.*" Herbert Simon [1].

A. MOTIVATION

"*Are you real?*" a man asks. "*If you can't tell, does it matter?*" This poignant reply comes from a character in the science fiction television series "Westworld" [2]. The character is either a robot or real human; we can't tell which. The scene evokes an eerie feeling that, some day in the future, we may live in a world where machines become so assimilated that we can no longer tell the difference between humans and machines. Such a world is already here. We live on vital infrastructures: energy, water, transportation, finance, communication, all of which are regulated by machines today, at least partially if not fully. These machines are increasingly becoming autonomous, meaning they are operating automatically on their own. We are increasingly becoming dependent on them in a world where unseen and unforeseen events could occur at any minute. So, it does matter that we recognize autonomous machines and their behavior to keep them accountable, especially the new generations of machines that learn.

There is no definitive definition of what an autonomous machine is [3] and [4]. Autonomy means a self-governing state [5]; thus, an autonomous machine implies a self-governing apparatus. Depending on what the apparatus is

The associate editor coordinating the review of this manuscript and approving it for publication was Alessia Saggese.

and what it is governing, the meaning of autonomous machines varies. The machine's purpose, functions, and behavior dictate its existential definition. For example, National Institute of Standards and Technology (NIST) defines autonomy of unmanned systems (UMS) as follows:

"*A UMS's own ability of integrated sensing, perceiving, analyzing, communicating, planning, decision-making, and acting/executing, to achieve its goals as assigned by its human operator(s) through designed Human-Robot Interface (HRI) or by another system that the UMS communicates with. UMS's Autonomy is characterized into levels from the perspective of Human Independence (HI), the inverse of HRI.*" [6].

According to this definition, autonomy is defined by the inverse of HRI, implying that if we eliminate the human-robot interface (i.e., denominator becomes zero), then the machine is given a unique level of complete autonomy. At this level, the machine is left with its own ability to maintain its existence. If such a machine exists, how will it acquire skills to survive?

Autonomous behavior used to be the kingpin of living organisms and our source of fascination. For example, when 320,000-year-old stone tools were found in northern Africa, the archaeologists discovered the evidence of ingenious behavior of early humans innovating new tools to adapt to their changing environment [7]. The tools were smaller, and the blades and points were more precise than their previous generations' bulky hand axes. They used raw materials like black obsidian, which they could not have accessed unless they had a trade network with other remote communities. These early humans survived and thrived by adapting to the environment by exploring and exploiting the resources available.

Autonomous behavior can also be observed in a microscopic environment. Viruses, an acellular organism, invade their hosts, change genetic structures, and take over cellular metabolism of their hosts [8]. Even though biologists in general do not consider viruses as living organisms [9], they exhibit the process of responding and adapting to the environment to survive and thrive.

These are the exemplary evidence of autonomous behavior. Organisms as behaving systems interact with the environment, manipulate the resources available, and act autonomously to respond and adapt to the environment. Their physical capacity and features vary significantly from one species to another and within the same species, making our biological universe enormously diverse and complex.

When considering nature's autonomous behavior, a distinction must be made between reactive and proactive behaviors. The former represents behaviors in response to physical changes in the environment as a feedback system by relying on the built-in mechanisms of the body. On the other hand, proactive behavior as we humans exhibit daily is not only reactive to physical changes but also proactive by anticipating future changes. It represents deliberate behavior in anticipatory, conceptual, and potentially collaborative actions over a longer time span than the reactive behavior. Because of the latent nature, proactive behavior is more difficult to observe and understand than reactive behavior.

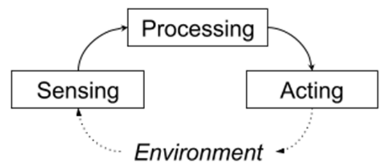

FIGURE 1. Basic components of machines exhibiting reactive behavior.

Essential functions that enable reactive behavior are relatively well understood. They can be abstracted to three components: sensing, acting, and processing (Fig. 1).

Sensing is for obtaining information about the external and internal conditions; acting is for actuating their means to interact with the environment; processing is for linking sensing and acting for situated behavior. Situated means that the behavior depends on the situation. These three components can be hardwired to autonomously execute reactive actions in response to sensed inputs. With this structure, it is possible to construct a machine as an electro-mechanical circuit with a set of behavior rules in the processing component to regulate actuators in response to sensing data.

Unlike reactive behavior, essential functions that enable machines to exhibit proactive behavior are not well understood. Human behavior is highly circumstantial because our inner drive to act, so-called motivation, is circumstantial. Our behavior is also developmental. It changes as we grow by learning and acquiring new experience by interacting with the environment. Human behavior is therefore both reactive and proactive, and overall developmental in nature. If we were to build a machine that can exhibit all these different kinds of behaviors, is the structure of Fig. 1 sufficient? If so, what does the processing component look like and how does it work to support developmental behavior?

B. HISTORICAL PERSPECTIVES AND PROGRESS

In the fields of artificial intelligence, cognitive science, and robotics, the processing component is universally regarded as the brain of the machine, taking inspiration from the human and animal brains. Over the years, a variety of machines have been constructed in attempts to replicate some of the cognitive features of the brain. To what degree these machines exhibit cognitive abilities is a subject of debate. One of the reasons why we have difficulty, aside from the apparent complexity of the subject matter, is attributed to considerable degrees of freedom in interpreting what cognition is. To provide the context behind the problem statement of this article, this extensive review offers a concise, unified story of the historical progress, tracing back to the roots of major theoretical frameworks and perspectives.

1) THE ORIGIN AND CLASSICAL FRAMEWORKS

In the beginning, from what we know from written records, the frameworks of thoughts on human behavior and cognition may have originated from a metaphysical question, *why is there something instead of nothing* [10]. Attempts to derive a rational answer only yield more questions than answers, thereby creating a variety of narratives such as universal laws and teleology [11]. One of the derivative questions is about our knowledge, where does it come from, experience or inference. In the 4th century BC, Aristotle laid down the framework of analytics as a logical process of dealing with facts from experience to derive a posteriori conclusion of universal laws [12], [13].

The perspective of experience and inference being the source of knowledge and behavior is shared among various cultures. Ancient Hindu philosophy teaches that cognition emerges from perception and volition when a desire crystallizes, which leads to behavior [14]. According to Avicenna, the 10th century Persian philosopher, "*the human intellect at birth is a pure potentiality that is actualized through education. Knowledge is attained through empirical familiarity with objects in this world from which one abstracts universal concepts*" [15]. Despite the apparent agreement in the process of knowledge acquisition in early frameworks, many thinkers subsequently picked up on specific attributes on inference and experience to construct various frameworks, creating divisional schools of thoughts: rationalism and empiricism.

In 1640, Rene Descartes examined human behaviors and recognized the conceptual distinction between the mental and physical domains [16]. Descartes wrote in his replies to his fellow philosopher Arnauld, "*When someone falls and holds out his hands so as to protect his head, he isn't instructed by reason to do this*" [17]. Even though Descartes acknowledged that not all human behavior arises from rational deliberation, he concluded that rational deliberation must take place in the mind that is distinct from the body. This idea is often referred to as mind-body dualism. Because of the emphasis on rationality, Descartes' framework is often categorized as rationalism [18].

In 1689, John Locke articulated his thoughts on the human mind in his book *An Essay Concerning Human Understanding* [19]. He claimed that the mind produces knowledge by putting together simple and complex ideas, building relationships, and generalizing them by abstracting out particulars [20]. In essence, Locke thought that human thinking is based on abstract ideas derived from our experience. Because of the emphasis on empirical observation and experience, Locke's framework is often categorized as empiricism.

In 1781, Immanuel Kant described his thoughts on human behavior in *The Critique of Pure Reason* [21], which synergizes the two frameworks of empiricism and rationalism. He argued that "*human understanding is the source of the general laws of nature that structure all our experience; and that human reason gives itself the moral law, which is our basis for belief in God, freedom, and immortality.*" [22].

In *The Critique of Practical Reason* (1788), Kant describes human autonomy by defining the causal relationship between actions and principles [22]. Kant argues that human actions are not directly caused by desires, but by principles (maxims) that specify the rule or policy of actions. According to Kant, there are two kinds of principles we act on: material and formal principles. Material principles describe how one acts to satisfy desires. Formal principles describe how one acts without referring to any desire. Kant argues that we are free in the sense that we can control our desire and choose our principles. However, our actions are not free or autonomous if we choose to act on material principles only. This is because actions based on material principles are following the law of nature in us, instead of giving ourselves the law by formal principles. Therefore, the only way to act freely in the sense of autonomy is to act on formal principles by identifying with a rational self, which means to act morally [23].

In 1878, Hermann von Helmholtz offered a deeper understanding of observation, sensation, perception, and cognition. He argued that we determine spatial relations among objects, not by physical sensation alone, but by learning to interpret signs from the sensation [24], [25], [26]. According to this so-called sign theory, observation does not give us direct copies of objects but signs of the objects from physical sensation. The fact that objects up close appear larger than the same objects far away is because perceptions of the objects are in fact signs. With signs we derive the law of causality from observed regularities that causes will be followed by effects. Helmholtz used this framework to derive many practical principles in fluid dynamics, optics, and conservation of energy. His "free energy" principle inspired numerous works on machine learning such as the Helmholtz Machine [27], [28] and Active Inference [29], [30].

2) MODERN FRAMEWORKS

The 20th century saw two broad scientific fields emerge, changing the way we view cognition and behavior for both living and nonliving things. Neuroscience emerged from empirical analyses and syntheses of anatomy, embryology, physiology, pharmacology, and psychology [31]. Computing and information science emerged from empirical analyses and syntheses of physics, dynamics, electronics, communication, control, and computer systems. The perspectives of both fields began to catalyze an idea of artificial machinery behaving like living things. Some notable developments occurred soon after the World War II ended.

In 1948, Alan Turing described in his report *Intelligent Machinery* a machine that exhibits intelligent behaviors by applying a human-analogous teaching process of rewards and punishment to machines [32]. Turing concisely articulated his ideas by first categorizing four types of machinery: discrete controlling, discrete active, continuous controlling, and continuous active machinery. Discrete machinery is a machine whose states are described as a discrete set;

continuous machinery operates in a continuous manifold. Controlling machinery is a machine that only deals with information without producing any physical effect; active machinery causes physical effects.

By focusing on discrete controlling machinery, Turing elaborated his thoughts on intelligent behavior in relation to brains and machines. He brought clarity to the notion of intelligence itself as an emotional concept, a subjective interpretation of the properties of the object. He wrote, "*If we are able to predict its behavior or if there seems to be little underlying plan, we have little temptation to imagine intelligence. With the same object therefore it is possible that one man would consider it as intelligent and another would not; the second man would have found out the rules of its behavior.*" [32] He prescribed a simple experiment to test this idea, which later became known as the Turing Test.

Also in 1948, Norbert Wiener theorized in his book *Cybernetics* that all intelligent behavior was the result of feedback mechanisms, and proposed a study of how humans, animals and machines control and communicate with each other [33]. The term "cybernetics" comes from the Greek word "*kyvernitis*", the steer master of a boat [33]. If a boat is aimed to the right or left, the steer man turns it to the left or right to correct its course. By comparing the actual direction with the intended direction, and by applying the negative correction, the steer man works as a negative feedback controller of the boat. Wiener's insights inspired active research on dynamic systems and automation in a variety of applications.

Also in 1948, William Grey Walter built autonomous robots *Elmer* and *Elsie* [34], [35]. With motorized wheels, touch and light sensors, and electrical control circuits, the robots exhibited complex behavior resembling the behavior of insects and animals. Walter's demonstration provided valuable insights that complex behaviors do not necessarily arise from complex processing [36]. His work inspired active research on robotics, known as behavior-based robotics (BBR) [37]. A good example of BBR is the commercially ubiquitous vacuum cleaning robots.

Also in 1948, Claude Shannon published two articles in the Bell System Technical Journal, combined known as *A Mathematical Theory of Communication*, in which he introduced a quantity that measures how much and at what rate information is produced by a process [38]. Known today as Shannon Entropy, he defined the quantity based on probabilities of the outcome of a random process. Shannon Entropy plays a central role in information theory as measures of information, choice, and uncertainty.

In 1949, Donald Hebb in his book *The Organization of Behavior*, proposed a theory of adaptation of brain neurons during the learning process [39]. Known today as Hebbian learning, his theory inspired active research on artificial neural networks that proliferated commercially and academically in recent years.

In 1956, perhaps motivated by all these recent developments in computing, information, and neuroscience, John McCarthy organized a workshop at Dartmouth College with a group of scientists to discuss computing machines and intelligence [40]. He introduced the term "*artificial intelligence*" in his proposal, thus making the Dartmouth Workshop the founding event for the field of artificial intelligence, also known as A.I. [41].

The workshop was driven by a conjecture that "*every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it*" [40]. Based on the framework, many cognitive architectures were introduced, including SOAR [42], ACT-R [43], EPIC [44], and ICARUS [45]. Because of the emphasis on symbol processing [46], the framework became known as symbolic AI.

In 1957, Frank Rosenblatt introduced perceptron, a machine learning algorithm inspired by Hebb's work that classifies signals based on artificial neural networks [47]. A formal analysis of perceptrons by Minsky and Papert [48] was unfortunately miscited and involved in controversy but nonetheless, their pioneering works paved ways for deeper understanding of learning behavior and inspired subsequent proliferations in artificial neural networks. Because of the emphasis on parallel connections of simulated neurons, the framework became known as connectionist AI.

Symbolic AI and connectionist AI stood rather independently from each other, like rationalism versus empiricism in the 17th century. The target of criticism against symbolic AI was the isolated treatment of symbolic processes without natural synergy with the physical environment. There were research efforts to combine symbolic AI and dynamic systems. For example, from 1966 to 1972, the Artificial Intelligence Center of Stanford Research Institute (now SRI International) developed an autonomous robot called Shakey [49]. The project was the first of its kind to combine logical reasoning and physical action by bringing multidisciplinary research from computer vision, natural language processing, information search, and robotics. Shaky is considered the first mobile robot to reason its own actions by breaking down commands into basic chunks by itself [50].

By recognizing the origin of symbolic AI and its targeted machinery, the perceived isolation of symbol processing is understandable. Based on Turing's classification of machinery [32], symbolic AI is intended to deal with discrete controlling machinery. Robotics and systems theories such as Wiener's cybernetics and Walter's robots deal with continuous controlling and active machinery. Neural circuitry such as Hebb's theory applies to continuous controlling machinery. In other words, these research efforts were targeting different classes of machinery from different perspectives and purposes. Under such circumstances, direct comparison of concepts and principles is misguided without first establishing a unifying framework. For this reason, it is worth highlighting general, multidisciplinary frameworks that appeared during this period.

In 1956, Ashby brought the concepts of information theory to cybernetics and introduced the general theories of regulation, so-called the law of requisite variety and the theory of minimum entropy [51]. The law of requisite variety states that "*only variety in R (regulator) can force down the variety due to D (disturbance); variety can destroy variety*" [52]. The theory of minimum entropy states that the optimality of a regulator is achieved by minimizing the entropy or surprise of the outcomes. The entropy in this context refers to Shannon Entropy in information theory. Ashby's theories generally imply that an optimal regulator is one that brings a system to the state of homeostasis, where its key characteristics are maintained and is resilient to external disturbances.

In 1970, Ashby and Conant presented the idea of error-controlled versus cause-controlled regulations [53]. Classical feedback controllers are a type of error-controlled regulation. Model-based regulators are a type of cause-controlled regulation. Conant and Ashby argue that error-controlled regulation is an inferior method because the entropy of the outcomes cannot be reduced to zero. On the other hand, cause-controlled regulation is a superior method because the entropy can be reduced if the regulator can mirror the system itself, implying that a good model of the environment is a necessity for optimal regulation and homeostasis. The framework on homeostasis and optimization influenced the subsequent research framework of active inference [29].

In 1972, A. Harry Klopf made a distinction between homeostasis and heterostasis in living systems and assumed that adaptation occurs by seeking maximal conditions as a goal rather than seeking steady state conditions [54]. He introduced a theoretical framework based on a heterostatic neuronal model that unifies neurophysiological, psychological, and sociological properties of living adaptive systems. Klopf made connections between brain functions and machines by highlighting three brain regions (limbic and hypothalamus, midbrain and thalamus, and neocortex), to offer a basis for the synthesis of adaptive intelligent machines. Klopf's framework on heterostasis and optimization influenced the subsequent research framework of reinforcement learning [55].

In 1975, John Holland in his book *Adaptation in Natural and Artificial Systems* showed how a computational model of evolution and natural selection, so-called genetic algorithm, exhibits unique self-organization characteristics [56]. A new variant called genetic programming was introduced by his students, Goldberg [57] and Koza [58] in the 1980's, inspiring new approaches in machine learning. The popularity also brought renewed attentions to previous research in the 1960's in evolutionary strategies by Rechenberg [59] and Schwefel [60] and evolutionary programming by Fogel et al. [61]. The field has since grown inclusively as evolutionary computation, collectively exploring the biologically and computationally synergistic approaches to complex problem solving in optimization, clustering, and code generation.

In 1978, Gerald Edelman introduced a biologically inspired theory of adaptive behavior, called the neuronal group selection (TNGS) [62]. TNGS posits that adaptive behavior of organisms is the result of selection in somatic time among synaptic populations. TNGS was applied to construct a cognitive architecture called synthetic neural modeling (SNM). SNM was then implemented in a series of autonomous robots called "Darwin" [63], [64], [65].

In 1983, A.G. Barto, R.S. Sutton, and C.W. Anderson showed problem-solving capacity of single neuron-like elements by applying classical and operant conditionings of animal behaviors [66]. This pioneering work continued to expand and subsequently became the foundation of reinforcement learning, providing an alternative learning method to supervised and unsupervised learnings of artificial neural networks.

In 1986, Rodney Brooks introduced the theory of subsumption architecture [67], an influential design structure for autonomous machines. In essence, the subsumption architecture follows the principles of sensory-response systems without symbolic mental representations of the world. This idea is closely related to behavior-based robotics, inspired by Grey Walter in the 1950's. Brooks' robot design and perspective inspired active research on autonomous machines that could carry out complex tasks in challenging environments such as planetary exploration and military applications.

The mid-20th century also saw significant outputs in the field of experimental psychology. Examples include the law of effect by Thorndyke [68], classical conditioning by Pavlov [69], latent learning by Brodgett [70], operant conditioning by Skinner [71], behavior development by Piaget [72], drive theory by Hull [73], sensory preconditioning by Brogden [74], vicarious trial and error by Muenzinger [75], and cognitive maps by Tolman [76]. Unfortunately, the field was marred by yet another dogmatic framework battle between behaviorism and cognitivism in psychology. It became vogue to place more focus on cognition than behavior, as the mind is considered an information processor for perception, attention, thinking, and consciousness [77], [78], [79].

3) CONTEMPORARY FRAMEWORKS AND TRENDS

The late 20th and early 21st centuries saw the proliferation of affordable, high-power hardware and software computing tools, resulting in an explosive growth, and accelerated expansion and dissemination of technical knowledge at a global scale. Particularly benefited from this trend are neural networks and robotics.

In 2006, a fast-learning algorithm for deep neural network architecture was introduced by Hinton and his colleagues [80]. Because of the demonstrated effectiveness in pattern recognition in image data, coupled with a high-performance computing environment, the approach was immediately expanded and applied to a variety of applications in object recognition and speech recognition. Commercial products were developed and marketed as ''smart'' products. A.I. quickly became a household name.

In robotics, autonomous robots became ubiquitous due to the improved algorithm and software, coupled with high-performance hardware components. This is evident in the proliferation of robots in warehouse fulfillment centers across the globe. These robots include automated guided vehicles, automated storage and retrieval systems, collaborative robots, articulated robotic arms, and goods-to-person technology [81].

Over the years, research efforts in cognition have shifted from the abstract symbolism to embodied beings [82], [83]. An embodied being means that cognition is an activity that takes place within a physical body, and that cognition arises because of physical interaction between the body and environment. This is the same thought framework in ancient philosophy as described earlier. Today however, few elaborate on physical elements of embodiedness [84]. The physical elements in cognitive system research are largely limited to neural circuitry in brains. According to the surveys by De Garis et.al. [85] and Goertzel et.al. [86], the current state of the art research on cognitive architectures and large-scale brain simulations still has a long way to go to understand our brains, not to mention artificially achieving human level intelligence.

This is perhaps a testament to how complex the bodies and brains of animals and humans are. Brains do not stand alone; they exist to support the connected bodies. Humans are equipped with millions of sensory receptors to sense the external environment and internal states of the organs and frames, via olfaction, gustation, vision, hearing, equilibrium, and somato-sensory receptors, including mechanoreceptors, thermoreceptors, proprioceptors, pain receptors, and chemoreceptors [31].

To provide movement, support, and stability in the environment, humans are equipped with complex musculoskeletal systems, consisting of bones, muscles, cartilage, tendons, ligaments, joints, and miscellaneous connective tissues [87]. Because of this, human brains are necessarily complex, made of billions of cerebral neurons [31], to control the enormously vast and complex parts. When comparing an artificial machinery to such complex structures and functions of humans, a great sense of humility is in order.

Recently a new research field called ''developmental robotics'' emerged [88], [89], [90], [91], [92], [93], [94]. This interdisciplinary research emerged ''*as a reaction to the inability of traditional robot architectures to scale up to tasks that require close to human levels of intelligence*'' [95]. The amount of engineering and programming requirement prohibitively escalates to construct a machine with an increasingly sophisticated cognitive capability. In addition, the more we try to hard-code intelligence in a machine, the potential of introducing more bias and shortcomings increases. The most representative framework of developmental robotics is demonstrated in mobile robots SAIL and Dav by Weng [89]. The field is at an early stage, facing many difficult issues in its attempts to scale up to human level intelligence.

4) DESIGN PRINCIPLES OF CONTEMPORARY FRAMEWORKS

Based on the long historical progress on intelligent and cognitive machines, researchers extracted key insights and wisdom for building such machines and summarized them in the form of design principles. Krichmar and Edelman [96] and Krichmar and Hwu [97] suggest the following design principles for brain-based machines:

- *a brain-based machine should incorporate a simulated brain with detailed neuroanatomy and neural dynamics that controls behavior and shapes memory,*
- *it should organize the unlabeled signals it receives from the environment into categories without a priori knowledge or instruction,*
- *it should have a physical instantiation, which allows for active sensing and autonomous movement in the environment,*
- *it should engage in a task that is initially constrained by minimal set of innate behaviors or reflexes,*
- *it should have a means to adapt the device's behavior, called value systems, when an important environmental event occurs, and*
- *it should allow comparisons with experimental data acquired from animal nervous systems.*

Pfeifer et al. [98] suggest the following design principles for autonomous cognitive agents:

- *agents must be designed for ecological niche, tasks, and agent itself,*
- *agents must be embodied, autonomous, self-sufficient, and situated,*
- *agents' intelligence must emerge based on a large number of parallel, loosely coupled processes that run asynchronously,*
- *all intelligent behavior must be conceived as sensory-motor coordination,*
- *design must be parsimonious and exploit the ecological niche,*
- *agents must be designed with partial overlap in functionality in subsystems,*
- *the complexity of the agent must match the complexity of the task environment, and the physical and neural dynamics must be balanced, and*
- *agent behavior must be motivated on a value system.*

Cangelosi and Schlesinger [93] suggest the following design principles for developmental robots:

- *Development as a dynamical system,*
- *Phylogenetic and ontogenetic interaction,*
- *Embodied and situated development,*
- *Intrinsic motivation and social learning,*
- *Nonlinear, stage-like development, and*
- *Online, open-ended, cumulative learning.*

These principles are notably similar by sharing many attributes. One of them that particularly stands out is value system and intrinsic motivation. Let us reflect on what Turing wrote in 1948. "*If the untrained infant's mind is to become an intelligent one, it must acquire both discipline and initiative. … discipline is certainly not enough in itself to produce intelligence. That which is required in addition we call initiative. … Our task is to discover the nature of this residue as it occurs in men, and to try and copy it in machines*" [32].

What Turing refers to as initiative is what the design principles call for as intrinsic motivation and value system. A popular supervised training approach in machine learning is to feed a large volume of data to train artificial neural networks. This is discipline-based learning. Another popular approach, reinforcement learning, is to optimize a reward function that represents specific tasks to be solved. Because the data and optimality are defined and supplied by humans, machines operate for a reason external to the machine. In other words, the initiative is extrinsic, not intrinsic. Machines are trained but not intrinsically motivated in learning.

What is intrinsic motivation? This question has been heavily investigated and debated over the years. See Ryan and Deci [99] for their well-cited definitions and treatments in psychology. However, as Oudeyer and Kaplan [100] point out, current definitions and treatments are not satisfactory for computational models. They wrote, "*the most pragmatic approach to intrinsic motivation from a computational point of view is maybe to avoid trying to establish a single general definition*" [100]. The principled treatment of intrinsic motivation is thus a fundamentally important yet unsolved topic for cognitive systems and developmental robotics.

C. PURPOSE, OUTLINE, AND APPROACH OF THE ARTICLE

The purpose of this article is to derive logical and plausible answers to why and how a machine can develop new behavior on its own without human intervention. Instead of isolating and zooming directly into the topic of intrinsic motivation, this article takes a broader scope of behavior development. Something must drive a machine to behave. To understand what it is that drives behavior, this article first defines and analyzes what behavior is, then explores its causality and mechanism.

Following the introduction, chapter II of this article analyzes developmental behavior in humans and machines and abstracts the categories of ontogenetic behavior necessary for developmental machines. Chapter III analyzes the process of behavior development. It identifies the causal factors and mechanisms of behavior development and transition. The results of the analyses are summarized as the principles of developmental autonomous behavior. Chapter IV addresses criticism, future work, ethics, and system architecture. The article concludes with a brief remark in Chapter V.

The approach taken in this article is systems science. Things and events are viewed and treated in terms of systems and processes. A process is a series of steps taken to proceed to the next end point. A system is a set of things working together to bring a process to the next end point. Processes are therefore realized by systems.

Human brains and bodies are extremely complex systems. There is no attempt or claim made in this article about replicating human behavior in machines. The word "cognition" is reserved exclusively for a mental process of humans and animals, not of machines in this article. Related phrases such as "cognitive system" and "cognitive architecture" may appear in reference to the relevant works by others in literature. Words that describe phenomena of organisms such as "sentience", "consciousness", and "belief" do not appear in this article because they are outside the scope of this article.

To analyze behavior, the word "motivation" must be clearly defined. This specific terminology, particularly in the form of "intrinsic motivation" has been central to certain frameworks yet their definitions and treatments are unclear.

Motivation is neither an observable entity nor a thing that engineers can build in a machine. Motivation is merely *a reason that causes actions*. Such actions may be observed as behavior. Reasons vary depending on the circumstances. Therefore, motivation is circumstantial and so is behavior. Motivation emerges when a certain circumstance is recognized. The circumstance presents the reason to act. It is therefore the process of recognizing the circumstance that causes actions. Processes are realized by systems. Behavior therefore requires a system to recognize the circumstance to give reasons to act.

Let us call such a system a *value system*. In this article, a value system plays a central role as the mechanism of detecting inherently meaningful signals from the environment that cause reactions or purposive acts. The key is to identify what an inherently meaningful signal is to the machine under what circumstances. Without a value system, there is no intrinsically motivated act. Engineers should be able to program and embed an initial value system in the machine at the time of its construction to trigger reactions to signals; however, once built, the innate value system shall evolve to a system that circumstantially and dynamically drives purposive acts. The innate value system thus defines the perspective and agency of the machine. This article explains why and how.

Finally, the motive of this research. The purpose of studying developmental autonomous behavior is to understand what a machine is from a different perspective, and to foster safe and healthy relationships between humans and machines. It goes back to the origin of a machine; why we build a machine in the first place. Human-machine collaboration is the ultimate reason for humans to build machines. We benefit from building machines by combining each other's strengths and filling in for weaknesses [101], [102]. However, empowering machine's capabilities without being fully aware of its consequences and repercussions impacts human lives and environment negatively. We are learning the lessons the hard way today.

As anyone with a pet knows, we appreciate how a rich and meaningful relationship can develop between humans and animals. This rewarding experience is a result of the pets' ability to exhibit responsive, adaptive, and developmental behavior, but such a relationship is built on our awareness of the inherent risk. Ultimately, the goal of this research is to inspire people to look at machines from different perspectives for responsible and sustainable human-machine collaborations.

D. DEFINITIONS OF TERMS

The following is a list of terms that appear frequently in this article. They are logically ordered to describe their meanings to avoid misunderstandings or unnecessary debate on terminology.

Behavior - observed phenomenon, exhibited by a system.

System - a set of things working together as part of a process.

Process - a series of steps taken to proceed to the next end point.

Embodied - the condition in which a system is represented in a tangible, visible, or accessible form.

Organism - an embodied system that satisfies the properties of life as defined in biology.

Machine - an embodied non-organismic system that exhibits behavior.

Environment - a habitat where organisms and machines exist.

Stimulus - a thing or event that evokes functional reactions in an organism or machine.

Sensor - a system that detects stimuli.

Response - a reaction to a stimulus in an organism or machine.

Actuator - a system that causes behavior.

Action - an output of an actuator.

Motivation - reasons that cause actions. Reasons vary depending on the circumstances; therefore, motivation is circumstantial.

Intrinsic motivation - motivation originating from internal circumstances of an organism or machine.

Extrinsic motivation - motivation originating from external circumstances to an organism or machine.

Value system - a mechanism to detect meaningful signals from the environment. Observed events are internalized by the value system to become an internal circumstance.

Signal - a representation of a stimulus as received by a sensor.

Sign - a representation of properties of a signal.

Symbol - a representation of properties and relationships of signs.

Information - a general term to represent an interpretable property of a signal, sign, or symbol.

Data - a general term to represent signals, signs, or symbols as an input or output of a system. Information is carried by data to be processed by a system.

Memory - a system to store data. Also refers to the storage as well as what is stored in the storage.

Processor - a system that processes data.

Concept - information conceived in memory.

Adaptation - a process or action of organisms or machines that alter their behavior according to observed changes in their internal or external conditions. Adaptive behavior is an act that results from adaptation.

Prospection - a process or action of organisms or machines that alter their behavior according to expected changes in their internal or external conditions. Prospective behavior is an act that results from prospection.

Developmental system - a system that progressively acquires new behavior through a process of adaptation and prospection.

Ontogenesis - development of behavior in a developmental system from its birth to maturity.

Phylogenetics - an evolutionary view of organisms. Borrowed from phylogenetics, this word is used in this article to define conditions required for a machine to exhibit certain behavior.

Cognition - a general term that implies certain capabilities of organisms to process signals, signs, and symbols.

Autonomous system - a system that acts on its own.

Developmental autonomous machine - a machine that acts on its own and progressively acquires new behavior through a process of adaptation and prospection. Why and how it acts and changes its behavior depend on the purpose, environment, and configuration. Developmental autonomous machines can be defined, classified, compared, analyzed, and synthesized based on three aspects:

- *Anthropogenic aspect* (Purpose) defines an autonomous machine in terms of the purpose of its existence and behavior,
- *Ecological aspect* (Environment) defines an autonomous machine in terms of the environment in which the machine is used, and
- *Phylogenetic aspect* (Configuration) defines an autonomous machine in terms of the sensory inputs, actuator responses, and signal processing systems to support its existence and behavior.

II. DEVELOPMENTAL BEHAVIOR

A. INTRODUCTION

Behavior is an observed phenomenon, exhibited by a behaving system. Machines and organisms, including humans, are all behaving systems. We may not see what is happening in their internals, but we can observe what they exhibit as behavior. To analyze behavior in biology, Nikolaas Tinbergen proposed four perspectives: causation, evolution, survival value, and ontogeny [103]. Causation is about internal mechanisms and external triggers that cause behavior. Evolution is about the behavior's evolutionary progression. Survival value is about the behavior's purpose. Ontogeny is about the behavior's development. In other words, the causation perspective asks how the observed behavior is triggered and accomplished mechanistically. The evolution perspective asks how the behavior evolved historically. The survival value perspective asks how the behavior contributes to its survival. The ontogeny perspective asks how the behavior develops over the lifetime of the organism [104]. In essence, Tinbergen's four perspectives of behavior address the study of cause, effect, and their development. The cause and effect in biological behavior was also articulated by Ernst Mayr in terms of ultimate and proximate explanations [105]. In essence, ultimate explanations are concerned with why a behavior exists, and proximate explanations are concerned with how it works [83], [104], [106].

Behavior analysis often accompanies a general concept of cognition. It refers to the mental process of reasoning and knowledge [107]. Cognition implies certain capabilities in organisms to process sensation and mental formulation that drive behavior. Cognition is a theoretical construct, not a tangible entity that one can observe directly; thus, varieties of interpretations have been derived and debated in philosophy and psychology. For this reason, it is important to clarify how cognition shall be treated in behavior analysis for machines.

Lea and Osthaus [108] offer a biologically grounded perspective in analyzing cognition. In their study of canine cognition, they defined what a dog is from three perspectives: phylogenetic, ecological, and anthropogenic views. Phylogenetic view considers where the animal fits in relationship with the evolutionary tree of biological species. This view represents a constraint on the animal's cognition in terms of its nervous systems, sensory inputs, and motor responses. Ecological view considers where the animal fits in relationship with the resource-driven environment. This view represents a constraint on the animals' cognition in terms of the purpose to which cognition is put to use in natural habitat. Anthropogenic view considers where the animal fits in human history. This view represents a constraint on the animal's cognition in relationship with humans as every animal on this planet is directly or indirectly influenced by humans in terms of their chance of survival.

By defining what a dog is from these three perspectives, Lea and Osthaus selected comparable animals and analyzed across a range of cognitive domains, including associative learning, sensory cognition, physical cognition, spatial cognition, social cognition, self-consciousness, and mental time travel.

This analytical framework of cognition is directly applicable to defining what an autonomous machine is in terms of its operating environment (ecological view), system configurations (phylogenetic view), and the purpose, values, and relationship to human designers and users (anthropogenic view). The focus on behavioral causation emphasizes principled explanations to, not only the "how" question on mechanisms but also the "why" question on behavior development. These biological perspectives of cognition and behavior provide a useful guideline to study developmental behavior in machines. This ethology-oriented perspective offers a renewed opportunity to construct a theoretical framework of developmental autonomous behavior from the ground up. Based on this perspective and foundation, ontogenies of human and machine behavior are analyzed in the next sections.

B. ONTOGENY OF HUMAN BEHAVIOR

Every animal goes through growth stages from newborn to maturity, and the behavior changes accordingly. For example, a newborn human baby exhibits limited types of behavior by crying, feeding, and sleeping. The baby gradually acquires new behavior by responding to smell, touch, sound, and sight. We can observe the change in behavior as she moves her arms, legs, makes sound, and changes facial expressions in response to her sensing. By the time she stands and walks, her behavior changes quickly in response to the consequence of her actions. Her movement becomes faster and more fluid, and the ability to communicate with the external world improves rapidly. As time goes by, she gains experiences interacting with the environment, and her behavior becomes anticipatory and prospective. She explores and exploits the world around her, and life continues. From this ontogenetic perspective, the behavior development can be seen as a progressive process of new skill acquisition.

Jean Piaget [72] is one of the early pioneers who analyzed and documented the behavior development process from observations. Piaget identified four stages: sensorimotor stage (infancy), pre-operational stage (toddler and early childhood), concrete operational stage (elementary and early adolescence), and formal operational stage (adolescence and adulthood) [109]. More recent analyses by Mascolo and Fischer [110] show a contemporary model of development stages: reflexive, sensorimotor, representation, and abstraction.

According to these models, infants at the reflexive stage exhibit basic reflex actions in response to social stimulations. In the sensorimotor stage, children begin to exhibit sensorimotor actions for goal-directed acts. In the representation stage, children exhibit multiple complex action patterns as a single representation of sensorimotor experience. By connecting sensorimotor experience to a representation, children can make a sound or picture stand out for an object or meaning of a word. Single representations allow children to form images and ideas of objects and meanings, including a representation of self. In the abstraction stage, older children begin to represent "*generalized, intangible, and hypothetical aspects of events, people, things, and processes*" [110].

The growth stages described above can be explained as part of a system transformation process (Fig. 2). It begins at birth with a basic system of sensorimotor reflexes and motor skill learning. The reflex system develops into a sensorimotor system of a lower-order with single actions to a higher-order with multiple action patterns. The higher-order sensorimotor system develops into a single representation system. The single representation system develops into a system of abstraction. And finally, the system of abstraction develops into abstract principles, which are the highest

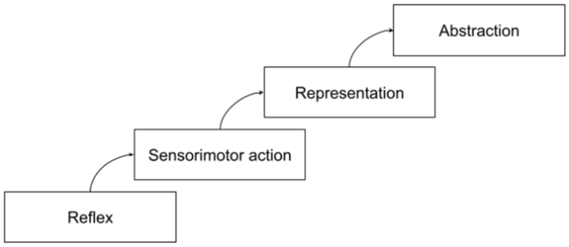

FIGURE 2. Human skill development stages by Mascolo and Fischer [110].

level of human skill development according to Mascolo and Fischer.

According to these models, the nature of information represented and used at each stage of the development process is speculated to change from sensory stimuli to interpreted representations to higher-order representations of abstract ideas and principles. Rasmussen conceptualized the nature of information representation in three key words: signals, signs, and symbols [111].

"*Signals are sensory data representing time-space variables from a dynamical spatial configuration in the environment, and they can be processed by the organism as continuous variables.*"

"*Signs indicate a state in the environment with reference to certain conventions for acts. Signs are related to certain features in the environment and the connected conditions for action. Signs cannot be processed directly, they serve to activate stored patterns of behavior.*"

"*Symbols represent other information, variables, relations, and properties and can be formally processed. Symbols are abstract constructs related to and defined by a formal structure of relations and processes - which by conventions can be related to features of the external world.*"

By using signals, signs, and symbols, Rasmussen illustrates three levels of the human performance model (Fig. 3). The skill-based behavior represents sensorimotor performance without conscious control. The rule-based behavior is controlled by a stored rule or procedure provided. The knowledge-based behavior takes place during unfamiliar situations when skill-based or rule-based behavior is not sufficient.

According to Rasmussen's signal-sign-symbol formulation, signals and symbols are treated as directly manipulatable entities while signs cannot be processed directly. This is because signs are interpretations of physically sensed signals. This brings us back to Helmholtz's sign theory [25] from the historical review of thought frameworks in Section I-B. Helmholtz separates sensation from perception, which is an interpreted version of the observed sensation. Rasmussen's treatment of signs is in alignment with Helmholtz' perspective. The separation of signals and signs conceptually justifies the need for different systems for behavior execution.

The process of extracting features in the environment is carried out by a system of feature formation (FF) as depicted in Fig. 3. The output of FF is a sign that triggers

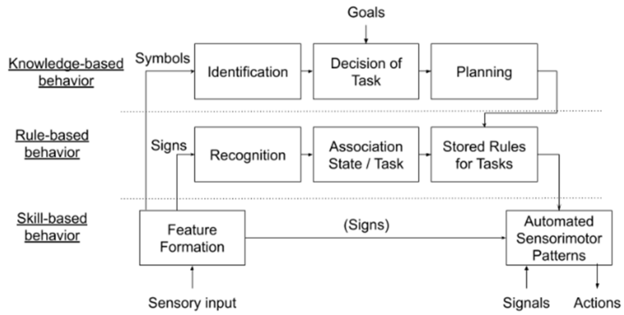

FIGURE 3. Three performance levels of skilled human operators by Rasmussen [111].

actions by the automated sensorimotor patterns (ASP). The process of extracting features from sensed signals by FF and triggering actions in ASP is observed as skill-based behavior, according to Rasmussen. The extracted feature can also trigger associated states for stored rules that specify predefined actions (tasks), which are executed by ASP. This process of recognizing signs for associated states that directly trigger action policy is observed as rule-based behavior.

If there are no associated states for the extracted signs, the system cannot trigger any action with the skill-based and rule-based setups alone. For example, suppose a worker is instructed to strictly follow an operation manual to perform a certain task. When he faces a situation that is not specified in the manual, he would not know what to do. This situation raises a need for more than skill-based and rule-based behavior. For humans, it typically involves a process of identifying the meaning of the situation as well as a desire for a change in the situation. The worker in the example can potentially identify what kind of changes are necessary to mitigate the situation. In this case, he would make his own decision to act outside of the predefined rules of actions. For him to do so, he would need to understand what the desired state might be under the specific circumstance.

Rasmussen uses the word "goal" to represent the desired state, and "symbol" to represent the properties and relations among signs. By identifying the properties and relations among signs, along with his awareness of the goal, the worker in the example can mitigate the situation without relying on the operation manual. This process is observed as knowledge-based behavior, according to Rasmussen.

The human skill development model is in alignment with the human behavior development models by Piaget and Mascolo-Fischer in principle. At the early stage of human skill development, sensory stimuli are signals that trigger reflex actions. A signal is a quantitative indicator of the time-space behavior of the environment, and has no meaning or significance as is, until it is interpreted, and a sign is extracted from it. A sign is a perceived information in the signal that represents an event or state of the environment. Signs are simply features in the signals from the environment, and by themselves do not provide much more information unless they are linked to other signs or actions. The relationships among signs and actions can be conceptualized as a higher-level representation of the signals, referred to as symbols. A symbol is a concept that represents relationships, properties, objects, and situations in the environment. At the later stage of human skill development, symbols are the primary representation of ideas and situations for reasoning and prediction.

C. ONTOGENY OF MACHINE BEHAVIOR

1) THE FRAMEWORK

Let us bring our attention to machines. If we were to build a developmental autonomous machine that learns from its experience and acquires new skills on its own, what growth stages would it go through, why, how? Before we proceed, let us impose a restriction on machines in terms of their hardware. Humans and animals physically change as they grow. It is conceivable that machines could alter their physical configuration on their own, intentionally or accidentally, but this complicates the matter of behavior change. For this reason, let us assume that the machine's hardware configuration remains the same and will not change in its lifetime. We will address the topic of morphology once we understand the basic principles of behavior development.

Even with the initial morphology restriction, we must consider a variety of system components and configurations for machines. In practice, a machine's system configuration is determined in accordance with its purpose and operating environment. Particularly sensitive to the determination is sensorimotor systems. Environmental properties of underwater, underground, atmosphere, outer space, and overland terrains all differ and contribute to special requirements for the machine to operate. Sensors, actuators, process units, chassis, and power supplies cannot be selected properly without knowing what kind of environment it is put to use, and why it is built in the first place. We must therefore accept such variations.

Despite the physical variations in machines, there is one component that is commonly required in all environments and purposes across the board: power supply. Energy is the ultimate reason that drives behavior. If an animal is hungry, it looks for food. If an obstacle prevents it from getting to the food source, it is seen as a threat. By the laws of thermodynamics, without a source of energy to supply the power needed to function, neither organism or machine can survive. It is no coincidence that energy has been and continues to be a major economic, political, and social issue in the human world.

Typically, it is the job performed by humans to maintain the power supply for machines. Eliminating the support of humans, machines are no longer autonomous in the sense that they cannot sustain their existence in the environment on its own. Deducing from this observation, energy plays a central role in machine behavior and autonomy. Let us now formulate a framework of study on the ontogeny of machine behavior based on this perspective.

Suppose there is a machine equipped with an arbitrary choice of sensing, acting, and processing components, and is placed to operate in an arbitrary environment. Suppose also that the purpose of the machine is to sustain its existence in the given environment on its own. The only thing the machine needs to do is to replenish its own energy so that it can operate continuously and indefinitely on its own. The environment should provide some resources if the machine can figure out how to exploit them. In other words, the machine's task is to exploit any resource it finds in the environment to operate continuously without completely depleting its energy. In short, this is a survival game.

There are two rules for this game. The first rule is that the machine is allowed to use whatever it has in its system configuration as well as whatever it finds in the environment. The second rule is that the machine is not given instruction about how to use its own system components except a few basic movements a priori. It is the machine's job to learn how to use its body, find useful resources in the environment, learn how to use them, and ultimately replenish its energy before it runs out. A surviving machine can claim itself to be a fully autonomous machine.

The analysis of human behavior development provides a useful framework to set up and analyze this game. It allows us to focus on the relationship between signal processing and skill acquisition. The signal-sign-symbol formulation conceptually justifies the need for specific systems to process raw sensor signals and bring them to higher-levels of representations to cause different types of necessary behavior. Before we can design such systems, important questions must be answered. First, what type of behavior is necessary to win this game? Second, how does a machine acquire such behavior? The first question is addressed in this chapter. The second question is addressed in Chapter III.

2) SELF-EXPLORATORY AND REFLEXIVE BEHAVIOR

Before a machine finds its energy source in the environment, it must learn to use its body. As the game's second rule states, a machine does not know how to use its body in the beginning. As seen in human skill development, the beginning of a growth path takes place in the sensorimotor system for basic motor skill learning. For a machine to survive and thrive in its environment, it must be able to move and control its own body precisely and purposefully. This skill must be acquired first if not given innately.

What we may observe at this early stage is a kind of behavior that appears random and uncontrolled. Meltzoff and Moore [113] used a term ''body babbling'' to describe the experiential process of human infants moving their limbs and facial organs in repetitive manners. What is learned from body babbling is a mapping between movements and body end states [114]. With proprioceptive and motion sensors in the body, the dynamic patterns of movements and what happens to the body as the end results can be monitored. The idea of body babbling is that such mapping must be known and remembered for purposeful actions. This mapping must be acquired from experience if not innately given.

Let us call this random and uncontrolled behavior self-exploratory behavior. Self-exploratory behavior is about motor skill learning. Because the machine cannot safely and effectively explore its outer environment yet, it explores its own body first to learn how to use it. It learns the relationship between body movement and internal sensation. Body movement is executed by its actuators, and the internal sensation is detected by internal sensors, such as proprioceptors, motion/balance sensors, and energy sensors. Self-exploratory behavior is a prerequisite to exploratory and exploitative behavior in the external environment that comes later.

As the machine exhibits self-exploratory behavior, its external sensors begin to detect signals from the external environment. These external signals can be interpreted in three basic ways: appetitive (positive), aversive (negative), and neutral. Appetitive signals are interpreted positively, aversive signals are interpreted negatively, and neutral signals cause no immediate interpretation at the time of reception. What is positive or negative depends on the machine's innate value system, a mechanism to detect intrinsically meaningful signals from the environment.

From an observational standpoint, the machine is expected to exhibit reactional behavior in response to appetitive and aversive signals. For example, detecting an appetitive signal may cause the machine to approach toward where the signal is coming from, like a hungry dog approaching toward a smell of food. Detecting an aversive signal may cause it to move away from where the signal was detected, like a hand twitching away from a burning stove. These are innate responses to specific features of sensor signals.

This type of behavior is called a reflex action or reflexive behavior. In human physiology, five major elements contribute to reflex actions: sensory receptors, sensory neurons, spinal cord, motor neurons, and muscles [31]. First at the point of sensing, for example a fingertip, sensory receptors trigger sensory neurons to carry a nerve impulse to the spinal cord. Neurons in the spinal cord pass the impulse to motor neurons. The motor neurons then carry the nerve impulse to the muscle and the muscle then contracts. As a result, your finger moves away from the sensing spot.

From these observations and analyses, it is reasonable to consider the self-exploratory and reflexive behavior as the initial stage of behavior development for developmental autonomous machines. Self-exploratory behavior is an early-stage motor skill learning by associating the movements with the internal effects by using internal sensors. Reflexive behavior is an innate, involuntary movement in response to an aversive or appetitive signals from the environment. It is primitive yet serves an important purpose for the machine's development and survival. The ability to map a relationship between signals reflects the machine's associative learning capability with memory. The innate patterns of aversive and appetitive responses reflect the machine's prototypical definition of value system, analogous to survival instinct in animals. Because these are innate behaviors, such abilities must be given by humans for a machine to play the game of survival.

3) OBSERVATIONAL ASSOCIATION AND EPISODIC BEHAVIOR

While positive and negative signals elicit predefined responses in reflexive behavior, neutral signals do not elicit any apparent response in the beginning, because there is no innate reflex act defined for neutral signals. It does not mean that neutral signals are useless. As the machine encounters new situations from self-exploratory movements, some of these neutral signals may turn out to be related in some ways to positive or negative signals. For example, the sound of a bell may be neutral at first, but if it consistently precedes the delivery of a food, the ringing bell sound may evoke positive interpretation, like a tell-tale sign of a good thing to come. The neutral signal may be perceived now as a meaningful signal and may elicit a conditioned reflex act. This type of behavior has been recognized as classical conditioning in experimental psychology [69].

Suppose there is a system that recognizes a certain feature in an observed signal that may be related to a positive or negative signal. Such a relation can be defined in different ways, such as the proximity of occurrence in time, space, or something else. The strength of relation can also be defined in different ways, such as the frequency of occurrence. Suppose that there is a system that stores related features as linked elements. With these systems, a process can be established to recognize, remember, and recall certain features from observed signals. Let us call this associative learning process an observational association. It's observational because the association is based on observed signals.

Suppose that some of the stored features act as conditions for predefined actions such as reflex acts. When a feature is detected in the observed signal, its linked elements are recalled, and trigger associated actions. Let us call this behavior episodic behavior. Episodic behavior is an acquired, involuntary behavior based on observed events from experience. The difference between innate and acquired behavior is that innate behavior is prebuilt at birth while acquired behavior is learned from experience. The etymology of the word episodic comes from Greek *epi* ("in addition") and *eisodos* ("a coming in, entrance") [115]. This implies that episodic behavior arises from a process of adding new experience into existing ones.

Episodic behavior is primitive yet serves an important purpose for the machine's survival and development. It reflects the machine's ability to exploit new signals from the environment by associating with the intrinsically meaningful signs. By identifying and associating novel signals with meaningful signs, the machine can anticipate in essence a meaningful event in advance. This capability not only improves the chance of survival, but also establishes the growth path for further development.

4) INTERVENTIONAL ASSOCIATION AND PROCEDURAL BEHAVIOR

The process of motor skill learning in self-exploratory behavior eventually leads to more precise motion control as the mapping solidifies between the body movements and the internal sensation. Conceptually speaking, the direction of mapping can be two ways: one from the movements to the sensation, and the other from the sensation to the movements. In neuroscience, the former is called a forward model, the latter an inverse model [116], [117], [118]. The forward model tells what happens to the body when moved. The inverse model tells what move causes the sensation (effect). When combined, the two models establish the causal relationship between the movements and the effects. In a sense, these models can function as a predictor of the consequences of a movement as well as an effect.

By associating its movement with observed changes in the environment as its consequences, the machine can map the causal relationships between them. The resulting mapping can then be used to voluntarily choose an action to obtain a desired outcome. In other words, mapping between actions and consequences enables purposeful behavior. In contrast to the passive and involuntary nature of reflexive and episodic behavior, the machine's behavior can become active and voluntary by associating the external and internal sensing signals with motor command signals.

For this to happen, however, there must be a system that defines what a desired outcome is under the circumstance. Without such a system, there is no basis or reason for machines to choose an action. Let us call this hypothetical system a circumstantial value system. In contrast to the innate value system in reflexive behavior, which is a fixed, binary mechanism of judging observed signals, the circumstantial value system is a dynamic, non-binary mechanism of judging own movements and observed signals. It computes expected values in sensor signals with respect to motor command signals, and by using the values, it determines an action that leads to a certain outcome under the specific circumstance. Associating the own movements with environmental cues allows the machine to purposefully select movements. Let us call this associative learning interventional association.

The basic tenet of this type of behavior can be described from the early works by Edward Thorndike and B. F. Skinner. According to Thorndyke, "*To explain fully why any human being thinks and feels and acts as he does, it is necessary to know what circumstances will give him the feelings of satisfaction and of discomfort. Having learned that connections productive of satisfaction are selected for survival and that connections productive of discomfort are eliminated, the final step is to learn what sort of result is satisfying*" [68]. According to Thorndyke, actions that produce pleasure are likely to be reproduced, while actions that produce pain are less likely to be reproduced. This so-called "Law of Effect" implies a process of "trial and error" or "search and select" to find the action that brings pleasure and avoids pain. This process is often referred to as "instrumental learning" and has been extensively studied and applied in reinforcement learning [55].

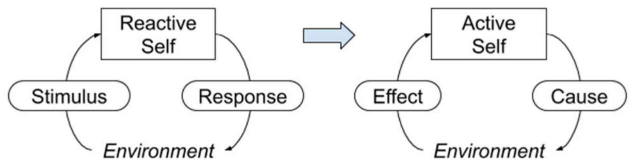

FIGURE 4. Transformation from reactive to active self.

Skinner introduced the term "operant conditioning" to describe the type of learned behavior that responds to non-eliciting stimuli [71]. He argued that not all behavior fit into the simple stimulus-response formula. He called the type of behavior that responds to specific eliciting stimuli respondent, and all non-respondent behaviors operant. In his description, respondent behaviors are associated with prior events, while operant behaviors are associated with posterior events. A typical example to describe instrumental learning and operant conditioning is a reward-punishment experiment. For example, let's say there are blue and red buttons. If you press the blue button, you get rewarded with food. If you press the red button, you get punished by an electric shock. Which one would you press? After some attempts of pressing both buttons, lab rats eventually learn to press the blue button more often than the red button.

From these observations and analyses, it is reasonable to consider the type of behavior resulting from interventional association as part of the machine's developmental growth path. Let us call this behavior procedural behavior. Procedural behavior is an acquired, voluntary behavior based on associated actions and consequences from experience. The etymology of the word procedural comes from Latin *pro-* ("forward") and *cedo* ("go, move") [119]. This implies that procedural behavior arises from an effort to proceed for a purpose.

The implication of the machine being able to actively choose its action is significant. The machine was a passive system of stimulus-response when exhibiting reflexive and episodic behaviors (Fig. 4, left). Associating actions to events promotes a new perspective of the response as a causal action, and the observed stimulus as a consequential effect (Fig. 4, right). As a result, what started as a random exploratory activity in the early stages of reflexive and episodic behaviors progressed to an opportunity-seeking trial-and-error activity. The transformation from the passive reactive self to an active self enables active learning and purposive actions.

5) HABITUAL AND AUTONOMIC BEHAVIOR

As an example of trial-and-error behavior, let us consider the case of riding a bicycle in humans. When we first learn how to ride a bicycle, we pay a lot of attention to every little movement and try to react to keep a balance. In the beginning, too much or too frequent reactions may cause the bicycle to lose its balance. The process is a repetitive series of sensing motion feedback signals and exerting motion control signals. As we get the hang of motion control, we gradually pay less and less attention to the motion feedback signals. The operation eventually becomes smooth and fluid, almost automatic. As the body experiences the sequence of actions repeated many times, the behavior becomes habitual.

Conceptually speaking, the basic idea is that the sequence of feedback-based movements becomes associated as one feedforward movement. Feedforward is faster than feedback. When a sequence of feedback-based movements is performed, a number of input-output feedback cycles must be executed. If the sequence is repeated often, the whole sequence of movements can be executed more efficiently by eliminating the process of checking a feedback signal each time before responding. The sequence can become a chain reaction of responses without signal checking. This feedforward action is faster and more efficient than a series of feedback actions.

From these observations and analyses, it is reasonable to consider this type of behavior as part of the machine's developmental growth path. Let us call this autonomic behavior. Autonomic behavior is an acquired, voluntary behavior based on habit and self-governance of motor skill control. The word autonomic is used to avoid confusion with similar words, automatic and autonomous, which are commonly used to represent flavors of self-governance such as automatic controllers and autonomous vehicles. Autonomic behavior emerges as a motor skill learning behavior because of repeated sequence of actions.

6) COUNTERFACTUAL AND CONCEPTUAL BEHAVIOR

Autonomic and reflexive behavior in machines can be considered similar to skill-based behavior in the human skill development model by Rasmussen, in the sense that the information used to cause actions is based on signals. Episodic and procedural behaviors in machines relate to rule-based behavior in the sense that the information used to cause actions is based on signs (features of signals). Signals and signs are theoretical constructs and may not reflect the underlying physical mechanisms directly, but the idea is to distinguish the levels of information used to cause different actions.

Exhibiting reflexive behavior is evidence of the machine responding to naturally meaningful signals. Episodic behavior is evidence of associative learning. Procedural behavior is evidence of the ability to choose actions by learning the causal relationship between actions and consequences. Autonomic behavior is evidence of motor skill learning. Combined with these types of behavior, it is not unreasonable to expect a machine to survive if there is sufficiently useful information in the signals from the environment. If this is the case, is there any other type of behavior necessary?

Consider a situation where a significant change occurs in such a way that previously learned behavior does not lead to expected consequences. For example, suppose that a mobile robot is on its way to a battery recharge station and suddenly loses one of its wheels. When the wheel stops turning, the actions that worked before no longer work. What type of behavior would the robot exhibit? If the machine is limited to the behavior types based on signals and signs, it would presumably relearn its motor skills and repeat exploring and exploiting the environment again.

From a human's perspective, this is not necessarily an unusual situation though not desirable. When a car breaks down, for example, we would probably think of what to do, like calling someone, checking the car to see if you can fix it, walking to a nearby station, or waiting for someone to help you. This is a process of imagining what-if scenarios with hypothetical actions and judging the possible consequences based on a value system. The process of "what if" scenario exploration takes place in the memory, not in the physical environment. We then proceed by prioritizing an action.

The process of "what if" scenario exploration is called counterfactuals. The dictionary definition of counterfactual is "contrary to fact" (merriam-webster.com). In philosophy, counterfactuals are modal discourse that concerns alternative ways things can be, as "what is not, but could or would have been" [120]. Judea Pearl defines counterfactuals as probabilistic answers to "what if" questions [121]. Counterfactuals are typically used in a sentence as a contrary-to-fact antecedent. For example, a sentence of the form "if X did not occur, then Y would have been different" declares a hypothetical situation with a contrary-to-fact condition.

Counterfactual reasoning does not involve sensorimotor systems. It takes place in a processing component without relying on sensors or actuators. Any behavior resulting from counterfactual reasoning can be considered similar to knowledge-based behavior in the human skill development model, in the sense that the information used to cause actions is based on symbols [111]. Symbols represent properties and relations of features of the external world. Symbols are a theoretical construct to represent a level of information used to cause a different type of actions. Let us call it conceptual behavior.

Conceptual behavior is an acquired, voluntary behavior that is performed entirely on internal memory. The word concept comes from Latin *concipio* from *con-* ("together") and *capio* ("to capture, to grasp, to derive") [122], [123]. In English, the word concept is defined as an abstract idea, something conceived in the mind [124]. It is implied that conceptual behavior is derived from abstract ideas conceived in the mind.

Conceptual behavior can be considered as prospective behavior in the sense that it hypothesizes scenarios based on expected changes in the imaginary world, rather than observed changes in the real world. The prospective nature of conceptual behavior enables machines to behave in uncertainty without a priori instruction. This is useful in finding the best course of actions in case of unforeseen events and situations. With the ability to imagine and hypothesize scenarios, it is possible for conceptual behavior to produce novel actions that have not been exhibited before.

7) MULTIPLE MACHINES AND SOCIAL BEHAVIOR

Suppose there is another developmental autonomous machine in the same environment. Suppose also that both machines are equipped with a microphone and speaker. Because all external elements in the environment can be either a potential resource or threat, machines may approach, avoid, or ignore each other, depending on their signal interpretations. In the meantime, each machine has been minding its own business, performing its motivated acts.

Now let us suppose that machine A detects a light coming from a battery recharge station (i.e., food source for machine A) and begins to move toward it. While moving, machine A emits a sound signal that represents its internal concept for "food". Machine B detects the sound signal but does not initially know what it means, so interprets it as neutral. After a while, it so happens that, if machine B approaches toward where the sound comes from, it often sees the light from the battery recharge station (also the food source for machine B). By observational association, the previously neutral signal of the sound becomes associated with the appetitive signal of the light; machine B associates the sound signal as a sign for food; machine B is now conditioned to approach machine A.

This is a simple example of how communication may emerge between two machines. Because of the interactive nature between the two machines, such behavior may appear "social" in the eyes of a third person observing them. The word "social" derives from Latin socius, meaning companion or ally (Merriam-Webster), which is originally derived from a Proto-Indo-European root word *sekor* or *sekw*, meaning to follow or go after [125], [126]. It can be said therefore that the simple behavior of machine B going after machine A is a prototypical social behavior.

Consider an example in human society. In baseball, a player sees his coach removing his cap. Mechanically speaking, the visual system of the player observed a video signal of his coach. The video signal contains certain characteristics in pixel changes which were associated as a sign to steal a base in the coach's mind. It so happens that the signal was already associated in the player's mind as a sign to steal a base. This mutual agreement of signals and signs allow the coach and player to communicate by visual signals. The opposing team's pitcher receives the same video signals but does not have a proper sign associated with the signal, so the pitcher does not understand what the signal means. The coach and player associate themselves as members of a group. They do so by establishing a mutual agreement of signals and signs to exchange meaningful signals to each other.

Social behavior is an acquired, voluntary behavior that interacts with other machines or organisms. When a group of machines interact with each other in a way new references are established through mutual understandings, a common language is born. While each machine acts for its own merit, some of them may share a similar motivation, such as finding a food source. If the group of machines agree with a common goal based on their shared objective, a new derived motivation emerges for each member of the group. This shared motivation drives new collective social behavior in each member of the group. The emergence of common language and shared motivation catalyzes the emergence of social behavior. Fig. 5 illustrates the social behavior emergence from signal exchange among three machines.

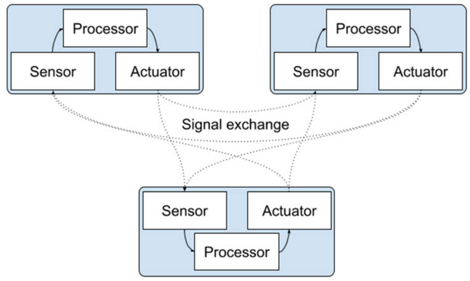

FIGURE 5. Social behavior by exchanging and interpreting signals.

8) ANTICIPATORY BEHAVIOR

There is a type of behavior that arises when anticipating, rather than observing, certain changes in the environment. To some extent, anticipating what would happen in the future implies prediction. Predicting the future is not trivial, but people and animals do anticipate future events and act accordingly. For example, when the sky gets darker and becomes blustery, we anticipate the rain coming and we prepare for it. When a child is young, we anticipate the growth and may opt to buy larger size shoes rather than tight-fit shoes. We can anticipate because we know from our experience that some events repeat with certainty, thus making them predictable, and that some future events have tell-tale signs. We also know that some future events are completely unknown or unpredictable, like when the next earthquake hits. However, we still try to anticipate and prepare for such events. Because of apparent complexity, anticipatory behavior is often considered a different kind of behavior from stimulus-response behaviors; however, such a view needs further analysis.

Four words commonly appear in scientific literature in relation to behaviors and future events: anticipation, prediction, preparation, and prospection. They imply similar but not the same meanings, yet clear distinction is often missing. The etymology of the word anticipate comes from Latin *anticipatus*, perfect passive participle of *anticipare*, from *ante* ("before") + *capere* ("to take") [127]. Therefore, the word anticipate implies an act of acting before something happens. The etymology of the word predict comes from Latin *praedīcere*, perfect passive participle of *praedictus*, from *prae* ("before") + *dīcere* ("to say") [128]. Therefore, the word predict implies an act of calling out future events beforehand. The etymology of the word prepare comes from Latin *praeparāre*, from *prae-* ("before") + *parāre* ("to make ready") [129]. Therefore, the word prepare implies an act of making ready in advance. The etymology of the word prospect comes from Latin *prospectus*, past participle of *prospicere*, from *pro* ("before, forward") + *specere, spicere* ("to look, to see") [130]. Therefore, the word prospect implies an act of looking forward.

Based on these origins and meanings, we can say that a behavior is predictive when a machine calls out future events or actions. A behavior is anticipatory when a machine takes an action before something happens. A behavior is preparatory when a machine acts to make ready for upcoming events. Therefore, prediction, anticipation, and preparation are related in that a machine predicts a future, anticipates a change in the future, and prepares for the anticipated change. On the other hand, a prospective behavior is slightly different. A behavior is prospective when a machine acts by looking into the future. It is not necessarily predicting a future, but rather it is expecting, hoping, or just imagining for a certain event or situation to occur. This active nature of prospection makes it different in behavioral characteristics in comparison to anticipation based on prediction. For example, when you are walking on a crosswalk and see a speeding car approaching, you walk faster to avoid a potential accident. This reactive behavior is anticipatory, but not necessarily prospective. When you hope to improve your football team's performance, you recruit a strong player to join the team. This active behavior is prospective, but not necessarily anticipatory. In other words, anticipation refers to the influence of predictions while prospection refers to the influence of expectation or imagination.

This etymological perspective of anticipation is in alignment with how an anticipatory system is defined in scientific literature. For example, Rosen defines an anticipatory system as "*a system containing a predictive model of itself and/or of its environment, which allows it to change state at an instant in accord with the model's predictions pertaining to a later instant*" [131]. Pelluzo uses Rosen's definition to argue that the sensorimotor state is insufficient to determine behavior and an anticipation of the future is needed in an anticipatory system [132]. Butz et al. [133] raised an important question of whether a predictive model is necessary for all anticipatory behaviors. From a broader perspective of anticipatory systems, they identified four different types of anticipatory mechanisms: implicit, sensorial, state, and payoff anticipation.

According to Butz, Sigaud, and Gerard, an implicitly anticipatory mechanism structurally arises without prediction about the future. Sensory inputs are directly mapped onto an action decision to generate anticipatory behavior. A sensorial anticipatory mechanism uses an implicit predictive mechanism to sense the future stimuli and states, which influence sensorial preprocessing, which in turn influence behavior. In other words, sensorial anticipatory behavior is a type of preparatory behavior that arises in response to sensory prediction of imminent changes in the environment. A state anticipatory mechanism uses an explicit predictive model of the environment to predict the future states, which directly influence behavior. A payoff anticipatory mechanism does not use a predictive model of the environment, but instead uses an expected payoff to choose an action [133]. Table 1 summarizes the classifications of anticipatory mechanisms.

TABLE 1. Anticipatory mechanisms and behavior.

Anticipatory mechanism	Predictive model	Ontogenetic behavior type
Implicit	*No*	*Reflexive / Episodic*
Sensorial	*No**	*Episodic*
Payoff	*No*	*Procedural*
State	*Yes*	*Conceptual*

** Prediction is done by an implicit sensory mechanism, not by an explicit predictive model.*

The classification mechanism helps decompose anticipatory behavior into four ontogenetic behavior types for developmental autonomous machines. The implicit anticipatory mechanism drives reflexive and episodic behaviors because of its direct stimulus-response process. The sensorial anticipatory mechanism drives episodic behavior also because of its stimulus-response process, in particular sensory preconditioning. The payoff anticipatory mechanism drives procedural behavior as exemplified by reinforcement learning. The state anticipatory mechanism drives conceptual behavior because of its explicit prediction model that resides in memory.

For example, as in classical conditioning, an experiential association between a neutral stimulus NS and an unconditioned stimulus US triggers a conditioned response CR. This behavior can be considered an implicit anticipatory behavior as CR arises in anticipation of US. It is also possible to regard the innate reflexive behavior as implicit anticipatory because the ultimate explanation of innate hard-wired reflex actions is broadly speaking anticipatory. Butz, Sigaud, and Gerard wrote, "*broad understanding of the term anticipation basically classifies any form of life in this world as implicitly anticipatory*" [133].

The sensorial anticipatory mechanism drives episodic behavior, and the process can be explained by sensory preconditioning (SPC) [74]. SPC is a three-stage conditioning of two neutral stimuli (NS1 and NS2) and a conditioned response (CR). At stage 1, NS1 and NS2 are presented serially or simultaneously to establish an association between them. At stage 2, NS1 is conditioned to elicit CR as in classical conditioning. At stage 3, when NS2 is presented at a later time, CR is triggered even though NS2 was not conditioned directly to elicit CR. Based on the SPC mechanism, the phenomena of implicit and sensorial anticipatory mechanisms can be explained by a stimulus-response mechanism of neutral association and its transfer property.

In SPC, CR arises in direct response to sensing NS2, which is preconditioned to CS (formerly NS1) in the sensory system, which triggers CR in anticipation of US. According to the recent research by Cuevas and Giles [134], the SPC mechanism was observed in human infants' anticipatory behavior and latent learning capabilities.

The implicit and sensorial anticipatory mechanisms are also described by Hull [73] as antedating reactions based on the principles of stimulus-response mechanism. According to Hull, the antedating reactions are learned reactions "*that appear in advance of the point in the original sequence at which they occurred during the conditioning process*" [73]. The theory is built on six behavior principles: stimulus trace, positive association, negative association, forgetting, mark of reinforcing state, and internal stimulation.

From these analyses, what has been simply called anticipatory behavior is a complex class of behavior types from reflexive to episodic, procedural, and conceptual behavior, each demanding a different ultimate-proximate explanation. In this article, adaptation is defined as a process or action of organisms or machines that alter their behavior according to observed changes in their internal or external conditions. Adaptive behavior is an act that results from adaptation. Therefore, anticipatory behaviors that arise from stimulus-response mechanisms, such as implicit and sensorial, are considered adaptive behavior. Prospection is defined as a process or action of organisms or machines that alter their behavior according to expected changes in their internal or external conditions. Prospective behavior is an act that results from prospection. Therefore, anticipatory behaviors that arise from an internal predictive model in memory are considered prospective. The difference between observation and expectation separates the real environment where sensors and actuators operate, from an imaginary environment where memory dominates its operation. While creative, imaginative actions may result in prospection, ill-intended use or malfunction in memory could result in undesirable behavior.

9) LATENT LEARNING BEHAVIOR

Another type of behavior that needs some consideration is latent learning. Latent learning is a type of learning that is not apparent at the time of learning, but the learned behavior appears later in reward situations. For example, suppose that you are reading a book of botany and see a picture of poison oak. A few days later as you are walking, you spot a plant that looks like a poison oak. You act by moving away from the plant. From an observer's point of view, it is not apparent why and how you took such an action. Brodgett [70] first introduced the concept of latent learning to describe animals' abilities to acquire new behavior in non-reward conditions and use them later when reward situations appear. Some consider latent learning as a different kind of learning from Pavlov's classical conditioning or instrumental and operant conditioning described by Skinner and Thorndike. Some consider latent learning as a type of anticipatory behavior.

To begin with, why does latent learning behavior exist and how does it work? Tolman [76] approached the ultimate-proximate explanations by applying a concept of vicarious trial and error (VTE), introduced by Muenzinger [75]. VTE is a behavior observed in rats in a maze that appears as if rats are thinking about what to do. According to Tolman, animals exhibit latent learning because their brains remember prior experiences and organize them as a map which allows the animals to explore and exploit the environment [76]. Tolman proposed that animals perform latent learning by building a set of hypotheses, called cognitive maps, at decision points by actively searching, comparing, and remembering stimuli by doing VTE. According to Tolman, latent learning is a different type of learning behavior from classical and operant conditioning.

Hull took a different approach and argued that the principles of stimulus-response mechanism could explain all behaviors, including anticipatory and habitual behaviors [73]. According to Hull, latent learning is the same type of learning behavior as classical and operant conditioning. By using his six behavior principles, Hull proved that "*the Pavlovian conditioned reaction and the Thorndikian associative reaction are special cases of the operation of the same principles of learning*" [73].

Tolman rejected this view of stimulus-response mechanisms and criticized Hull's theory and Thorndike's as oversimplification of the laws of conditioning. Tolman wrote "*Hull, like Thorndike, passes from O′s and B′s to S′s and R′s with no clear statement of his justification for doing so . . . we must be told why and how the actual gross O′s can be reduced to simple S's, and the actual gross means-end B′s to simple R's.*" [76]. What he means by O, B, S, and R are object, behavior, stimulus, and response, respectively. Tolman's view gained popularity over the years, while Hull lost his.

Tolman constructed his theory by observing the behaviors of rats in various maze configurations. Because of the spatially oriented nature of the experiments, as well as the use of the word map, the concept of cognitive maps was literally interpreted by other researchers as a spatial representation of the physical world. However, cognitive maps, in their original form, are meant as a metaphor for information storage in memory, not necessarily exclusive to geometric maps of the physical world. Since O'Keefe and Nadal [135] published an extensive treatment of hippocampus as the neural substrate for cognitive maps, the hippocampus is now widely regarded as the primary brain region to handle spatial navigation for animals. However, many phenomena of impaired non-spatial behavior are also related to the hippocampus, and recent research shows that the hippocampus functions more than spatial navigation [136].

Based on the analysis of anticipatory behavior and human skill development, it is reasonable to consider latent learning as a type of behavior that can be observed at multiple stages in organisms' ontogenetic development. It could occur as episodic behavior in a form of sensory preconditioning, as described by Brogden. At other times, latent learning

TABLE 2. Latent learning mechanisms and behavior.

Theorist	Mechanism	Ontogenetic behavior type
Brogden	*Sensory precondition*	*Episodic*
Hull	*Sensorimotor principles*	*Episodic / Procedural*
Tolman	*Cognitive maps*	*Conceptual*

could occur as procedural behavior in a form of operant conditioning, or as autonomic behavior in a form of habitual behavior as described by Hull. And finally, it could occur as conceptual behavior in the form of hypotheses, as described by Tolman. Table 2 summarizes the categorical view of latent learning mechanisms and corresponding behaviors.

D. SUMMARY OF THE CHAPTER

This chapter sets up the framework to study developmental autonomous behavior for machines. The term "developmental autonomous behavior" categorically encapsulates the general ability of a machine to acquire new skills and behavior from its birth to maturity on its own without human intervention. The study is inspired and influenced by an ethological view on causations of behavior in biology. Considering the broad and diverse nature of machines in general, the framework is formulated by focusing on energy as the central driver of behavior. The internal energy level of a machine influences its motivational state. In this framework, machine autonomy is defined in terms of an ability to resupply its own energy without human intervention. It is assumed that a machine is equipped with an arbitrary choice of sensing, acting, and processing components, and is placed to operate in an arbitrary environment. Key findings in this chapter are listed below.

1) The framework is formulated as a game of survival with two rules. Rule #1: a machine can use whatever resource it finds in the environment. Rule #2: a machine is not given any instruction on how to use its body except a few basic movements a priori. End game: a machine wins if it can sustain its operation indefinitely without depleting its energy. The survivor can claim to be a fully autonomous machine.
2) Because a machine must learn to use its body to find a resource in the environment, it must be equipped with an ability to learn basic motor skills. Behavior observed from this learning activity is called self-exploratory behavior.
3) Because a machine needs a system to recognize meaningful circumstances that give reasons to act, it must be equipped with an innate value system that defines the most basic features, positive and negative, of the environment. Behavior observed from the innate action driven by the innate value system is called reflexive behavior.
4) Because a machine must be able to recognize novel features of the environment, it must be equipped with an ability to associate observed signals with meaningful features. Behavior observed from the acquired involuntary action driven by observational association is called episodic behavior.
5) Because a machine must be able to recognize a circumstance that gives a reason to choose an action, it must be equipped with a circumstantial value system that determines a desired outcome and action under the circumstance by associating actions and consequences. Behavior observed from the acquired voluntary action driven by the circumstantial value system and interventional association is called procedural behavior.
6) Because a machine must move precisely and efficiently to explore and exploit the environment, it must be equipped with an ability to learn motor skills in such a way that a series of frequently repeated actions becomes a fast habitual motion. Behavior observed from the automatic habitual action is called autonomic behavior.
7) Because unseen, unforeseen events occur at any moment in the environment, a machine must be equipped with an ability to perform counterfactual reasoning by processing high-level representations of the prior experience in memory to derive novel actions. Behavior observed from the acquired voluntary action driven by such a process is called conceptual behavior.
8) Because there might be more than one machine in the environment which can be a resource or threat for survival, a machine must be equipped with an ability to exchange signals to communicate with the others. Behavior observed from the action driven by communication is called social behavior.
9) Various forms of anticipatory behavior can be observed and classified in developmental behaviors. Implicitly anticipatory behavior can be observed in reflexive and episodic behavior. Sensorial anticipatory behavior can be observed in episodic behavior. Payoff anticipatory behavior can be observed in procedural behavior. State anticipatory behavior can be observed in conceptual behavior.
10) Various forms of latent learning behavior can be observed and classified in developmental behaviors. Latent learning behavior described by sensory preconditioning can be observed in episodic behavior. Latent learning behavior described by the sensorimotor principles can be observed in episodic and procedural behavior. Latent learning behavior described by cognitive maps can be observed in conceptual behavior.

The general idea of behavior observation and the developmental behavior categories are depicted in Fig. 6.

III. BEHAVIOR DEVELOPMENT

A. INTRODUCTION

The previous chapter formulated a framework of study on developmental autonomous behavior for machines. It was set

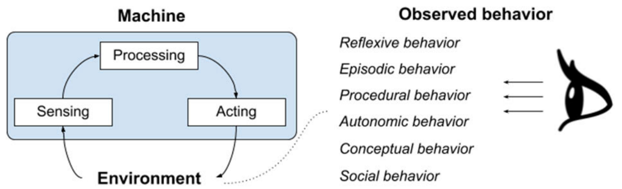

FIGURE 6. Six categorical behaviors for developmental autonomous machines.

up as a game of survival that challenges a machine to prove its autonomy in the sense of operation continuity. Within the framework, categorically different behaviors were identified as necessary for machines to play and win the game. Building on the set up, this chapter addresses the question of why and how machines develop those behaviors. We analyze the operating principles of developmental behaviors and the circumstances in which these behaviors emerge within the framework of a survival game.

Over the years, we have been accustomed to expecting machines to behave predictably within the bounds of their designs and specifications. With the emergence of autonomous machines, we begin to face situations where machine behavior is not exactly predictable or understandable. Unlike classic fixed-behavior machines, the behaviors of autonomous machines are highly circumstantial.

Human behavior is also circumstantial, but safe and meaningful relationships can be built on our awareness of each other's behavior and behavior development. For example, in the eyes of parents and caregivers, the behavior of their children may not be predictable but recognizable and understandable. This is possible because of the knowledge of the children's behavior development, providing a perspective and context that helps them recognize the circumstances that may have caused the observed behavior.

Knowing human physiology is not sufficient to understand human behavior. For the same reason, knowing machine specification is not enough to understand machine behavior, especially if they are built to adapt and learn from the environment. Studying behavior development provides vital insights to deal with future generations of autonomous machines.

1) ASSUMPTIONS

A machine needs essentially three building blocks to exhibit autonomous behavior: sensing, acting, and processing components. A sensing component detects signals from the environment, an artificial equivalent of sensory receptors and neurons. An acting component executes an action that influences the environment, an artificial equivalent of motor neurons and muscles. A processing component converts sensory signals to motor control signals, an artificial equivalent of the spinal cord, brain stem, cerebellum, and cerebrum.

In continuous active machinery in Turing's sense such as robots, sensing is typically achieved by electro-mechanical devices such as photoresistors and ultrasonic sensors. These sensors operate by converting physical properties of environmental elements to electrical signals. They are not typically designed to perform perception in Helmholtz' sense, so conversion from signals to signs in Rasmussen's sense must be handled by the processing component to make use of the signals.

Acting is also performed by electrical-mechanical devices, such as motors, pumps, speakers, light emitters, and various signal transmitters. Since most contemporary devices are electronically controlled, actuators typically operate by converting electrical control signals to change properties of a physical medium. It is a job of the processing component to feed the control signals to the actuators.

Processing is typically performed by signal and data processing devices. Contemporary versions of these devices, such as microcontrollers and microprocessors, are built on integrated circuits with embedded peripheral functions such as memory, analog-digital and digital-analog conversion, and communication interfaces. These functions are often packaged as one body unit, known as system-on-a-chip or SoC. While sensors and actuators are designed and manufactured for the purposes of specific use and performance, processing components are typically built for general purposes and therefore operation instructions must be developed and installed. Such instructions are generally called software, while the physical components are referred to as hardware.

It seems trivial but is important to recognize that this software-hardware separation is mainly a byproduct of digital computer design, not a natural principle of information processing. As the software-hardware paradigm has become the mainstream practice since the age of digital computers, it seems to have created a mental bias toward viewing software as cognition and/or cognition as software. If a machine is built to simulate a cognitive process, it is done so by using systems working together to enable the process. Software is a system and so is hardware. The two systems work together to enable a desired process. Therefore, software is not the only component that enables the process of simulated cognition (if there is such a thing).

Machines that play the survival game are built by humans. It is assumed that each machine is equipped with an arbitrary set of sensing, acting, and processing components, and is placed in an arbitrary environment. Software is a design element, along with hardware elements of sensing, acting, power supply, and chassis.

B. ORIGIN AND MECHANISM OF INNATE BEHAVIOR

1) OVERVIEW

To play the game of survival, humans must decide what basic functions the machines should possess. Varieties of machine designs are possible, but because of the game rules, all machines must first learn to move their bodies and experience the environment before they can effectively learn to survive. For machines to develop the skills necessary to explore and exploit the resources in a given environment,

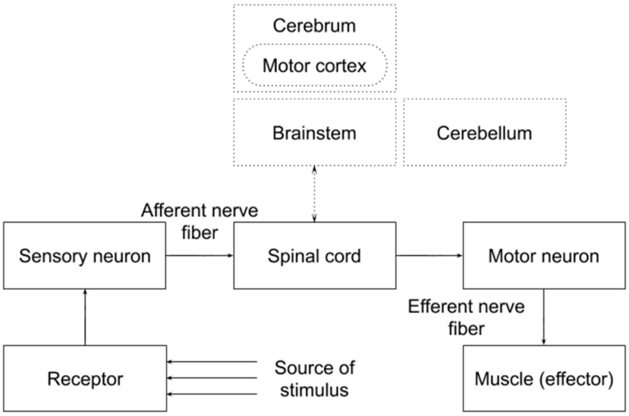

FIGURE 7. Neural mechanism of reflex acts in humans.

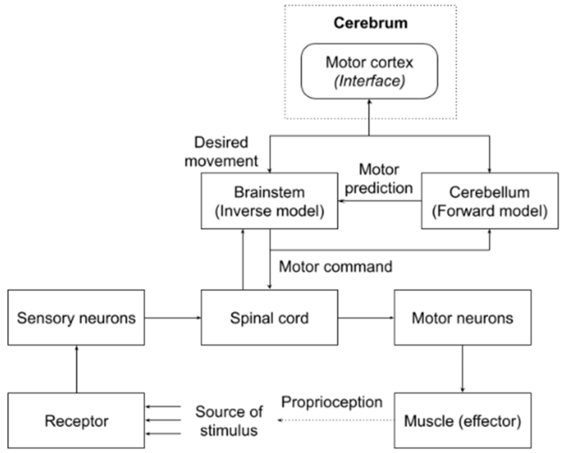

FIGURE 8. Neural mechanism of motor skill learning and reflex acts in humans.

what kind of basic, innate behavior should machine exhibit in the beginning?

From the observation of human behavior development, two types of innate behavior are considered necessary. First is reflexive behavior as an innate, involuntary movement in response to signals from the environment, which are pertinent to the machine's survival. Second is self-exploratory behavior as an innate, random movement of the machine's actuators, which promotes motor skill learning. These basic sensorimotor behaviors should minimally help machines get by the initial encounters in a given environment and serve as the foundation for subsequent behavior development.

2) NEURAL BASIS OF REFLEXIVE AND SELF-EXPLORATORY BEHAVIOR IN HUMANS

Since the concept of reflexive and self-exploratory behavior originates from the observation of human infants, let us first review what is known in neuroscience about such behaviors. Reflex acts in humans are coordinated, involuntary motor responses to a stimulus applied to sensory receptors [137]. Sherrington [138] was among the first to recognize simple reflexes as the basic units of movement. Fig. 7 is a conceptual diagram of neural mechanisms for reflex act. It involves sensory inputs, motor outputs, and spinal cord without the intervention of the Cerebrum (shown in solid lines, Fig. 7). However, some movements are mediated by supraspinal centers, i.e., brainstem nuclei, cerebellum, motor cortex (shown in dotted lines, Fig. 7), and this convergence of sensory signal processing at spinal cord and supraspinal centers allows reflexes to be smoothly integrated with centrally controlled motor commands [137].

In humans, sensory experience begins with sensory receptors, each responding to a specific type of energy and transducing electrical signals. Signals are quantitative representations of the time-space behavior of the environment detected by the sensing mechanism [111]. Signals have no meaning or significance until they are translated to signs. A sign is detected based on physical properties of the signals, such as intensity, size, temporal frequency, or detection threshold [139]. For behavior to arise, sensor signals must be translated to a sign that signifies a meaningful event, object, or state of the environment. Reflex acts therefore represent the meaningfulness of environmental cues in the most fundamental way.

Sensors are not limited to detecting signals from the external environment but also include the internal environment, i.e., conditions in the body. The internal signals are necessary to regulate our own body. For example, when we stand on one foot, our equilibrium tells our balance, while our proprioception tells our muscle movement so that we can balance even with our eyes closed. Without proprioception, we wouldn't know where our bodies are and how to move them. Without equilibrium, we wouldn't know how our bodies are moving or situated. Without an ability to sense our hunger, we wouldn't be able to manage our energy level to operate.

As discussed in the previous chapter, the body babbling behavior by human infants is the evidence of early motor skill learning. The neural basis of motor control is not completely understood and still an active area of research in neuroscience. Fig. 8 is a conceptual diagram constructed by adapting multiple hypotheses from Kim et. al. [118], Pearson and Gordon [139], and Lisberger [140]. It shows the conceptual neural mechanisms of reflex acts and motor skill learning. The brainstem and cerebellum are thought to be the primary brain regions involved in motor control. The cerebellum does not directly control the motor system, but instead functions as a predictor of the states of the motor system. This promotes fine motor control by the brainstem [118]. Proprioception and motor association are hypothesized to be taking place in the cerebellum.

3) OPERATION PRINCIPLES OF REFLEXIVE BEHAVIOR IN MACHINES

Reflexive behavior is innate and involuntary. The movement is innate because it is given to the machine at its birth. It is involuntary because the movement is generated by a stimulus-response mechanism, not a deliberate voluntary action. In other words, reflexive behavior is a programmed movement, designed by humans.

The reason why the machine exhibits reflexive behavior is because: (a) the machine experiences a negative situation that is harmful to the machine, or (b) the machine experiences a positive situation that is beneficial to the machine. In case of (a) where the observed signal is aversive, the machine exhibits a reflexive aversive response. In case of (b) where the observed signal is appetitive, the machine exhibits a reflexive appetitive response. Based on these ultimate-proximate explanations, the machine's reflexive behavior operation can be conceptualized as follows:

$$\textit{If signal is appetitive}, \quad \textit{do appetitive response}, \tag{1}$$
$$\textit{If signal is aversive}, \quad \textit{do aversive response}, \textit{ and} \tag{2}$$
$$\textit{If signal is neutral}, \quad \textit{do not respond} \tag{3}$$

Suppose we have external sensor signals,

$$S^{ext}(t) = \{s_1^{ext}(t), s_2^{ext}(t), s_3^{ext}(t), \ldots s_M^{ext}(t)\} \tag{4}$$

and basic motor command signals,

$$S^{mbase}(t) = \{s_1^{mbase}(t), s_2^{mbase}(t), s_3^{mbase}(t), \ldots s_N^{mbase}(t)\}, \tag{5}$$

where $s_i^{ext}(t)$ is the i-th external sensor signal at time t, $i \in \{1, 2, \ldots, M\}$ and $s_j^{mbase}(t)$ is the j-th basic motor command signal, $j \in \{1, 2, \ldots, N\}$. Then a machine can exhibit reflexive behavior by executing the behavior rules of (1), (2), and (3) by a function,

$$S^{mbase}(t+1) = H(S^{ext}(t)) \tag{6}$$

In essence, the function H is a direct mapping between observed signals and movements. The movements are predefined by the basic motor command signals, defined and programmed by humans. For example, suppose that the machine has a touch sensor to detect an obstacle. When the touch sensor is triggered (aversive signal), the machine moves away from it (aversive response). Suppose also that the machine has a photo sensor to detect a light signal coming from the battery recharge station. When the light is detected (appetitive signal), the machine approaches toward the light (appetitive response).

Reflexive behavior reflects the machine's innate value system, a mechanism that detects meaningful features of signals. It is a design task for humans to define meaningfulness in terms of the machine's sensing and acting capabilities in the given environment. The innate value system is then constructed as a signal-response behavior program as described above. By defining the innate value system in this way, humans can always observe the machine's internal judgment system.

4) OPERATION PRINCIPLES OF SELF-EXPLORATORY BEHAVIOR

Self-exploratory behavior is innate and random. The movement is innate because it is given to the machine at its birth, like reflexive behavior. The random motion causes the internal sensors to detect signals that indicate actuator positions (proprioception), whole body movement, and energy consumption. By associating its body movement with internal sensation, the machine learns how to move in the environment. The ability to use the signals to learn motor skills is defined and programmed by humans.

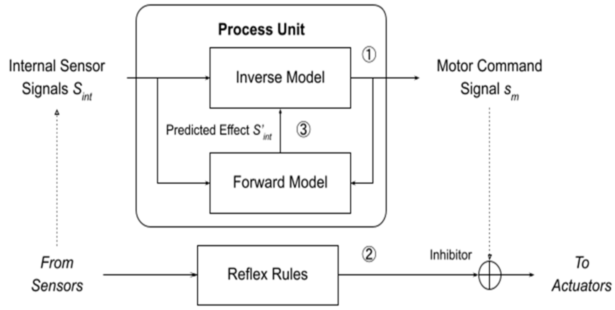

FIGURE 9. Motor skill learning by forward-inverse mechanism.

Suppose that the machine randomly executes the basic motor movements $S^{mbase}(t)$. The movement influences the values of internal sensor signals,

$$S^{int}(t) = \{s_1^{int}(t), s_2^{int}(t), s_3^{int}(t), \ldots s_P^{int}(t)\}, \tag{7}$$

where $s_k^{int}(t)$ is the k-th internal sensor signal at time t and $k \in \{1, 2, \ldots P\}$. The internal sensor signals include proprioception signals from motor encoders, motion signals from 3-axes accelerometers, and energy signals from battery voltage sensors, for example. The relationships between the executed movements and internal sensor signals can be interpreted in two directions: from the movements to the observed signals, and from the observed signals to the movements.

In the first case, the motor command signals are treated as input, and the observed internal sensor signals as the outputs. In this case, the learned input-output relationship is a function,

$$S^{*int} = f_{int}\left(s_k^{mbase}, S^{int}\right), \tag{8}$$

where S^{*int} is the predicted value of the internal sensor signal. In the second case, this model is reversed by treating the internal sensor signals as inputs and the motor command signals as the output. In this case, the learned input-output relationship is a function,

$$S^{*mbase} = g_{int}(S^{*int}, S^{int}), \tag{9}$$

where S^{*mbase} is the necessary motor command signal, given the desired effect S^{*int} and the internal sensor signal S^{int}.

Both models represent causal relationships but in different ways. The former is often called a forward model and the latter an inverse model [116]. Together the two models provide the ability to control the motor behavior precisely and purposefully. In particular, the inverse model can serve as a feedforward controller of the motor actions [117]. Fig. 9

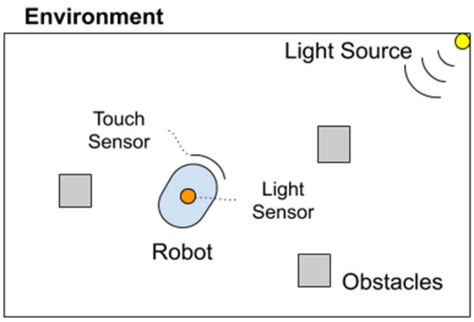

FIGURE 10. Operation environment of "Machina Speculatrix."

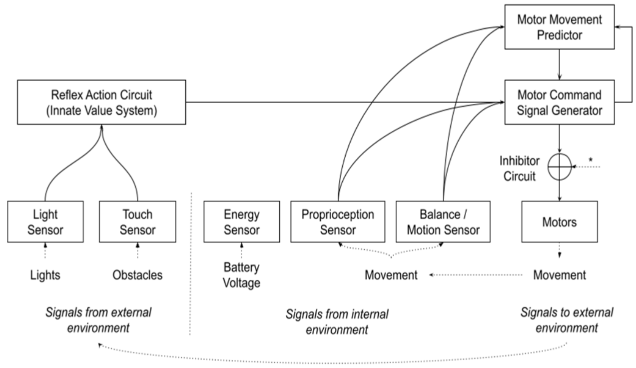

FIGURE 11. Illustration of operation process for self-explanatory and reflexive behavior.

shows the conceptual diagram of motor skill learning at the early stage of behavior development based on the forward-inverse mechanism.

In the beginning, the machine exhibits babbling movements as self-exploratory behavior. Motor command signals are randomly sent from the inverse model of the process unit as the model has not learned anything yet (noted ① in the diagram). While the machine moves around, it may detect signals that cause reflexive behavior (noted ② in the diagram). The babbling movements and reflex actions perturb internal sensor signals, which allow the forward model in the process unit to map the relationship between the motor command signals and their effects. The forward model feeds the predicted effects to the inverse model, completing a feedback loop (noted ③ in the diagram). Note that the reflex behavior rules are placed outside of the process unit, indicating the direct hardwired stimulus-response mechanism of reflex actions. In practice, this is not necessary, and it is a design choice whether to place it inside or outside of the process unit.

Conceptually speaking, the forward-inverse mechanism is a system of acquiring motor skills for desired outcomes in the form of multiple inverse models. The increased number of inverse models is a natural consequence of accumulating procedural knowledge about the movement. More complex the movements and effects become, more inverse models to store the knowledge. In this context, the associative learning between movements and the internal effects by the forward-inverse mechanism can be considered a prototype of motor skill learning. Based on the prototype, the learning mechanism can in principle be extended to more advanced motor skill learning by using the external sensor signals, as the motor movements by self-exploratory and reflexive behavior also elicit changes in external sensor signals.

5) SYSTEM AND PROCESS REQUIREMENTS

Based on the analysis, the system components required for machines to exhibit self-exploratory and reflexive behavior are listed below:

1) sensors to detect aversive and appetitive signals,
2) actuators to execute predefined movements for aversive and appetitive actions,
3) sensors to detect internal conditions, such as proprioception, motion, and energy status,
4) at least one processor as a process component to convert sensor signals to actuator control signals,
5) a power supply for sensing, acting, and processing components,
6) a chassis to hold the above components as a single operation body, and
7) a recharge mechanism for the power supply (to play the game of survival).

The process component provides three major functions. First is the innate value system that triggers predefined actuator command signals in response to predefined sensor signal conditions. Second is the self-exploratory mechanism that sends basic actuator command signals randomly. Third is the motor skill learning mechanism that builds a model between the basic actuator command signals and the internal sensor signals.

6) EXAMPLE OF REFLEXIVE AND SELF-EXPLORATORY BEHAVIOR

Let us look at an example that illustrates how a machine with the required system components exhibits self-exploratory and reflexive behavior. In 1950, Grey Walter constructed a robot called *Machina Speculatrix* which exhibited reflexive and self-exploratory behavior (but lacks motor skill learning) with two sensors and two actuators [34]. Fig. 10 shows the experimental configuration.

The touch sensor generates a signal that represents an aversive unconditioned stimulus. This signal carries an environmental cue that indicates negative values (e.g., poison, predator) to the robot's value system (e.g., health, tasks). The light sensor generates a signal that represents an appetitive unconditioned stimulus. This signal carries an environmental cue that indicates positive values (e.g., food, mate) to the robot's value system. These stimuli are called unconditioned because they do not depend on previous experience. The reflex action circuit takes unconditioned stimuli as inputs and triggers aversive or appetitive unconditioned responses. For example, if the robot's touch sensor is turned on, it triggers an aversive response by moving away from the direction of touch because touching an obstacle is a negative event as it blocks the robot from seeking food (lights). Similarly, if the robot's light sensor detects a light source, it triggers an appetitive response by moving toward the light source because the light is where the food is.

There is an additional response called inhibition in *Machina Speculatrix*, which inhibits either an appetitive or aversive response due to the value system's priority. For example, an appetitive response is inhibited when an obstacle is detected in the direction of the light source because the value system prioritizes the aversive over appetitive. This could change if, for example, the robot is extremely hungry and pushing over obstacles can be accepted in such an emergency.

Fig.11 illustrates the operation process of self-exploratory and reflexive behavior for the example machine.

7) VALENCE, EXTRINSIC-INTRINSIC MOTIVATION, AND VALUE SYSTEM

The innate value system that enables a machine to exhibit reflexive behavior is programmed as a simple binary mechanism based on positive and negative sensations. This binary judgment mechanism is inspired by human behavior. Meaningfulness of observation is described and analyzed as valence in psychology. Valence is a conceptual term that refers to the quality of an event, object, or situation, representing its attractiveness and averseness [141]. Valence derives from a person-situation relation [142]. In other words, the positivity or negativity of an event is subjectively determined by the perceiver of an event and is not necessarily inherent to an event. For example, a rainy day may be perceived negatively by a girl who was planning on outdoor play, but it may be seen positively by her brother who wanted to stay home and read a book instead. Or the same girl may find the rainy day positive later when she had fun stepping on a mud puddle.

Brendl and Higgins identify four principles of judging valence: goal supportiveness, membership status, referential status, and response elicitation [142]. The principle of judging valence based on goal supportiveness refers to a condition of whether an event facilitates or impedes the satisfaction of a goal. Applying this principle to a machine as an example, a self-driving vehicle is approaching a charging station and suddenly sees an obstacle ahead. The object in this case is perceived negatively as it impedes the satisfaction of recharging its battery.

The principle of judging valence based on membership status refers to a condition of whether an event is associated with an already valenced representation. Applying this principle to a machine as an example, a mobile robot is roaming around in a playground and suddenly hears a whistle sound. A moment later the robot collides with an obstacle. After experiencing the same incident numerous times, the previously neutral event of a whistle sound becomes associated with a negative event of obstacle collision. Next time a whistle sound is heard, the robot perceives it negatively because the event is associated with a negative valence experience.

The principle of judging valence based on referential status refers to a condition of an event in comparison to a reference point. Applying this principle to a machine as an example, a mobile robot that just navigated through a maze in 30 seconds may perceive the time positively if it was the shortest time of all its previous tries. However, the record time may be perceived negatively if the average time of all competing robots was much shorter at 20 seconds.

The principle of judging valence based on response elicitation refers to a condition where an event's valence is inferred from the observation of other instances. Applying this principle to a machine as an example, a mobile robot is navigating through a maze and approaches a T junction. Prior to that, the robot observed that all other robots turned left at the T junction. By making an inference, the robot perceives the left turn positively and the right turn negatively.

Brendl and Higgins' framework highlights the complexity of how circumstances could result in different perceptions of observed events. Particularly relevant to machine behavior is the goal supportiveness principle. Machines are typically designed to perform tasks to achieve a goal specified by humans. Positive and negative valence can be determined by successful completion of the assigned task. Ryan and Deci define extrinsic motivation as "*doing something because it leads to a separable outcome*" [99]. Successful completion of an assigned task results in a separable outcome. Based on valence in goal supportiveness and motivation, it can be said that a machine at its creation behaves in the Ryan-Deci sense of extrinsic motivation.

A comprehensive value system for developmental autonomous machines shall not be static. While the innate value system reflects the separable outcome that the machine seeks, it is expected to evolve from the initial binary mechanism to a complex system of judgment by learning from various circumstances that the machine experiences in its lifetime. A machine associates novel signals with known features because it is meaningful to the machine. It is an inherent process to do something because it is meaningful. According to Ryan and Deci, "*doing something because it is inherently interesting or enjoyable*" is intrinsic motivation [99]. Based on this definition and logic, if the process of behavior development is driven by a mechanism that associates novel signals with known meaningful features, then the subsequent behaviors arising from the process of association can be considered intrinsically motivated. According to Ryan and Deci, intrinsic motivation is "*a critical element in cognitive, social, and physical development because it is through acting on one's inherent interests that one grows in knowledge and skills*" [99].

C. CAUSALITY AND MECHANISM OF EPISODIC BEHAVIOR

1) OVERVIEW

Once a machine is released to the environment, the game of survival begins. Initially, the movement is expected to appear

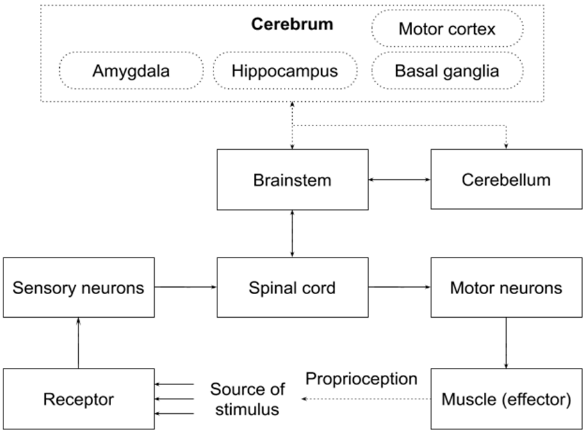

FIGURE 12. Neural mechanism of conditional behavior.

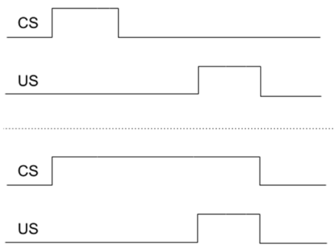

FIGURE 13. Trace and delay conditioning.

primitive and random, and the machine only reacts to certain signals, judged by the value system designed by humans. The rest of observed signals are treated as neutral, and no response is observed. If there is a meaningful feature in the signal, the machine should exploit it, if it knows how. For example, if a sound of a whistle consistently precedes an aversive signal of obstacle collision, then the sound event can be a resource for the machine to anticipate the obstacle event. By exploiting such a resource, the machine could save itself from a collision. The machine needs to learn this skill from observation to survive. What additional functions should a machine possess to exhibit this episodic behavior?

2) NEURAL BASIS OF EPISODIC BEHAVIOR IN HUMANS

The recent findings in the neural basis of episodic behavior posit that the brainstem and cerebellum are the primary regions involved in conditioning [143], [144]. As in reflex actions, sensory signals are processed through the sensorimotor mechanism with signal pathways from sensory neurons to the spinal cord, and from motor neurons to the muscles (Fig. 12). Conditioning occurs when conditioned stimulus (CS) and unconditioned stimulus (US) converge in various regions of the cerebellum and brainstem. However, the exact mechanism of conditioned behavior is still unclear and appears to vary depending on the types of conditioning involved.

For example, conditioning of two different types of stimuli presentations involve different brain regions. In case of trace conditioning where CS and US do not overlap in time (Fig. 13, top), the hippocampus appears to be involved in conditioning in addition to the primary regions of cerebellum and brainstem, but in case of delay conditioning where CS and US overlap in time (Fig. 13, bottom), the hippocampus is not involved [145]. In fear conditioning, the amygdala and other high brain functions appear to be involved [144].

The fact that the exact mechanism of conditioned behavior is still unclear in neuroscience highlights the complex nature of the brain and behavior. What appears to be different between reflexive and episodic behaviors in the brain is the possible involvement of hippocampus and amygdala in episodic behavior. These two regions are located deep inside the cerebral cortex, just above the brainstem [146]. The current understanding is that hippocampus is known to be involved with memory, amygdala is involved with emotion, and basal ganglia, also located in the same region, is known to participate in motor control [146]. Therefore, the involvement of these three deep-lying structures indicates the role of memory and emotion in episodic behavior.

It has been argued that the conditioned behavior is not a simple memory-based reflex and involves complex learning processes, according to Rescorla [147]. Gallistel and Balsam [148] shared this view and argued that the temporal contiguity of events, as has been widely discussed, may not be the sole mechanism of conditioned behavior, and instead suggested a learning mechanism that maps the temporal organization of sequential events. As Thompson observed [145], the involvement of hippocampus in trace conditioning indicates the role of episodic memory for discontiguous events, while the absence of it in delay conditioning may indicate a different mechanism for contiguous events.

Traditionally, the hippocampus has been considered the primary component for episodic memory formation, and this view continues to be dominant in research (see Moscovitch, et. al. [149] for overviews of this research thread). However, there is a growing trend to consider synergistic interaction between hippocampus and amygdala. From an evolutionary perspective, brain mechanisms have evolved to store information that are more "interesting" than trivial events. The interestingness is often reflected in emotion in animals and humans, and many experimental findings show that emotionality enhances memory [150].

From these findings, Richter-Levin and Akirav [151] proposed a combined system of amygdala and hippocampus that modulates episodic memory. Yang and Wang [152] describes the neural circuits of amygdala-hippocampus interplay for memory modulation. The system of amygdala-hippocampus interaction for memory is based on the hypothesis that, when an event elicits emotional arousal, the amygdala mediates the release of adrenal stress hormones, e.g., epinephrine and glucocorticoids, which in turn modulate the hippocampal memory function [151]. In other words, while the hippocampus may be the primary operator of memory, the amygdala may be the regulator of memory formation and retrieval by judging the interestingness of an event.

3) OPERATION PRINCIPLES OF EPISODIC BEHAVIOR IN MACHINES

The key functional difference between reflexive and episodic behaviors is memory. Episodic behavior would not arise if a machine had no means to remember its experience. The machine also needs sensors to recognize events that may be meaningful but not yet appraised by the innate value system.

The processing component temporarily stores the occurrence of neutral signals and extracts potential features that correlate with known signals. Such correlation can be temporal, spatial, or both, depending on the configuration. In case of high correlation above certain thresholds, the observed feature in the neutral signal becomes a sign that is associated with the established signal. Because the innate value system is a trigger mechanism that issues basic motor command signals, when the newly associated (formerly neutral) signal is detected, it is interpreted as a sign to trigger the associated motor command, as if the original appetitive or aversive signal was detected.

This process of observational association is made possible by having a memory storage and an ability to extract features that correlate with known signal features. The concept is inspired by the neural basis of episodic behavior in humans. Even though the precise understanding still needs further research, the hypothetical system of amygdala-hippocampus circuitry in human brains conceptually implies an appraisal mechanism that identifies meaningful features in novel experiences. In a simplified view, episodic memory emerges because of positive or negative valence of signals received from the environment, based on the innate value system defined for reflexive behavior. The associative learning process measures the inherent values of novel signals with respect to their correlations with the known signals.

To illustrate this mechanism, let us consider a simple example by using probability measures. Suppose we have three external sensor signals,

$$S^{ext}(t) = \{s_1^{ext}(t), s_2^{ext}(t), s_3^{ext}(t)\} \tag{10}$$

and two basic motor command signals,

$$S^{mbase}(t) = \{s_1^{mbase}(t), s_2^{mbase}(t)\}. \tag{11}$$

Suppose also that we have a reflexive behavior function,

$$S^{mbase}(t+1) = H(S^{ext}(t)) \tag{12}$$

that map $S^{ext}(t)$ to $S^{mbase}(t+1)$ according to the following behavior rules:

$$\text{If } s_1^{ext}(t) \neq 0, \quad \text{then } s_1^{mbase}(t+1) = 1 \tag{13}$$

and

$$\text{If } s_2^{ext}(t) \neq 0, \quad \text{then } s_2^{mbase}(t+1) = 1. \tag{14}$$

This means that a non-zero value in the external sensor signal s_1^{ext} at time t triggers a movement of s_1^{mbase} at time $t+1$. Similarly, a non-zero value in the external sensor signal s_2^{ext} at time t triggers a movement of s_2^{mbase} at time $t+1$. Whatever values in the external sensor signal s_3^{ext} do not trigger any movement at this point.

Now suppose we program a processing component to calculate three probability measures $P\{s_1^{ext} \neq 0\}$, $P\{s_3^{ext} \neq 0\}$, and $P\{s_1^{ext} \neq 0 | s_3^{ext} \neq 0\}$. $P\{s_1^{ext} \neq 0\}$ is the probability of s_1^{ext} occurring, $P\{s_3^{ext} \neq 0\}$ is the probability of s_3^{ext} occurring, and $P\{s_1^{ext} \neq 0 | s_3^{ext} \neq 0\}$ is the conditional probability of s_1^{ext} occurring when s_3^{ext} also occurs. With these probability measures, one can determine:

(a) the significance of s_3^{ext} with respect to s_1^{ext}, and

(b) the relationship between s_3^{ext} and s_1^{ext} in terms of their occurrences.

By using the Bayes theorem, we can establish a relationship among the probability measures,

$$\begin{aligned} P\{s_1^{ext} \neq 0 | s_3^{ext} \neq 0\} &= P\{s_3^{ext} \neq 0 | s_1^{ext} \neq 0\} \\ &\quad \times P\{s_1^{ext} \neq 0 \div P s_3^{ext} \neq 0\} \end{aligned} \tag{15}$$

If the probability of s_1^{ext} occurring is similar to the probability of s_3^{ext} occurring, i.e.,

$$P\{s_1^{ext} \neq 0\} \simeq P\{s_3^{ext} \neq 0\} \tag{16}$$

then the neutral signal s_3^{ext} is as significant as s_1^{ext}, confirming the condition (a). If the conditional probability

$$P\{s_1^{ext} \neq 0 | s_3^{ext} \neq 0\} \geq k, \tag{17}$$

where k is an arbitrary threshold number, then it means that the neutral signal s_3^{ext} is a predictor of s_1^{ext}, confirming the condition (b). If (a) and (b) are both confirmed, then s_3^{ext} is considered a reliable predictor of s_1^{ext}. In this case, s_3^{ext} is associated with s_1^{ext} and the movement s_1^{mbase}. As a result, the machine exhibits episodic behavior based on the following condition and action formula:

$$\text{If } s_3^{ext}(t) \neq 0, \text{ then } s_1^{mbase}(t+1) = 1 \tag{18}$$

For example, let s_3^{ext} be a whistle sound and s_1^{ext} an obstacle collision signal. If the experience of s_3^{ext} and s_1^{ext} meet the conditions of (a) and (b) by the equations (16) and (17), then the next time the machine hears the whistle sound, it executes the back off movement as if it detected an obstacle.

The probability measures can change in a reverse mode. For example, suppose the whistle sound stops occurring. In that case, the conditions (a) and (b) no longer hold, and the behavior rule (18) becomes invalid, and s_3^{ext} no longer triggers s_1^{mbase}. This is the situation of forgetting a learned behavior.

As a primitive method of observational association, the process above can be generalized by an event relation map shown in Table 3. For n-dimensional events e_i, $i \in [1, n]$, $p\{e_i\}$ is the probability of an event occurring, and $p\{e_i | e_j\}$ is the conditional probability of e_i given e_j, where $i, j \in [1, n]$.

A mechanism for machines to exhibit episodic behavior can be done by referring to the values specified in the map. For example, suppose e_1 is an aversive event and e_3 is a neutral event. If $p\{e_1\}$ and $p\{e_3\}$, the probabilities of the events

TABLE 3. Event relation map.

	e_1	e_2	…	e_n
e_1	$p\{e_1\}$	$p\{e_2 \mid e_1\}$	…	$p\{e_n \mid e_1\}$
e_2	$p\{e_1 \mid e_2\}$	$p\{e_2\}$	…	$p\{e_n \mid e_2\}$
…	…	…	…	…
e_n	$p\{e_1 \mid e_n\}$	$p\{e_2 \mid e_n\}$	…	$p\{e_n\}$

e_1 and e_3 occurring are similar, and $p\{e_1|e_3\}$, the conditional probability of e_1 conditioned on e_3 is high, then e_3 triggers an action associated with e_1 as if e_1 occurred even in the absence of e_1. Note that the method described is for an illustrative purpose only, and not a claim for a universal or general mechanism for episodic behavior.

To exhibit episodic behavior, a machine needs the following components:

1) the system of reflexive and self-exploratory behaviors,
2) sensors to detect neutral signals, and
3) memory storage.

4) EXAMPLE OF EPISODIC BEHAVIOR EMERGENCE IN MACHINES

Grey Walter demonstrated episodic behavior with his second robot *Machina Docilis* [35]. *Machina Docilis* has two new elements added to the previous robot *Machina Speculatrix*: a neutral stimulus and an associative memory. A newly added sound sensor is to detect a whistle sound, which is blown just before the robot sees a light, or just before the robot touches an obstacle. Initially the whistle sound is a neutral stimulus that causes no effect on responses. When a whistle is blown just before the robot sees a light, the robot continues its behavior of moving toward the light without being affected by the whistle. After repeating the same whistle event 10 to 20 times, the whistle event is associated with the light event. The whistle event is no longer neutral and becomes a conditioned stimulus. As a result, when the whistle event is detected, the robot moves toward the light as if it saw a light.

Another example is when the whistle is blown just before the robot touches an obstacle. By the same mechanism described before, after repeating the event 10 to 20 times, the whistle event is associated with the aversive unconditioned stimulus, triggering a behavior of moving away from an obstacle when the whistle event is detected. If the condition continues with the same follow-up event (i.e., see a light or touch an obstacle), the conditioned stimulus is strengthened, stays in the memory, and the episodic behavior continues. However, if there is no light or obstacle after the whistle for a number of times, the association will be weakened and forgotten eventually, and the episodic behavior will not be exhibited as a result. Fig. 14 illustrates the operation process of episodic behavior for the example machine.

There are alternative mechanisms to implement episodic behavior in machines. For example, Pritzel et. al. [153] incorporates episodic memory in a deep reinforcement learning framework and shows significant performance improvement. The mechanisms illustrated so far seem oversimplified, and the probability calculation may appear biologically unsound. Indeed, this sentiment is articulated by Rescorla as he wrote, "*Most of us are not comfortable with the notion that organisms take in large blocks of time, count up numbers of unconditioned stimulus events, and somehow arrive at probability estimates*" [154]. It is clear in our objectives that machines need not replicate biological mechanisms unless there is such a requirement in the machine's purpose.

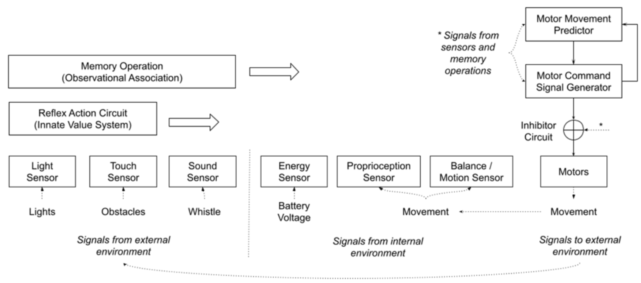

FIGURE 14. Illustration of operation process for episodic behavior.

For developmental autonomous machines, it is only the beginning to learn to utilize the ability to remember signals that are potentially related to its intrinsically meaningful signs (i.e., features of appetitive and aversive stimuli). It slowly accumulates observed signals that are related to appetitive and aversive signs in terms of their contiguity of events or proximity of objects. These are some of the resources immediately available for the machine to exploit from the environment at this point. This changes soon when the machine learns to associate its own actions with observations.

D. CAUSALITY AND MECHANISM OF PROCEDURAL BEHAVIOR

1) OVERVIEW

With episodic behavior emerging, the machine begins to exploit the environmental cues by associating novel signals with intrinsically meaningful signals. However, such behavior is still involuntary and not a purposive act. To survive, the machine must learn to drive itself to choose an action that finds the energy source while avoiding threats. For this to happen, the machine must be able to associate its own movement and the consequence. Because a meaningful event can be consequential to its action, a new way to exploit the environment could emerge by associating its actions with resulting changes in the environment. The machine takes an action, observes the consequence, chooses another action, and repeats the process. The basic tenet of procedural behavior is that some actions may be chosen more often than the others based on the chance of resulting in a favorable outcome. But how does a machine know what a favorable outcome is?

For non-developmental machines, this is not a problem because desired outcomes are determined a priori directly or indirectly by humans in the form of goals, tasks, or rewards. All conceivable circumstances are predetermined and factored into the programmed procedure by human designers. Designing a non-developmental machine therefore is reduced to finding an action that leads to the desired outcome. This can be solved directly by programming the best-known actions, or indirectly by programming an algorithm that determines possible actions. Most commercial machines are built by the former approach, and an increasing number of experimental machines are built by the latter, which is often known as the machine learning approach.

The situation is different for developmental machines playing the game of survival. Favorable outcomes depend on circumstances which are unknown a priori and change dynamically. Moreover, at the early stage of behavior development, the notion of outcome or consequence does not exist in a developmental machine. It can only move its actuators randomly for basic motor skill learning, and innately react to appetitive and aversive signals and their episodically associated signals. Unless something inside changes, developmental machines do not exhibit purposive acts.

The notion of outcome or consequence could emerge if the machine has an ability to recognize the causal relationship between actions and observation signals. The associated observation signals can be perceived as outcome or consequence, which can be evaluated subsequently by a value system to determine what is favorable or not. The problem is that the innate value system is not designed to perform this function. This is the missing piece for developmental machines to generate purposive acts.

2) NEURAL BASIS OF PROCEDURAL BEHAVIOR IN HUMANS

Trial-and-error behaviors can be described as a repetitive and coordinated act of sensing, moving, judging, and remembering. As identified in the reflexive and episodic behaviors, the major components in the brain that handle sensing and moving are the spinal cord, brainstem, cerebellum, and sensory and motor neurons. They function together as a sensorimotor system, and possibly take part in conditioning as well. As identified in the episodic behavior, the amygdala and hippocampus are known to be involved with judging, remembering, and conditioning. Obviously, the brain mechanism cannot be simplified as such, and the exact mechanism is still an active topic in neuroscience.

Recent studies have shown that a neurotransmitter called dopamine is found to be associated with reward situations as a reward prediction error [155], more general prediction error [156], and even generalization and bonus [157]. More recently, dopamine is also found to be involved in penalty situations [158], [159]. Dopamine is mainly synthesized in dopaminergic neurons located in a region called substantia nigra in the basal ganglia and is projected to another structure in the basal ganglia, called striatum [160]. The role of substantia nigra and striatum as a dopaminergic circuit has been computationally modeled and applied as reinforcement learning [55]. These components are shown in Fig. 15.

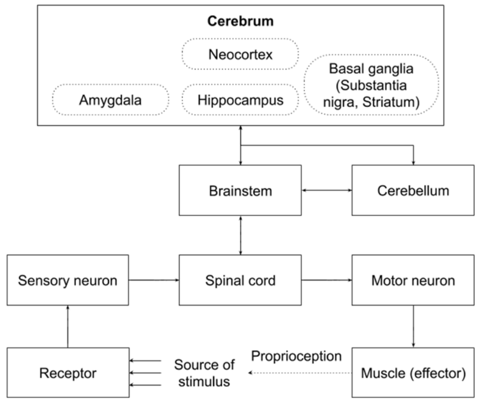

FIGURE 15. Neurostructural mechanism of procedural behavior.

In contrast to the structural primacy in neural mechanisms for reflexive and self-exploratory behavior, the neural dynamics for procedural behavior appear to be strongly influenced by the biochemical actions of neurotransmitters. Neurotransmitters are signal carriers and fundamental in all types of behaviors, but the current research in neuroscience appears to indicate that the dynamics of purposeful activities are more explainable by biochemical signals than the structural properties of the brain alone. Perhaps it may appear so due to the lack of understanding in neural structures. However, purposive actions are highly dynamic and circumstantial. To cope with fast changing events in the environment, fluid and dynamic properties of biochemical processing seem critical in explaining such phenomena. In other words, the neural mechanism behind purposive behavior is no longer just a matter of structures and connections in regional neural clusters. What values signals carry begin to matter.

3) OPERATION PRINCIPLES OF PROCEDURAL BEHAVIOR IN MACHINES

a: NAIVE STATISTICAL APPROACH

What causes a developmental machine to exhibit purposive acts? Before diving into the solution, let us review a few classical methods to highlight what is missing in the current machine learning approaches.

Purposive acts are driven by the expectation about future outcomes with respect to its own actions. Altering behavior based on future outcomes is a type of anticipatory behavior, and prospective in the sense that the expected change is caused by its own purposeful actions. Based on the classification of anticipatory behavior as discussed in Section II-C, this type of behavior can be considered a payoff or state anticipatory mechanism. A payoff anticipatory mechanism uses an expected payoff to choose an action, while a state anticipatory mechanism uses an explicit predictive model of the environment to predict the future states [133].

To illustrate the nature of the problem, let us consider a simple approach with probability measures as a payoff function. Suppose we have external sensor signals,

$$S^{ext}(t) = \{s_1^{ext}(t), s_2^{ext}(t), s_3^{ext}(t), \ldots s_M^{ext}(t)\} \tag{19}$$

and basic motor command signals,

$$S^{mbase}(t) = \{s_1^{mbase}(t), s_2^{mbase}(t), s_3^{mbase}(t), \ldots s_N^{mbase}(t)\}, \tag{20}$$

where $s_i^{ext}(t)$ is the i-th external sensor signal at time t, $i \in \{1, 2, \ldots, M\}$ and $s_j^{mbase}(t)$ is the j-th basic motor command signal, $j \in \{1, 2, \ldots, N\}$. Note that the motor command signals are still basic movements used in reflexive and self-exploratory behavior.

Before a machine can choose an action, it first needs to learn how to associate its actions and their consequences so that it knows what action leads to a certain outcome. One way to approach the problem is to treat the relationship between actions and consequences from a statistical point of view. For example, we can calculate the conditional probability $P\{s_{desired}^{ext}(t)|s_{executed}^{mbase}(t-1)\}$ of desirable event $s_{desired}^{ext}(t)$ occurring, conditioned upon an executed motor command signal $s_{executed}^{mbase}(t-1)$.

Because a desirable event is unknown, it must be determined somehow. A simple method, though admittedly not a good one, is to compare the conditional probabilities of all events and choose the highest value. To formulate this approach, we define a value function V,

$$V(t) = \max\left\{P\{s_i^{ext}(t)|s_j^{mbase}(t-1)\}\right\} \tag{21}$$

for all i and j, where $i \in \{1, 2, \ldots, M\}$ and $j \in \{1, 2, \ldots, N\}$.

By assuming that signal detection is considered an event, the statistical relationship between events and actions can be expressed by the conditional probability of e_i given a_j, $p\{e_i|a_j\}$ for M-dimensional events e_i, $i \in [1, M]$ and N-dimensional actions $a_j, j \in [1, N]$, as illustrated in Table 4.

If i and j in (21) are not restricted, the highest probability might occur for an unwanted event of negative consequences. The machine may end up choosing an undesirable action. Since the innate value system specifies appetitive and aversive signals, it can be applied to choose an action that leads to an appetitive event only. For example, suppose that the second external sensor signal $s_2^{ext}(t)$ is an appetitive signal, then the value function can be restricted and becomes

$$V_2(t) = \max\left\{P\{s_2^{ext}(t)|s_j^{mbase}(t-1)\}\right\} \tag{22}$$

for all j, where $j \in \{1, 2, \ldots, N\}$.

This naive approach described above has many shortcomings. First, the statistical measure disregards the exact state of where the machine is. An action with the highest conditional probability over all states is not necessarily the best action at a given moment. Second, the payoff of an action is limited to the known outcome defined by the innate value system. Desired outcomes are circumstantial, unknown in advance.

TABLE 4. Action-event relation map.

	e_1	e_2	…	e_M
a_1	$p\{e_1\|a_1\}$	$p\{e_2\|a_1\}$	…	$p\{e_M\|a_1\}$
a_2	$p\{e_1\|a_2\}$	$p\{e_2\|a_2\}$	…	$p\{e_M\|a_2\}$
…	…	…	…	…
a_N	$p\{e_1\|a_N\}$	$p\{e_2\|a_N\}$	…	$p\{e_M\|a_N\}$

b: REINFORCEMENT LEARNING

Reinforcement learning is a type of machine learning approach that chooses actions based on a payoff anticipatory mechanism and in some cases with a state anticipatory mechanism. In a sense, the naive statistical approach described above is analogous to a greedy policy of reinforcement learning. As intuitively understandable, greedy policies are known to achieve sub-optimal performance [161]. In general, a reinforcement learning method is formulated as an optimization problem.

Richard Bellman formulated an iterative method, called dynamic programming, to solve the optimization problem [162]. The problem is formulated as a Markov Decision Process (MDP). MDP is a sequence of 4-tuples,

$$s_t, a_t, P(s_{t+1}|s_t, a_t), r_t \tag{23}$$

where s_t is a state at time t, a_t is an action, $P(s_{t+1}|s_t, a_t)$ is a probability that the action a_t will lead the state s_t to the next state s_{t+1}, and r_t is the reward received if s_t reaches s_{t+1}. The problem is formulated as to find a series of actions, called an optimal policy, that maximizes the total reward received by bringing the initial state s_0 to its final state s_T.

One approach to find a solution to the MDP problem is called backward induction, where a sequence of optimal actions can be found by reasoning backward in time. Bellman proposed a value iteration method to iteratively compute a value function V for all states s until V converges to a Bellman equation:

$$V_{t+1} = max_a \sum_s P(s_{t+1}|s_t, a_t) \times (r_t + \gamma V_t) \tag{24}$$

where V_t is the value function at iteration t and $\gamma \in [0,1]$ is a discount rate. If all states and rewards are known, then the Bellman equation converges to its optimal value. This is often referred to as the model-based approach because of the use of a priori knowledge of the environment.

Alternatively, instead of using the knowledge of the environment, actual reward received after an action can be used to update the value function. This approach is often called model-free reinforcement learning. One approach is to run simulations to collect sample state and reward information. Based on the sampled data, one can compute the optimal policy. This approach is called the Monte Carlo method [163]. Because it can learn the optimal policy from experience without the knowledge of the environment, this approach can be effective but may not be efficient as running simulations can be resource intensive.

Alternatively, a temporal difference (TD) approach overcomes the efficiency problem and achieves convergence by iteratively updating the value function with a reward prediction error [55]. A simplified learning process can be outlined as follows:

1) *Define a value function,*

$$V(s) = V(s) + \alpha(r' + \gamma V(s') - V(s)) \tag{25}$$

2) *Initialize $V(s)$ with arbitrary values for all s, except the terminal state s_T.*
3) *Choose a step size $\alpha \in$ [0,1] and a discount rate $\gamma \in$* [0,1].
4) *Choose an action.*
5) *Observe the state and reward.*
6) *Update the value function V.*
7) *Set s = s'.*
8) *Repeat 4∼7.*

The model-free reinforcement learning approach chooses an action based on a payoff anticipatory mechanism without relying on an explicit state prediction model. The model-based reinforcement learning approach chooses an action based on a state anticipatory mechanism with a state prediction model. Reinforcement learning is a learning method for machines to execute purposive acts that maximize rewards by supplying the state and reward information with a payoff anticipatory mechanism or an explicit state prediction model. For the state-reward relationship to function as the primary causal mechanism, the mathematical framework of reinforcement learning assumes the existence of a reward signal as a given. For this reason, reinforcement learning requires a supplemental process that generates reward signals. This is an active area of research in relation to the topic of intrinsic motivation in machines.

c: ACTIVE INFERENCE

While reinforcement learning attempts to optimize rewards, there is an alternative approach called *active inference* that attempts to optimize a different entity. Instead of rewards, active inference attempts to minimize so-called an information entropy that represents free energy or surprise to a machine based on the likelihood of possible outcomes [29]. The likelihood of possible outcomes represents a model of the world that the machine expects. A surprise is thus defined as the difference between the world model and observation [164]. The basic principle of active inference is to purposefully choose actions that minimize the surprise based on an explicit state prediction model.

As active inference continuously corrects the world model by actions, the environment is perceived as a system of expected states with minimum surprise or the entropy in the sense of Shannon Entropy. In this sense, active inference can be considered a model-based regulator with its supporting principle derived from the minimum entropy theory by Ashby [51] and the model-based regulator concept by Conant and Ashby [53]. However, this leads to the question of practicality from a control systems perspective.

Stability and uncertainty are two important factors to consider in control systems. The stability of a control system allows the system to reach and remain at the steady state under variations in input variables. The uncertainty of a control system poses difficulties in achieving stability, thus robust control methods are introduced to manage modeling errors. In other words, stability is about minimizing transitional and stead state errors, while robustness is about minimizing uncertainty.

Minimizing surprises in active inference therefore relates to robust control principles. However, active inference does not address transitional or steady state error reduction. This is because the actions are derived from information entropy, which is a measure of variety. The measure of variety is based on the probability distribution of random variables. While minimizing an entropy may reduce the fluctuation of random variables, it has no relation to the values of such variables. Therefore, minimizing an entropy alone has no impact on transitional or steady state error reduction in controlling continuous dynamic systems.

The main challenge of optimal control systems is to achieve stability while minimizing uncertainty. A model-based regulator may minimize uncertainty, but it needs an additional mechanism in principle to address the stability issue. This practical shortcoming is a topic of further research in active inference.

Nevertheless, active inference offers an important clue to the causal mechanism of purposive acts. The question of what drives a machine to trigger a purposeful action can be answered and explained in terms of surprise or deviation from the machine's innate value system.

d: EMERGENCE OF PURPOSIVE ACTS

Let us now explore the puzzle of what causes a machine to exhibit purposive acts. Specifically, we shall investigate what happens when the innate reflex acts fail to achieve the designed outcomes.

Suppose that we have a mobile robot machine with three sensors that detect external signals,

$$S^{ext}(t) = \{s_l^{ext}(t), s_t^{ext}(t), s_s^{ext}(t)\}. \tag{26}$$

$s_l^{ext}(t)$ is the light sensor signal, $s_t^{ext}(t)$ the touch sensor signal, and $s_s^{ext}(t)$ the sound sensor signal at time t. It also has six sensors that detect internal signals,

$$S^{int}(t) = \{s_{bat}^{int}(t), s_{lmot}^{int}(t), s_{rmot}^{int}(t), s_x^{int}(t), s_y^{int}(t), s_z^{int}(t)\} \tag{27}$$

where $s_{bat}^{int}(t)$ is the rechargeable battery voltage indicating the machine's energy level, $s_{lmot}^{int}(t)$ and $s_{rmot}^{int}(t)$ are the left and right motor rotations, respectively indicating the machine's proprioception, $s_x^{int}(t)$, $s_y^{int}(t)$, and $s_z^{int}(t)$ are the x-, y-, and z-axis acceleration, respectively indicating the machine's motion. The machine has four basic motor command signals,

$$S^{mot}(t) = \{s_f^{mbase}(t), s_b^{mbase}(t), s_l^{mbase}(t), s_r^{mbase}(t)\} \tag{28}$$

where $s_f^{mbase}(t)$ moves the machine forward, $s_b^{mbase}(t)$ moves the machine backward, $s_l^{mbase}(t)$ turns the machine left, and $s_r^{mbase}(t)$ turns the machine right.

Suppose that the machine is currently playing the game of survival, and it observes the internal signal $s_{bat}^{int}(t) < v_{low}$ where v_{low} is a threshold value for low battery voltage. The observed signal implies that the machine's energy level is low and that the machine needs to find a battery recharge station quickly before the battery runs out, or else the machine loses the game. Let us assume that the machine is currently operating with only self-exploratory, reflexive, and episodic behavior.

Suppose that the machine detects a light signal while randomly executing the basic motor command signals as self-exploratory behavior. The observed event immediately causes the machine's innate value system to execute a predefined movement as reflexive behavior; in this case a forward move to approach toward the light as it points to the direction of a battery recharge station. It so happens that the forward movement is not straight enough to keep the light in sight. The machine consequently loses the light signal and resorts back to its random babbling actions. In this scenario, the machine experienced two meaningful events: finding a light signal and losing the light signal.

Let us take a closer look at this scenario. At time $t-1$, the light sensor signal $s_l^{ext}(t-1)$ is less than or equal to a predefined threshold k, or $s_l^{ext}(t-1) \leq k$. For the subsequent discussion, set $k = 0$ for simplicity and brevity. Because the light signal points to the direction of a battery recharge station, observing that the light signal changes to $s_l^{ext}(t) > 0$ is a meaningful event for the machine under the circumstance of its battery level being low. The innate value system responds to this appetitive signal by triggering a reflex act $s^{mot}(t+1) = s_f^{mbase}$ at time $t+1$. By design, the outcome from this innate action at time $t+2$ is expected to be $s_l^{ext}(t+2) > 0$, but instead it observes $s_l^{ext}(t+2) = 0$. Losing the light signal is also a meaningful event for the machine under the circumstance. It indicates that the innate reflex act failed to achieve the designed outcome, and the machine needs to find the light signal again to move toward that battery recharge station.

The mismatch between the designed outcome and the observed outcome is a cause for a new desirable outcome. Under this circumstance, finding a light signal or observing a positive change in the light signal $\Delta s_l^{ext} = s_l(T+n) - s_l(T+n-1) > 0$ is desirable while n being smaller the better from the current time T. In this case, finding an action that leads to the desirable outcome can be considered a purposive act.

Similarly, two observable conditions for aversive signals $\Delta s_t^{ext} > 0$ and $\Delta s_t^{ext} < 0$ are also meaningful events. When the machine touches an obstacle, it observes a condition $\Delta s_t^{ext} > 0$ and an innate reflex act kicks in to back away from it. The designed outcome for this reflex act is to cause $s_t^{ext} = 0$. However, if the reflex act does not immediately cause $s_t^{ext} = 0$ to happen for some unknown reason, the machine consequently observes a new condition $\Delta s_t^{ext} = 0$ and $s_t^{ext} > 0$. Under the circumstance when the non-purposive innate reflex act does not lead to a designed outcome, a cause for a desirable outcome and an associated purposive act emerges. In this scenario, any attempt to find an action that leads to the desirable outcome $\Delta s_t^{ext} < 0$ and $s_t^{ext} = 0$ can be considered a purposive act.

The logic also applies to episodically associated actions. For example, suppose that an external sound sensor signal $s_s^{ext}(t)$ is initially neutral but becomes associated with the touch sensor signal $s_t^{ext}(t)$. When the machine observes a condition $s_s^{ext} > 0$, it executes a motor command signal s_b^{mbase} to back away as if it touched an obstacle. This episodically associated action inherits the designed outcome $s_t^{ext} = 0$ from the associated reflex act. If for some unknown reason the observed outcome from the episodic act leads to $s_t^{ext} > 0$, then this is a meaningful event and a cause for a new desirable outcome and a purposive act.

The above examples illustrate the ultimate-proximate explanations of why and how a machine can develop purposive acts. It begins with the innate value system that recognizes inherently meaningful events by triggering reflex acts. When the innate reflex acts fail to achieve their designed outcomes, the unachieved outcome becomes a desirable outcome. The desirability arises due to the circumstance in which the machine is operating at the given moment, e.g., low battery levels. Since actions to achieve the desired outcome are unknown to the machine at the time, it uses an arbitrary mechanism to find favorable actions. Consequently, such an attempt to find favorable actions emerges as a purposive act. Although such a mechanism is arbitrary, the ability to effectively find favorable actions directly influences the machine's chance of survival; therefore, the quality of purposive acts matters in the game.

From the perspective of anticipatory behavior mechanisms as discussed in Section II-C, purposive acts can be executed based on a payoff and/or state anticipatory mechanism. A payoff anticipatory mechanism uses an expected payoff to choose an action, while a state anticipatory mechanism uses an explicit predictive model of the environment to choose an action. As described earlier, model-free reinforcement learning is based on a payoff anticipatory mechanism, model-based reinforcement learning employs both state and payoff, and active inference uses a state anticipatory mechanism. Non-purposive acts such as the innate reflex act can also be considered anticipatory as it uses an implicit anticipatory mechanism without an explicit predictive mechanism. The episodically associated acts are also anticipatory based on a sensorial anticipatory mechanism by using an implicit predictive mechanism to sense the future signals and states. These classifications of anticipatory behaviors are based on the foundational work by Butz et al. [133].

e: CIRCUMSTANTIAL VALUE SYSTEM

The innate value system is a static appraiser of observed signals from sensors. The system's output is designed and fixed a priori based on the predefined judgment rules. While the system is static, the environment is dynamic. Circumstances dictate what a desirable outcome is at any given moment, and for this reason, the innate value system alone is not sufficient to derive favorable actions. It needs a supplemental value system that dynamically recognizes meaningful events and defines what a desirable outcome is at a moment. Let us call such a system a circumstantial value system.

The circumstantial value system is a dynamic appraiser of states. States are conditions that the machine is in at a given moment in time. The system identifies conditions that arise as meaningful to the machine in relation to its innate value system. The system defines desirable conditions that the machine needs to achieve and triggers an arbitrary mechanism that finds such actions. The desirable conditions as outputs of the system are expressed in terms of signals so that a mechanism can be chosen arbitrarily.

When a circumstantial value system identifies a desired condition, a new motivational drive arises and triggers a mechanism to search for motor command signals that result in favorable outcomes. Since numerous approaches and algorithms are possible for such purposive acts, the choice is arbitrary and up to the human designer. With that said however, the operation of procedural behavior in principle begins with an arbitrary action because a good action is not known initially. It tries an arbitrary action and observes the consequence and repeats the process until it finds one that works. As a result, the initial procedural behavior might resemble trial-and-error behavior.

f: SYSTEM AND PROCESS REQUIREMENTS

The system components required for machines to exhibit procedural behavior are listed below:

1) *external sensors to detect aversive, appetitive, and neutral signals from the environment,*
2) *sensors to detect internal conditions, such as proprioception, motion, and energy status,*
3) *actuators to execute movements,*
4) *at least one processor to convert sensor signals to actuator control signals,*
5) *a memory storage,*
6) *a power supply for sensing, acting, and processing components,*
7) *a recharge mechanism for the power supply, and*
8) *a chassis to hold the above components as a single operation body.*

They are essentially the same hardware elements needed for episodic behavior. The difference would be in the process component where the circumstantial value system operates. While the innate value system triggers involuntary reactions in response to fixed sets of conditions, the circumstantial value system responds to conditions that arise when the involuntary reactions fail to achieve desirable outcomes. This means that procedural behavior may not emerge when the innate value system and the corresponding actions

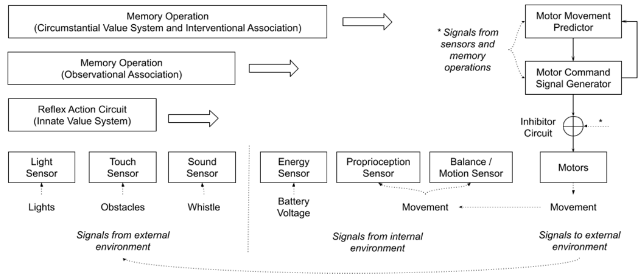

FIGURE 16. **Illustration of operation process for procedural behavior.**

sufficiently maintain the machine in a desirable state. When everything is working fine, there is no need for new behavior. However, when the circumstantial value system recognizes a certain condition, the machine engages in voluntary acts. Fig. 16 illustrates the operation process of procedural behavior.

4) EXAMPLE OF PROCEDURAL BEHAVIOR EMERGENCE IN MACHINES

Consider the previous example where the mobile robot developed episodic behavior. The machine's desirable state is when the battery is sufficiently charged. To remain in that state, desirable actions would be to drive the machine toward a battery recharge station by approaching a light signal while backing away when the touch sensor signal is detected. Because the behavior of self-exploration and reflex acts are only basic and not tuned to the environment, the machine may not find the light signal effectively. Furthermore, when the machine finds a light signal, it may not be able to keep the light in sight while approaching due to poor motor skills.

In this scenario, when a reflex act fails to achieve the designed outcome while the battery level is low, the circumstantially meaningful condition forms a circumstantial value system that can be expressed as:

$$s_{bat}^{int} < v_{low} \ \& \ \Delta s_l^{ext} < 0 \text{ when } s^{mot} = s_f^{mbase} \tag{29}$$

This identified condition triggers a mechanism to find actions s^{mot} that result in a desired outcome $\Delta s_l^{ext} > 0$.

Similarly, another meaningful condition when failing to avoid an obstacle can be expressed as:

$$s_{bat}^{int} < v_{low} \ \& \ \Delta s_t^{ext} > 0 \text{ when } s^{mot} = s_b^{mbase} \tag{30}$$

This identified condition triggers a mechanism to find actions s^{mot} that result in a desired outcome $\Delta s_t^{ext} < 0$.

These circumstantial conditions and desired outcomes are derived from the innate value system, which defines reflex acts in response to sensor values. Since the desirable ranges of internal and external sensor values are known a priori from the innate value system, a circumstantial value system can be programmed as a mechanism to monitor sensor values that deviate from expected values. However, it is not possible to generalize the equations (29) and (30) for all circumstances, unless there is a separate mechanism that provides alternative outcomes. In essence, the role of the circumstantial value system is to recognize circumstantially meaningful events to define desirable outcomes, express them in terms of signal changes that can be addressed by an arbitrary mechanism to find favorable actions.

a: PROCEDURAL BEHAVIOR AND MOTOR SKILL LEARNING MECHANISM

Modeling after the basic motor skill learning mechanism in self-exploratory behavior, procedural behavior can be structured by mapping the causal relationship between the motor command signals and their consequences. By treating the motor command signals as inputs and the observed external sensor signals as the outputs, the input-output relationship,

$$S^{*ext} = f_{ext}(S^{mbase}, S^{ext}) \tag{31}$$

shows the predicted external sensor signals, given motor command signals S^{mbase}. Similarly, the relationship can be reversed by treating the external sensor signals as inputs and the motor command signals as the output. In this case, the input-output relationship,

$$S^{*mbase} = g_{ext}(S^{*ext}, S^{ext}) \tag{32}$$

where S^{*mbase} is the necessary motor command signals that result in desired effects S^{*ext}, given the external sensor signals S^{ext}.

By substituting the predicted external sensor signals S^{*ext} with the function f_{ext}, the necessary motor command signals S^{*mbase} that achieves desired effects can be expressed as a function of external sensor and motor command signals,

$$S^{*mbase} = g_{ext}(f_{ext}(S^{mbase}, S^{ext})). \tag{33}$$

By completing a feedback loop between the motor commands and external signals, the inverse model can generate a motor command that results in a desired effect.

Once a circumstantial value system identifies a desirable condition, motor command signals that result in such an outcome should be associated and memorized in a memory storage. The accumulation of the stored relationships between the desired effects and their causal motor command signals is in principle the process of inverting the forward model. Under the circumstances described in the previous paragraph, the inverse model can retrieve the motor command signals that cause the desired effect from its memory. Because the forward model is not precise or accurate in the beginning, the inverse model is not precise or accurate either. As a result, the trial-and-error behavior may take some time to achieve desired outcomes. As the machine gains experience, the forward model improves its prediction capability, and so does the inverse model.

Fig. 17 shows the conceptual diagram of advanced motor skill learning based on the forward-inverse mechanism. The forward model for the external environment signals maps the relationship between motor signals and the effects on the external sensor signals (noted ① in the diagram). As the meaningful circumstances call for a desired effect, the inverse model issues a motor command signal that is associated with the desired effect, based on the forward model. The motor command is sent to the forward model of the internal environment signals (noted ② in the diagram). The forward model prioritizes this as an interventional signal, generates a feedback signal to the inverse model to generate the motor command signal on to the actuator (noted ③ in the diagram). In this system configuration, all four models are continuously learning and adjusting their models as new data keep coming in. The behavior may be slow and inefficient in the beginning as trial-and-error activities, but eventually the models improve their precision, resulting in more efficient, fluid motions. Ultimately, the inverse models become less reliant on the feedback signals from the forward models, resulting in more feedforward actions as seen as autonomic behavior.

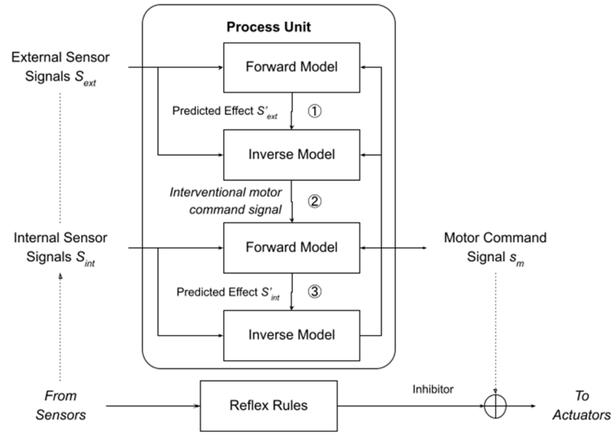

FIGURE 17. Forward-inverse mechanism for interventional and habitual association.

E. CAUSALITY AND MECHANISM OF AUTONOMIC BEHAVIOR

Purposefully achieving a desired outcome is a skill. If the desired outcome requires a fast, smooth physical movement of a body, the skill required is a refined motor skill. In this sense, procedural behavior is not only about selection of actions but also about motor skill learning. Once a suitable action is determined from the trial-and-error iteration, what is needed next is to perform the action efficiently. Autonomic behavior emerges because of procedural behavior and motor skill learning.

The basic principle of autonomic behavior is feedforward execution of learned movements. A movement of a body is initially controlled by a feedback mechanism that regulates actuator control signals based on internal sensor signals. In this case, a sequence of movements is executed as a series of feedback processes. When a certain sequence is frequently repeated over and over, feedback regulation makes the whole process slow and inefficient. Instead, the need for feedback

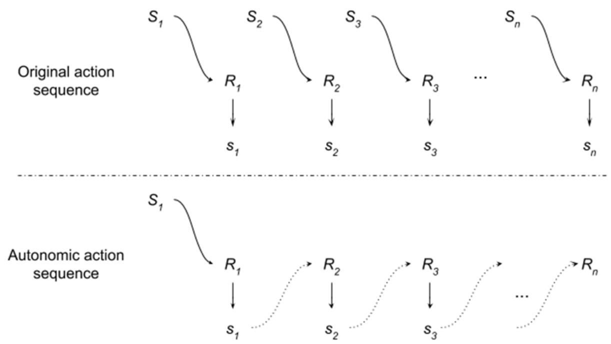

FIGURE 18. Illustration of Hull's habitual association.

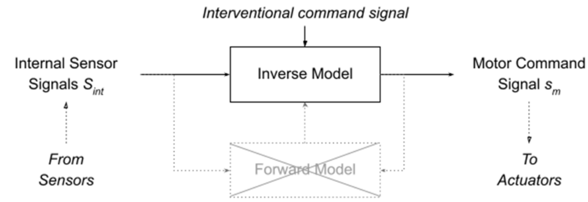

FIGURE 19. Autonomous behavior by inverse model only.

from sensor inputs at each transition can be reduced by associating the sequence of movements. As a result, the whole sequence can be executed by a chain of actuator outputs as a feedforward process. Feedforward is faster than feedback, and the behavior becomes automatic.

Clark Hull originally described this mechanism in a stimulus-response framework. According to Hull [73], when repeating a particular sequence of actions many times, the internal stimulus (i.e., proprioception) from each action becomes associated with the subsequent action. This leads to a condition where a sequence of actions automatically emerges because of chained associations between actions and their internal stimuli.

In Fig. 18, S is a sequence of external stimuli that originally triggers a sequence of responses R. Each response triggers an internal stimulus s by the machine's proprioception. After many repeated executions of the action sequence R, each internal stimulus s_i is associated with the subsequent response $\mathrm{R_i}$. Once the initial response $\mathrm{R_1}$ is executed, the subsequent responses automatically emerge without the presence of external stimuli S.

Recalling the basic motor skill learning mechanism in self-exploratory behavior, autonomic behavior can be structured as a process of reducing the reliance on predicted effects by the forward model when determining the motor command signals. In Fig. 19, the forward model provides predicted effects on the internal sensor signals for the inverse model to generate precise motor command signals. Theoretically speaking, if the precision of the inverse model improves, then the feedback process can become a feedforward process by reducing the role of the forward model.

F. CAUSALITY AND MECHANISM OF CONCEPTUAL BEHAVIOR

1) OVERVIEW

With successful execution of procedural and autonomic behavior, it is conceivable that the machine reaches full autonomy. It could find a way to reach the battery recharge station in time to replenish its power supply. Once the battery is recharged, it could leave the station and explore the environment again. When the battery level becomes low, it could find its way back to the station and recharge again. If the machine can repeat this cycle indefinitely, then according to the rule of the survival game, it can claim itself as a fully autonomous machine.

This is not necessarily a difficult task for non-developmental machines. Their process components are fine-tuned by humans to optimally achieve such tasks, as evidenced by the vacuum cleaner robots of the early 2000's. For developmental machines on the other hand, this has been a long challenging path to reach this level of autonomy, strictly by learning from experience on their own. The game is not over yet though in either case. There are three possible scenarios that could disrupt the machine's full autonomy:

1) *a significant change occurs in the external environment by an exogenous cause,*
2) *a significant change occurs in the internal environment by an exogenous cause, and*
3) *a significant change occurs in the external or internal environment caused by the machine's own actions.*

An example of the first scenario might be a failing light signal from the recharge station. The machine may not be able to find the recharge station without a guiding light. Any alteration in the operating space by removing, inserting, or modifying objects for example, could pose a problem for the machine to maintain its full autonomy.

The second scenario is a reasonable possibility where changes in the internal system occur due to mechanical failure, software malfunction, or any alterations by humans. For example, one of the wheels could get stuck with debris or fall off from the axle, causing it unable to move the way it used to. Things eventually break, and it is likely that machines encounter such unexpected situations during their lifetime. When this happens, the acquired behavior in the past may not be sufficient to maintain its full autonomy.

The third scenario might occur when, for example, the machine accidentally pushes objects or walls of the environment and changes their shapes, locations, or functions. The machine could also move or alter its body parts that may end up changing the behavior of the body.

When the state changes to a point where previously acquired skills become inadequate to achieve the desired state, the circumstantial value system drives a machine to seek new actions. The value system could initially drive the sensorimotor system as procedural behavior to achieve a desired outcome. This physical attempt is in a sense relearning the motor skills to adapt to the new environment.

There are two problems with physically executing procedural behavior. First, it may not always be possible to rely on motor skills due to physical constraints or energy availability. Second, the mechanism to find favorable acts in procedural behavior may be completely irrelevant in the new environment. For example, for a mobile robot that relies on a light signal to locate the battery recharge station, the previously learned behavior would not work if the light signal completely disappeared as the signal source becomes out of order. In other words, a possible solution for the new situation may not exist in the solution space for an optimization scheme in the procedural behavior mechanism.

Another issue is data availability. Signal-based learning requires signal data, but perpetually incoming raw sensor signals cannot be kept in memory forever. What a robot has experienced in the past may not be available when it needs it if the machine's learning is solely based on signal data. For these reasons, machines need a new type of behavior to achieve and maintain their autonomy.

Humans are adept at handling these situations. Natural disasters wreak havoc in our lives (first scenario). We get sick or injured (second scenario). We may break things with our own actions (third scenario). Not all situations are negative. A new technology emerges and brings better ways to do things (first scenario). Living in a healthy environment and digesting good foods could change how the bodies function (second scenario). We build dams, roads, and houses, and as such, we constantly change our environment by our own actions (third scenario).

What humans commonly and naturally do in challenging situations is deliberation, recalling past events, imagining manipulating objects, posing hypothetical what-if questions, and constructing a plan of actions. Such behavior is characterized as a memory process without sensorimotor involvement. The basis of such act is not raw sensor signals but a higher-level representation of the prior experience.

2) NEURAL BASS OF CONCEPTUAL BEHAVIOR IN HUMANS

The neural equivalents of the previous four stages are primarily the combination of reflex pathways (brainstem, spinal cord, sensory and motor neurons), cerebellum, amygdala, hippocampus, and striatum as sensorimotor and implicit memory systems. This is because the behaviors at these stages directly interact with and respond in real time to the environment through sensory receptors and muscle effectors. In contrast, conceptual behavior is primarily a memory operation. What is stored in memory is explicitly declared to be manipulated.

Fig. 20 shows a taxonomy of memory systems and their hypothetical connections to brain structures according to Milner et al. [165]. According to this diagram, the neural basis of conceptual behavior relates to the declarative memory system, while the neural basis of the previous four stages relate mostly to the nondeclarative memory systems. However, the notion of declarative and nondeclarative memories is a matter of conceptual distinction, not necessarily a structural separation in the brain.

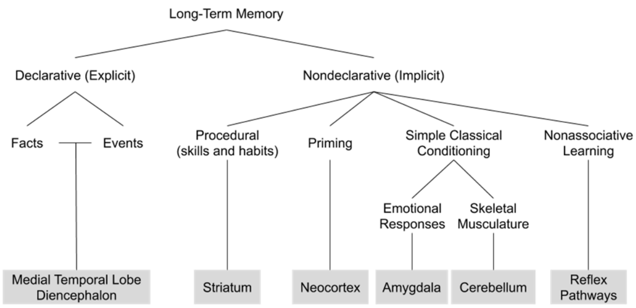

FIGURE 20. Memory systems and brain structures.

As the humans and animals mature to the point where they can function autonomously in the environment with procedural and autonomic behaviors, the major components in the cerebrum are expected to have matured. It is then natural to expect the neural mechanisms for episodic, procedural, and autonomic behaviors to support conceptual behavior by using the existing circuitries, instead of developing new ones specifically for conceptual behavior. This hypothesis of neural reuse appears to be shared among researchers in the names of neural exploitation hypothesis [166], shared circuits model [167], neuronal recycling theory [168], and massive redeployment hypothesis [169]. These hypotheses share a common view that evolution favors reusing existing components over creating new circuits.

3) OPERATION PRINCIPLES OF CONCEPTUAL BEHAVIOR IN MACHINES

To avoid confusion, let us clarify first that what is discussed here as conceptual behavior is strictly for developmental machines. It neither implies human thinking, nor conforms to the philosophical treatment of concepts in cognitive science. Conceptual behavior is simply a type of behavior that takes place in machines' internal memory without sensorimotor involvement. In principle, conceptual behavior emerges because of the process of temporarily halting physical sensorimotor operations and transitioning to in-memory operations. The need for such a process arises when the sensorimotor system becomes inadequate or irrelevant for adequate actions. Instead of evaluating the values of inflowing sensor signals, the machine evaluates the memory content as if it were an environment to explore and exploit.

Machines thus far have developed previous behavior types by relying on physical interactions with the environment. Sensorimotor systems were necessary for such physical interactions to learn new skills to survive in the environment. Regardless of the learning algorithms used for previous behaviors, the working memory of a machine has the contents necessary to execute actions. Such contents are identifiably stored in an implicit, non-declarative form, and accessed indirectly by their locations in the storage. For the memory content to be operable, there must be a system that declares such memory content and makes them accessible for operations. Let us call it a system of concepts.

When sensorimotor learning is no longer enough to acquire new skills, machines must exploit their internal memory for new actions. The basic operation of conceptual behavior is thus to declare implicit, non-declarative content in the working memory as explicit objects that point to the locations where the content is stored in memory, and to computationally operate on them. Instead of using raw signal data from the sensorimotor system in real time, a system of concepts makes use of the declared objects in memory.

a: DEFINITION OF A SYSTEM OF CONCEPTS

System is a set of things to perform actions as part of a process. Process is a series of actions to proceed to the next end point. For our purpose, we define concepts simply as information conceived in memory. Information is an interpreted representation of signals and their derivatives. A system of concepts is thus a set of interpreted signals declared as addressable objects in part of a memory process to drive new behavior.

Let us draw an analogy in human terms. Information originates from events observed in the eye of the beholder and is stored in the memory of the beholder. In our eyes, the world is a continuum of events in the space-time universe. When we think about an event and describe it to someone else, we need a system of concepts as a tool to organize and communicate the information. Language is an example of the system of concepts for expression, deliberation, and communication.

For example, mathematics is a language of science. Computer programming language is a system of concepts about operating computers. We use natural language to communicate with others by using concepts to describe events, which may involve tangible objects, intangible thoughts, and situational and contextual elements. Every element in natural language is a concept, organized by the rules of lexicon, phonetics, morphology, syntax, and semantics [170]. According to Holyoak and Morrison, human behavior of thinking is defined as "*the systematic transformation of mental representations of knowledge to characterize actual or possible states of the world, often in service of goals*" [171]. For a general discussion and review of concepts in broader domains, see Murphy [172].

To reason is to construct and manipulate concepts. As new words can be constructed by manipulating words, new concepts can be built by manipulating concepts. In other words, a system of concepts is a recursively generative system of new concepts. By using such a system, developmental machines generate new concepts and actions when circumstantial value conditions are not satisfied by previously known actions.

b: IMPLICIT AND EXPLICIT CONCEPTS

Most computing machines today operate by executing instructions programmed in computer languages. Machines therefore have the capacity to process concepts declared as variables and functions that represent objects, events, properties, and relationships. For example, events can be represented in labels in the form e_i, $i \in [1, n]$ for n-dimensional events. The property of events such as the probability of events occurring can be labeled as $p(e_i)$. The relations between events such as the conditional probability e_i given e_j can be labeled as $p(e_i|e_j)$, where $i, j \in [1, n]$. In addition, actions can be labeled in the form a_k, $k \in [1, m]$ for m-dimensional actions. The relations between actions and events can be labeled as $p(e_i|a_k)$, where $i \in [1, n]$ and $k \in [1, m]$.

These labels do not mean anything by themselves until they are computationally combined and linked to other labels to produce outputs that influence actions. For example, suppose we have a conditional relationship by the following expression:

$$\begin{aligned} &If\ p(e_1) \simeq p(e_3)\ and\ p(e_1|e_3) > k, \\ &\qquad then\ e_1 \sim e_3 (e_1\ and\ e_3\ are\ associated) \end{aligned} \tag{34}$$

The meaning of the expression (34) can be colloquially explained as, if the probability of an event e_1 occurring is similar to the probability of another event e_3 occurring, and the conditional probability of an event e_1 given that e_3 has occurred is greater than a threshold value k, then the event e_1 is linked to the event e_3 in memory. An action rule can also be expressed with various labels such as:

$$\begin{aligned} &When\ e_i\ is\ desired,\ do\ a_x\ such\ that\ p(e_i|a_x) \\ &\qquad = max\{p(e_i|a_k)\}\quad for\ all\ k \end{aligned} \tag{35}$$

To execute expressions like (34) and (35) in a computer program, the labels or objects that represent variables and functions are explicitly declared so that their values can be accessed and processed to generate an output.

For developmental autonomous machines, the variables and functions necessary for reflexive and self-exploratory behavior are explicitly declared by human programmers in a program at the machine's birth. Once the machine is released in the environment to operate autonomously, events are observed, and features are extracted from sensor data in an intermediary process. How such data is processed depends on the signal-based learning used, but once the process is completed, the used data are typically thrown away or kept temporarily in a volatile storage.

This is a system of implicit concepts where signals are processed computationally without explicitly declaring or storing the signal attributes. Implicit concepts represent low-level quantitative information about observed signals. They allow fast and concrete execution of information for physical actions. According to this view, all sensorimotor systems operate as a system of implicit concepts.

To declare and store an implicit, non-declarative memory content is to assign an arbitrary label to an addressable object as a container that points to the memory content and signal attributes. Once the pointer and signal attributes become directly accessible as an addressable object, it can be manipulated in relation to the other objects. Because of the signal attributes inside an object, an arbitrary operation of objects could derive a novel object with new attributes.

This is a system of explicit concepts where stored data is processed computationally by creating addressable objects as containers of pointers to signals and their attributes. Because the objects are created only for signals that relate to meaningful events, they take less space than raw signal data and can be kept in a long-term memory.

An explicit concept is a high-level representation of what the machine has experienced in the past. The process can be slower than the implicit concept system but can generate new concepts for novel actions. This generative capability allows a machine to potentially overcome situational challenges that sensorimotor systems alone could not overcome directly.

Concepts are manipulated by operators. In computer programs, the most used operators are arithmetic, assignment, logical, and relational operators. Arithmetic operators perform numerical operations on variables, such as addition (+), subtraction (−), multiplication (∗), division (/), and modulo operation for remainder after division (%). Assignment operators assign values to variables. Logical operators determine logical relations between variables, such as AND (&&), OR (||), and NOT (!). Relational operators determine quantitative relations between operands, such as equal to (==), not equal to (!=), less than (<), greater than (>), less than or equal to (<=), and greater than or equal to (>=).

These operators must be extended to work with objects. For example, suppose there is an assignment operation on a program variable x with arithmetic operators, $x = 2 * y + z$. In a system of implicit concepts, the program variables y and z can take any values during the execution of a program. The value of x is directly computed from the values of y and z. However, the program cannot alter the arithmetic expression by itself.

In a system of explicit concepts, the pointers to the addresses in a memory where the values of y and z are stored are assigned with explicit objects Y and Z, for example. Thus, the objects Y and Z have unknown but fixed values. An object X is then associated with a symbolic expression $2 * Y + Z$. The value of X is unknown but can be computed from the attributes associated with Y and Z. Instead of dealing with quantitative values of Y and Z, the semantic value of X can be altered by substituting Y and Z with other objects, say A and B for example. The new symbolic expression $X' = 2 * A + B$ carries a completely different semantic value from the original expression. In this way, symbolically manipulating explicit concepts could generate new concepts.

c: SYMBOLIC EXECUTION, INDIRECTION, AND SEMANTIC POINTERS

Using pointers as explicit concepts is a practical solution to symbol emergence in machines, as the concept itself has been used and investigated in various domains. For example, symbolic execution is a widely used technique in automated software testing. The idea originates from the need for efficiently testing the integrity of software operations. In a conventional approach, a program is tested by a programmer who feeds sample data to observe the outputs to see if the program operates properly. It is a tedious task, and the success depends on the programmer's ability and the coverage of sample data. A better, more efficient approach is to automate the process by symbolically executing a program to simulate a large number of test cases.

Balzer [173] introduced the idea of extendable debugging and monitoring system (EXDAMS), in which the behavior of software programs is captured in a record to identify flaws in a program from what happened and how it happened in execution time. King [174] introduced an interactive debugging and testing system called EFFIGY, in which a software program is symbolically tested by supplying symbols instead of numerical values to a program written in PL/1. Boyer et al. [175] introduced a software testing system called SELECT, which generates test data to create a symbolic representation of the output variables for programs written in LISP. Clarke [176] introduced a system that generates test data and symbolically executes the path for programs written in ANSI Fortran. More recently, Cha et al. [177] introduced a system called MAYHEM, which automatically finds exploitable bugs in programs by reasoning about symbolic memory indices. Kuts [178] applied the MAYHEM approach in a symbolic execution tool to evaluate different memory modeling methods.

Indirection is another example of using pointers as explicit concepts. It is a hypothetical neural model of prefrontal cortex and basal ganglia that exhibits symbol-like processing. Kriete et al. [179] proposed a neural network architecture that learns to perform pointer-like operations called indirection. It is a multi-stage reasoning process where the first stage contains instructions on subsequent tasks. By using sentence encoding and decoding as an example, they demonstrated that the neural network achieved full combinatorial generalization by indirection and variable bindings.

Zhang et al. [180] advanced the idea of indirection as a benchmark to test the reasoning capabilities of artificial neural networks. The benchmark is called pointer value retrieval, set up as a two-stage reasoning process where the first task contains instructions for solving the second task. The benchmark tests an arbitrary network's ability to learn the indirection process by reading and interpreting the pointer's roles and values, and finally completing the tasks.

Blouw et al. [181] proposed a computational model of explicit concepts by using a symbol-like representation called semantic pointers. Similar to indirection, semantic pointers point to lower-level representations of perception networks that process sensor, lexical, and motor signals, while higher-level semantic information is retained by the compression process. Semantic pointers can be symbolically manipulated independently without reactivating lower-level compression networks. The framework is designed to integrate traditional connectionist and symbolic approaches.

4) EXAMPLE OF CONCEPTUAL BEHAVIOR EMERGENCE IN MACHINES

Consider the previous example where the mobile robot developed procedural behavior. Suppose in that example that the light emitter at the battery recharge station suddenly fails to emit the light signal. The machine does not know what happened, so it continues searching for the light signal but to no avail. This condition can be expressed as:

$$\Delta s_l^{ext} \leq 0 \text{ for all } s^{mot} \tag{36}$$

The expression (36) means that the observed change in the light signal is less than or equal to zero for all motor actions. The motivational drive for procedural behavior was to find motor actions that achieve a desired condition $\Delta s_l^{ext} > 0$. It was derived by the circumstantial value system based on the circumstances where the innate value system was not satisfied by the machine's reflexive behavior. The machine is now facing a new circumstance where the desired condition is no longer attainable because of the significant change in the environment due to an exogenous cause. However, the machine has no way of detecting the event occurred at the light emitter. It can only detect the presence or absence of light signals.

As the machine continues to drive its motors in search of a light signal, its internal battery level eventually reaches a critically low level after a while. This condition can be expressed as:

$$s_{bat}^{int} \leq v_{low}, \tag{37}$$

where v_{low} is a threshold value that indicates a critically low battery level. At this point, the machine is in a situation where two value systems, innate and circumstantial, fail to be satisfied by reflexive or procedural behavior. Furthermore, the critically low energy level prevents the machine from executing motor actions. In other words, the skills acquired from the previous stages of development are no longer adequate to maintain the machine's full autonomy.

The expressions (36) and (37) are computable conditions and thus they can be used to trigger a system, the conceptual value system. In principle, the role of the conceptual value system is the same as the ones for innate and circumstantial value systems. The purpose and function of a value system is to establish a new desirable condition as a goal and to drive new actions to achieve the goal. The difference among the three systems is the means with which the actions are carried out. For the conceptual value system, the available resource is its internal memory. Under the current circumstance, a desirable condition is to reverse the expression (37),

$$s_{bat}^{int} > v_{low} \tag{38}$$

The conceptual value system then initiates a series of memory actions by working with two other systems, the conceptual storage and registry. All pointers to the working memory are explicitly labeled and linked to the existing variables and functions used in the previous development stages. For example, suppose that the machine's working memory has a record of actuator signals and their post-execution sensor signals in the past. Such a record was used in physical trial-and-error behavior in the previous stage. Except the identifiable variables such as sensor and actuator identifiers, actual signal values in the record are only linked to them without explicitly being declared. By declaring objects as a container that points to these signal attributes, they can be symbolically manipulated to generate new objects and expressions.

For example, suppose that the symbolic operation searches for objects that link to non-motor actuator events. Suppose that the search retrieves an object that points to a speaker action event that caused an increase in the battery signal. Such an event may have occurred early in the past when a random speaker action as part of self-exploratory behavior caused a human to plug the battery charger. The event was captured and kept as an object in memory. The robot can now purposefully execute a speaker action.

In a different scenario, suppose that the robot has a built-in compass but never used it in procedural behavior. The historical data would show correlations between light signals and compass signals. In-memory symbolic operations may uncover such correlations from the past data and lead to a set of new actions. These scenarios are simplistic and rather convenient, but nonetheless, they explain the logical workflow of causality, mechanism, and potential of conceptual behavior. Fig. 21 illustrates the concept of a machine that exhibits conceptual behavior.

G. CAUSALITY AND MECHANISM OF SOCIAL BEHAVIOR

1) OVERVIEW

A purely in-memory operation has its limits. The potential search coverage inside a single machine's memory is small compared to the vast universe outside the machine. With the ability to process high-level representations of observed signals, machines can extend symbolic operations to newly observed signals from the environment. Observational association as evidenced in episodic behavior from the earlier stage can significantly expand possible associations among what the machine has seen before, and it is seeing now. In this way, developmental machines can increase the size of their conceptual storage and registry. This in turn enables the machines to generate novel actions as they explore and exploit the resources they encounter in the environment.

One of the potential resources they may encounter is another machine with similar capabilities. Each machine acts on its own value systems and executes purposive acts to achieve their own desired outcomes. They have the means to exploit the signals emitted from the others. Initially,

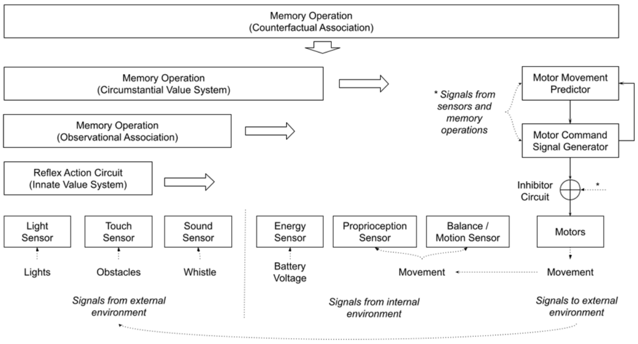

FIGURE 21. Illustration of operation process for conceptual behavior.

such signals are unknown and neutral. With the ability to observe and associate unknown signals with episodic events, the signals may eventually be interpreted as appetitive or aversive. Their natural tendency to approach appetitive signals causes the machines to approach toward each other. If some signals are deemed aversive, their inherent tendency to avoid aversive signals causes them to move away from each other.

In the eyes of an observer, such interactive behavior may appear social, but it is only episodic behavior responding to the observed signals. This could change to purposive acts if the machine recognizes any consequential change in transmitted signals due to its actions. If the purposive acts by one machine is reciprocated by another machine, the two machines' purposive acts are aimed at each other, resulting in a communication between the two machines.

Communication is a process of sending and receiving signals between a transmitter and a receiver. For the process to take place, there must be more than one machine operating and serving as a transmitter and a receiver in an environment. The receiver detects a transmitted signal and interprets according to the internal value system. The signal may not be transmitted intentionally by the transmitter. It could be a one-way observation of the transmitter's behavior by the receiver. The observed signal could trigger a purposive act if it is deemed meaningful. Such an act by the receiver could then be observed by the transmitter. It interprets the observed signal according to the internal value system and may also trigger a purposive act. Such an act by the transmitter could be observed again by the receiver and results in a subsequent action. This scenario could repeat itself as an iterative process of sending and receiving signals by way of purposive acts between the two parties.

This mutual act of signal exchange is built on two assumptions. First, a transmitter and a receiver share a common medium where signals can be exchanged. Second, the transmitted signal must relate to the internal value systems of the receiver. For example, suppose that a transmitter sends a light signal in different colors like red, green, and blue. For this to be a medium of communication, the receiver must have a sensor to detect the different colors of light. Such a signal may be a neutral signal initially, but once it relates to the value system by way of observational and interventional associations, the signal becomes meaningful and subsequently results in a purposive act. The process is reversed when the transmitter becomes a receiver by detecting a signal from the receiver's purposive act and interprets it according to its value systems.

2) EMERGENCE OF LANGUAGE

The receiver of a transmitted signal must be able to interpret the meaning in terms of its own value system. Such translation is possible even without a priori agreement if there is a common reference for the transmitter and receiver to relate the meanings to their respective value systems. A popular word guessing game of Charade is an example of how this can be done. While a player of the game sends visual signals by body movements, the other players try to interpret the movements to guess what the message is. The game is fun but challenging to establish a reference to what the visual cues represent. When a group of machines interact with each other in a way new references are established by understanding each other's signals, such signals serve as a "language" among them. In baseball or football, using body movements to send play calls is an example of such.

Communication is a process, enabled by a system. Language is a system of communication between the transmitter and receiver on how signals represent meanings. There is no cardinal rule per se on how exactly this must be done. In natural language for humans, some weigh high on structures such as grammar and vocabularies, others may weigh more on tonal qualities than structures. This applies to other animals as well. The fact that so many expressions and dialects exist and differ among various communities is evidence that anything goes as far as the rule of language is concerned [182].

3) EXAMPLE OF SOCIAL BEHAVIOR EMERGENCE IN MACHINES

Suppose that two developmental autonomous machines M_A and M_B are operating in proximity within the environment. Let us assume that they have been exploring and exploiting the environment for a long time and that they have already achieved a high level of conceptual behavior. Initially, signals regarding the fellow machine are neutral because they are neither appetitive nor aversive at this point.

Let us now imagine a situation where M_A's obstacle sensor is not working and is about to collide with an object. M_B as an observer recognizes the two objects about to collide. Because of its observational association regarding the object events, the episodic response triggers an alarm sound. M_A senses this alarm sound but such signal is yet to be conditioned and as a result, M_A collides with an object. After repeated encounters of object collision and preceding alarm sound, M_A finally associates the sound alarm with collision events. The next time M_A senses the alarm sound from M_B, it executes an avoidance response.

Although primitive, the scenario could extend to more complex situations like the game of *Hot and Cold*, where M_B guides M_A to locate the recharge station by sending signal patterns of "Hot" and "Cold" as M_A moves towards or away from the station. It is conceivable that more complex social behavior could emerge from extended interactions between the two machines if they possess the capabilities of behavior development.

In principle, it is reasonable to expect machines to eventually develop behaviors that may appear to form a community. At an early social stage, a particular machine, M_{alpha}, finds a food source more often than the others in a group. Others begin to follow M_{alpha}. Then new changes occur in the environment or the internal systems. Each machine is forced to go through a conceptual behavior stage to adapt to the change. Another machine, M_{alpha2}, finds a solution to overcome the change and send off appetitive signals to the others. Others begin to follow M_{alpha2}.

This process repeats for a while and eventually, a few machines emerge as leaders that attract more members than the others based on their strength to adapt to the changes. Machines begin to separate into groups, each led by a leader. Within each group, machines continue to exchange signals in response to the nature of their environment and changes that occur. A new leader may emerge depending on the circumstances.

This process continues for a while and eventually, some groups may settle in environments that pose less threats and require less adaptation than the other environments. Clusters of machines settling in various locations emerge. Their behavior collectively exhibits different characteristics to the others, reflecting their own adaptation to their native environment. In the eyes of an observer, this phenomenon may appear as if machines formed a community.

H. SUMMARY OF THE CHAPTER

This chapter described the causality and mechanisms of behavior development. In principle, behavior emerges when a machine detects a circumstance that drives the machine to act. Motivation is merely a reason that causes actions, which are then observed as behavior. Reasons vary depending on the circumstances; therefore, motivation is circumstantial and so is behavior. For new behavior to emerge, there must be a system that recognizes a certain circumstance that gives the machine a reason to act. Such a system is called a value system.

A value system is a mechanism that detects inherently meaningful signals from the environment. Signals are observed by a sensorimotor system as the machine explores its environment. The meaningfulness of observed signals is initially defined as the innate value system, programmed by human designers. Reflexive behavior is a type of behavior directly emerging from the innate value system.

1) LEVEL 1: MACHINE EXHIBITS SELF-EXPLORATORY AND REFLEXIVE BEHAVIOR

A developmental autonomous machine of Level 1 exhibits pre-programmed actions of basic motor learning and reflexive behavior. Reflexive behavior is an innate, involuntary behavior in response to a stimulus. A machine begins its life with self-exploratory motor learning behavior, exhibiting random movement with innate behavior patterns of aversive and appetitive responses. This initial stage is completely under control by human designers, who define the machine's purpose and agency by programming the innate value system and subsequent learning mechanisms. The system components required for machines to exhibit self-exploratory and reflexive behavior are:

1) sensors to detect aversive and appetitive signals,
2) actuators to execute predefined movements for aversive and appetitive actions,
3) sensors to detect internal energy status, motion, and proprioception,
4) a process component to convert sensor signals to actuator control signals,
5) a power supply for sensing, acting, and processing components,
6) a chassis to hold the above components as a single operation body, and
7) a recharge mechanism for the power supply (optional to play the game of survival).

The process component provides three major functions: innate value system, self-exploratory mechanism, and motor learning mechanism. The innate value system triggers predefined actuator command signals in response to predefined sensor signal conditions. The self-exploratory mechanism sends basic actuator command signals randomly. The motor learning mechanism builds a model between the basic actuator command signals and the internal sensor signals.

2) LEVEL 2: MACHINE DEVELOPS EPISODIC BEHAVIOR

The meaningfulness of signals dynamically evolves as the machine encounters new events in the environment. Episodic behavior is a type of behavior emerging from the process of associating neutral signals with the known meaningful signals. This passive observational association is a result of the machine's memory system computing the similarities and relatedness of observed signals in a manner similar to unsupervised learning. A developmental autonomous machine of Level 2 autonomously develops episodic behavior, which is an acquired, involuntary behavior, to passively explore and exploit the environment. In order to exhibit episodic behavior, the following components are necessary:

1) the system of reflexive and self-exploratory behaviors,
2) sensors to detect neutral signals,
3) memory storage, and
4) a processing mechanism for observational associative learning.

3) LEVEL 3: MACHINE DEVELOPS PROCEDURAL BEHAVIOR

A developmental autonomous machine of Level 3 autonomously develops procedural behavior, which is an acquired, voluntary behavior, as purposive acts of motor skill learning. Procedural behavior is a type of behavior emerging from the process of associating its own actions with the consequential changes in observed signals. This active interventional association is a result of the machine's memory system computing the difference between expected and observed values in signals. Due to the lack of precision in early-stage motor skills, actions derived directly from the innate value system may not result as designed. Under such circumstances, the value system evolves to a circumstantial value mechanism that dynamically defines desirable conditions, which results in purposive acts in a manner similar to trial-and-error behavior in reinforcement learning and supervised learning. To exhibit procedural behavior, the machine needs the system that supports episodic behavior, a motor skill learning system, and a mechanism to promote a circumstantial value system.

4) LEVEL 4: MACHINE DEVELOPS AUTONOMIC BEHAVIOR

A developmental autonomous machine of Level 4 autonomously develops autonomic behavior, which is an acquired, voluntary behavior in a self-governing manner. As the machine refines its motor skills to achieve desired conditions, certain actions are frequently repeated over and over. New skill learning is no longer necessary for such repeated actions. Autonomic behavior is a type of behavior emerging from habitual association between the initial trigger act and the series of subsequent actions. By eliminating intermediary processes of feedback signal evaluations, the entire operation becomes a sequence of feedforward actions, in a manner similar to feedforward neural networks. The process of habitual association is a result of the machine's memory system organizing and compacting the memory content for fast, energy-efficient execution.

5) LEVEL 5: MACHINE DEVELOPS CONCEPTUAL BEHAVIOR

A developmental autonomous machine of Level 5 autonomously develops conceptual behavior, which is an acquired, voluntary behavior, performed on high-level representations of the prior experience in memory. The machine may achieve its full autonomy solely from signal-based learning and the resulting episodic, procedural, and autonomic behavior, yet the meaningfulness of signals still evolves due possibly to significant environmental changes. Under certain conditions, previously learned skills may become inadequate to achieve desired conditions, or signals from the past may no longer be available for re-learning. Conceptual behavior is a type of behavior emerging from the process of using pointers to meaningful signals and their attributes as addressable objects and by symbolically manipulating objects to create novel actions. Conceptual behavior needs a storage and registry to manage explicit concepts as well as a conceptual value system that appraises the resulting symbolic expressions.

6) LEVEL 6: MACHINE DEVELOPS SOCIAL BEHAVIOR

A developmental autonomous machine of Level 6 autonomously develops social behavior, which is an acquired, voluntary behavior that interacts with other entities in the environment. As the machine refines its ability to conceptualize the stored data as explicit concepts, it begins to interpret observed signals based on similarities and relatedness with the known explicit concepts. Social behavior is a type of behavior emerging from the process of exchanging signals when two or more machines encounter in the same environment. With the ability to observe and associate unknown signals with explicit concepts, the exchanged signals can be exploited as a resource. This mutual act of signal exchange is built on the assumptions that the transmitters and receivers share a common medium among the machines, and the transmitted signals must relate to the internal value systems of the receivers.

IV. DISCUSSION

A. CRITICISM AND FUTURE WORK

The previous chapters laid out a theoretical framework of developmental autonomous behavior, providing the logical and plausible explanation of why and how new behavior can emerge in machines. It is an ambitious attempt on challenging topics, but by delineating the core subject in strict terms of machines without mixing intangible human quality, the proposed framework offers a clear, precise treatment of what it means for machines to be developmental. While it may have alleviated potential confusion and philosophical debates to some degree, it is not free from issues and shortcomings. Admittedly there is a lot of work to be done at this early stage of research.

First, the notion of autonomy needs careful examination. Since a wide variety of machines exist, the meaning of autonomy varies significantly depending on phylogenetic, ecological, and anthropogenic perspectives as mentioned in Sections I-A and II-A. Autonomy not only defines machines' agency but also indicates the risk and benefit of machines operating autonomously. The degree of autonomy must be declared in such a way that creators and users can confidently build and use such machines. This issue becomes urgently critical as the machine becomes a learning system. On one hand, learning is desirable because it is not reasonable to blindly trust a pre-programmed apparatus to safely operate autonomously under unforeseen circumstances. On the other hand, blindly trusting a self-learning machine that proactively changes its behavior is not reasonable either.

For this reason, this article addressed the question of what it means for a machine to be developmental as a learning system. Autonomy was framed as a survival game for clarity, simplicity, and generality to identify and explain the basic elements of developmental behavior. In this way, the notion of autonomy is delineated with generality. However, this is just a first step. Formalizing the characterization of autonomous systems is an important future work.

Kuguke et al. [183] formalizes the taxonomy of autonomous systems by distinguishing between systems that learn and ones that do not learn. In this framework, intermittent and eventually autonomous levels are introduced in between non-autonomous and fully autonomous levels, comprising four levels of autonomy. Building on this formal analysis is one possible direction.

The next criticism of this article is utility. This article did not elaborate on specific techniques or algorithms. Instead, the article kept its focus on principles and theoretical explanations on causality and mechanisms of behavior development for generality. There are literally millions of different kinds of machines at work for a wide variety of purposes and environments. Techniques and algorithms become pertinent when their utilities are defined and specified in terms of the machine's purpose and operating environment. Solving a toy problem in a lab environment has its place and value, but again the utility, what kind of a real-world machine and application for which the algorithm is intended, needs to be declared.

Even though there is no universal one-size-fits-all algorithm for developmental autonomous machines, this article illustrated a few practical methods, including probabilistic and unsupervised networks for observational association, reinforcement learning for interventional association, and symbolic execution for counterfactual association. These are just a tip of the iceberg. There exists a broad spectrum of techniques in control systems and machine learning that could serve as operational engines for developmental machines. For this reason, a rich and exciting area of future work is to explore opportunities in specific domains where developmental capabilities provide useful values. With specific utilities in mind, discussions on algorithms and techniques become meaningful.

With that stated however, formal analyses of cognitive systems and machine learning techniques must take place in relation to developmental autonomous behavior. The challenge is their diversity. Their forms and complexities vary significantly because of their underlying inspirations, such as artificial neural dynamics, state machines, symbolics, and probabilistic reasoning. To formally analyze the overall system, how the specific cognitive or learning system exactly supports the processes of symbol emergence and subsequent behavior transition from physical to conceptual and social behavior must be explained logically.

Values and memory systems represent the key fundamental mechanisms for behavior development. Something must drive machines to behave. Behavior does not develop without learning. Learning does not occur without memory. While machine learning techniques address a variety of technical challenges and some have been applied successfully, these tools have not been unified or categorized in a way that provides insights to how they function as value and memory systems.

Noori [184] described a working memory system for high-order cognitive tasks with symbolic representations. The system consists of a serial symbolic storage, state registry, and symbolic schema learning system. Blouw et al. [181] proposed a computational model of concepts as semantic pointers, which integrates traditional connectionist and symbolic approaches. Building on these ideas of symbolic operations in working memory is one possible direction.

Following on the line of utility discussion, scalability is another topic of importance. Real-world machines for practical purposes widely vary in their scales from single sensorimotor systems to high-dimensional, multi-sensor-actuator systems. The scales and complexity of physical systems vary and so does the data processing capability. In addition to dimensionality, time scales of skill development must also be understood. How can we build a scalable developmental machine for practical purposes?

How about morphology - physical changes in machines during their lifetime? Adding, removing, or modifying parts of the body can alter the behavior of a machine. It is conceivable that such physical changes occur due to exogenous and endogenous causes. Humans can add, remove, or modify sensors, actuators, batteries, chassis, processors, or any other parts. The environmental elements can also cause functional change or failure due to temperature, pressure, air movement, or other physical changes in the space. These are exogenous causes. It is also possible for machines to alter their physical configuration on their own. For example, an assembly robot with arms and hands can intentionally or accidentally add, remove, or alter parts. This is an endogenous cause. Can a machine fix itself or improve its functions by dealing with exogenous and endogenous causes of morphology? This is one possible question to be addressed in future works.

Last but not least, ethics. How do we maintain moral principles in machines that govern their behavior? We must be able to implement, not just ideological principles, but actual mechanisms that warrant safe operations of machines. The question of ethics also applies to human developers and users as well because they are the ones with the ability to control the behavior of the machines they create or deploy in the real-world environment. Let us elaborate on this topic in the next section.

B. ETHICS

1) WHY ETHICS

As we anticipate an environment in coexistence with machines that make their own decisions to act, we ought to be concerned with the uncertainty of how our social norms are maintained. For this reason, ethics is an important area of future work. There have been discussions in the past on ideological principles on machine ethics. The most notable example is Asimov's Three Laws of Robotics [185]:

- *First Law: A robot may not injure a human being or, through inaction, allow a human being to come to harm.*
- *Second Law: A robot must obey the orders given by human beings except where such orders would conflict with the First Law.*
- *Third Law: A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.*

These laws are presumably designed to protect humans but placing the task of protecting humans at the core of the machine's responsibility is not practical as a fundamental principle of machine ethics. For instance, the second part of the first law, which requires robots to attentively protect humans, is not a practical design requirement for robots built for simple tasks. The second law, which demands robots to be aware of the consequences of commands from humans, is an overbearing and impractical design requirement. The third law, which requires robots to actively defend themselves from threats, poses itself as the cause of social disorder when robots fail to recognize the conflicts. The only pertinent principle in Asimov's laws is the first part of the first law: a robot may not injure a human being.

The three laws of robotics also fail to address the human aspects of ethical guidelines. Machines are built and used by humans. Historically speaking, certain machines have been built and used or misused for the purpose of harming humans, animals, and properties. No law will help if there are no moral principles imposed on the creators and users of machines. It is thus critically important to deepen our understanding of ethical principles for both machines and humans.

2) HOW HUMANS FAIL IN ETHICS

Before we address machines, let us first reflect on how we humans exercise our moral principles. We are emotional and social creatures by nature, sharing the environment with others. Self-referential emotions such as pride, shame, guilt, jealousy, and empathy play a major role in how we act [186], [187], [188]. Human societies create a code of ethics to maintain the behavioral standards. A good social balance relies on collective efforts of the members, controlling their emotion and behavior in accordance with their code of ethics.

We all know that we shall not harm others. We all know that we shall not steal or harm the properties of others. In real life however, humans often violate these moral principles, especially when conflicting objectives occur. While prioritizing objectives is a key problem-solving skill for humans, circumstances strongly dictate how we behave, failing in ethical conduct.

If we allow machines to prioritize conflicting objectives, then they may fail in ethics by overriding the ethical principles just as humans do. The stories from "*I, Robot*" [185] is a good reminder of what could happen if machines were allowed to reason to prioritize. Unless there is a secure, reliable mechanism to prevent machines from crossing the ethical lines, machines will make mistakes just as humans do. The only mechanism that guarantees machines not to disobey the ethical code is to install such a code in the machine where it cannot be modified or overridden.

3) IMPLEMENTABLE ETHICAL CODES

There are three areas where ethical principles can be implemented as default mechanisms that the machines cannot override by themselves: value systems, learning mechanisms, and physical sensorimotor functions. An innate value system drives reflexive behavior, which is executed without reasoning from its birth to the end. Just like your hands twitch away from a hot stove, machines must instinctively stay away from crossing ethical lines. A circumstantial value system may under certain situations override reflexive behavior, but that can be controlled by learning mechanisms. Limiting the physical capacity of sensors and actuators is another element for humans to keep a machine under control.

In addition to the elements to be implemented in machines, humans must follow certain ethical principles. The first area is to make sure that producers of machines implement the ethical codes properly in their products. Second, the operating environment and proper usage of machines must be properly communicated to the users. Alphanumeric codes commonly used in industrial standards can be an effective way to communicate the level of risks associated with the machine. For example, the *IPxx* designation for ingress properties of electrical devices uses numerical values, e.g., *IP68*, to indicate the level of protection against dust and water [189]. With such a recognizable designation and display, manufacturers and users can be informed of necessary risks and guidelines for implementation, operating environment, and proper use.

4) COMMUNICATION AS A SAFETY TOOL

It is also conceivable to implement a system that allows humans to send aversive signals at a certain moment so that the machine interprets and associates the signal with its innate value system to avoid. This is in a way how we train our pet dogs to follow our social rules. It works because dogs have associative learning capabilities. Machines could do the same in principle. What we need to develop is a mechanism to communicate with machines to convey our signals to prevent accidents from occurring.

Inspecting the machine's internal integrity is an important safety measure. This can be accomplished by communicating with machines. With proper mechanisms, it is possible for machines to expose their internal concepts to humans, like a "explain yourself" command. Because humans can see and hear, there are two options: visual/audio language and natural language. Visual/audio language can be any recognizable visual/audio cues that humans can see or hear and understand. For example, LED flashing in patterns, displaying certain shapes or motions, and beep patterns like the Morse code or machine signs in sci-fi movies. Natural language is an obvious choice of communication for humans, but it requires a proper translation mechanism for machines. Exposing machines' internal concepts and their relationships in a way that humans can understand is an important consideration, not only for convenience but for building trust between machines and humans.

C. EMBODIEDNESS

Embodiedness is a much-discussed topic in cognitive science [83], [84], [190], [191], [192], [193]. An embodied being implies cognition as a product of interactions between the body and environment. Indeed, in developmental autonomous behavior, conceptual behavior cannot emerge without signals obtained from the experience of episodic, procedural, and autonomic behaviors. If the machine does not have a means to interact with its environment, conceptual behavior may not exist. Does this mean that the environment must be physical?

Consider a machine entirely constructed as a computer program with no physical sensor or actuator. Suppose that the machine lives in a stock market environment with many computer systems connected by data networks. In this case, the machine's sensor is a computer program that reads real-time stock values and other miscellaneous data from external sources. The machine's actuator is another computer program that executes a series of transactions to buy, sell, or hold tradable stocks. Let us suppose that the machine's innate value system is planted in such a way that appetitive and adverse signals are derived from the stock values gained or lost from a series of transactions. The machine is programmed in such a way that it learns to exhibit episodic, procedural, and autonomic behaviors, which yield signals and signs from the records of stimulus-response transactions, as well as world events that the machine can associate with its experience in the past. With a memory full of signals, signs, and symbols derived from its own experience, is it possible that conceptual behavior emerges in this machine? Is this machine embodied?

From this example, being embodied may not strictly imply physical entities and environment. What was conjectured in 1955 by the original AI innovators [40] cannot be rejected simply from the embodiedness argument. By Turing's classification of machinery [32], the above example of a stock trading machine deals with discrete controlling and active machinery, while physical robots with cognitive capabilities deal with continuous controlling and active machinery. What is more important than the notion of embodiedness is that concepts do not arise in vacuum. Concepts emerge from signals in a transformational process that one experiences. Conceptual behavior is situated in context.

D. SYSTEM ARCHITECTURE

1) STRUCTURE

To facilitate consistent, comparable, and reproducible experiments and analyses for developmental autonomous machines, a common general architecture would be useful. Such an architecture is preferably simple, scalable, transparent, and easily implementable in the contemporary computing environment. It is also important for an architecture to address the issues highlighted earlier in this chapter, such as scalability, morphology, integration, and ethics.

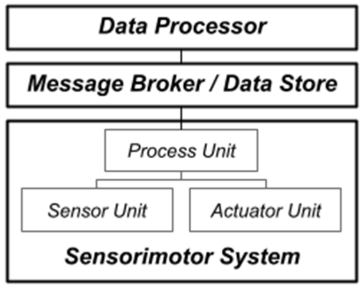

FIGURE 22. General architecture with a single sensorimotor system.

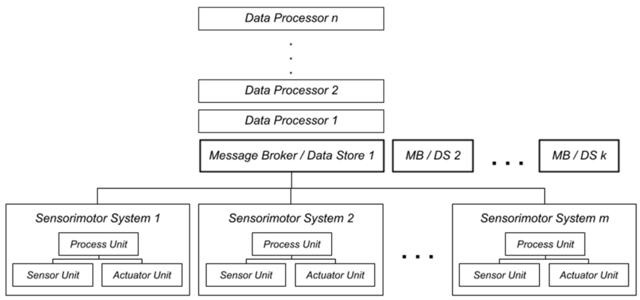

FIGURE 23. General architecture with multiple subsystems.

Fig. 22 shows a block diagram of a system architecture with a single sensorimotor system. Fig. 23 shows a block diagram of a system having multiple subsystems. Both structures assume a familiar 3-tier model of the 3T Intelligent Control Architecture for robots [194], [195]. The subsystems and their relationships are designed specifically to support the principles of developmental autonomous behavior. Three major subsystems are included: sensorimotor system, message broker / data store, and data processor.

The sensorimotor system interacts with the environment via sensor, actuator, and process units. Each sensorimotor system assumes a general structure that by itself can represent a wide array of real-world machines. For example, for continuous systems, sensors and actuators can be an electro-mechanical device that interfaces with the physical environment. For discrete systems, sensors and actuators can be a software program that interfaces with external information systems. The sensorimotor system can stand alone as an independently operating machine. However, it is not by itself a learning system. The behavior does not change unless specific instructions or execution programs are sent from the data processor via the message broker.

The message broker receives data from the sensorimotor system and stores them as a data store. The data processors perform various associative learning processes by using data stored in the message broker. The outcome of associative learning can be packaged as an instruction set, which can be uploaded to sensorimotor systems via the message broker or other means of data transport. With the updated instruction sets, the behavior of sensorimotor systems can be altered. In this sense, the message broker can be seen as a gatekeeper between physical and virtual environments. These three subsystems can be linked by wires, wireless, or a combination of both, so that all three need not be in the same location together.

2) PRINCIPLES OF THE PROPOSED ARCHITECTURE

The 3T robot architecture is particularly suitable for developmental autonomous machines for a few good reasons. First, it represents the architecture of complexity by the concept of decomposable hierarchic system, proposed by Herbert Simon [1]. The idea is that complex systems are hierarchical in nature, and that they can be approximated by nearly decomposable subsystems. The decomposable hierarchic system facilitates comprehensibility because of the visibility that the system provides to its subsystems and its behavior. Comprehensibility is an important feature for the purpose of research. Because of the decomposable hierarchy, it is possible to analyze where and how the learning is taking place in reflection to the machines' environment and experience.

The proposed architecture reflects contemporary computing environments and technologies. With the increased computing power and decreased costs, it has become possible to construct a complex high-speed high-power system with small, inexpensive off-the-shelf components. In addition, with industry-standard protocols such as message queue telemetry transport (MQTT) and internet protocol (IP), large-scale remotely located computing systems can be easily and inexpensively connected to local systems. In the proposed architecture, all these components are considered subsystems that fit within the decomposable hierarchy.

In Section I-A, important questions were posed without answers: "*If we were to build a machine that can exhibit all these different kinds of behaviors, is the structure of Fig. 1 sufficient? If so, what does the processing component look like and how does it work to support developmental behavior?*" The proposed architecture provides practical answers to these questions. By decomposing the processing component into three subsystems and by specifying their roles and relationships, it is shown that the same structure can be used to build a developmental autonomous machine, as illustrated in Fig. 24.

With the proposed system architecture, many of the topics discussed earlier in Section IV-A can be empirically studied.

Scalability: The whole system can scale up or down by the number of sensorimotor systems, message brokers / data stores, and data processors. Having multiple subsystems provides added complexity and functionality. For example, via the message broker linkage, sensorimotor systems as message clients can exchange messages (data) to each other for collaborative work. Sensorimotor 1 can be a mobility system with motors, while Sensorimotor 2 can be an assembly system with arms and hands, for example. Each system can operate independently to perform its own tasks, but by exchanging messages via the message broker, they can coordinate and collaborate to perform more complex tasks together. All the messages and data exchanged are stored in the message broker. Having multiple data processors and message brokers allows distributed processing either locally or remotely. With multiple data processors in a cluster configuration, the computation can be made faster by distributing the load among computing nodes.

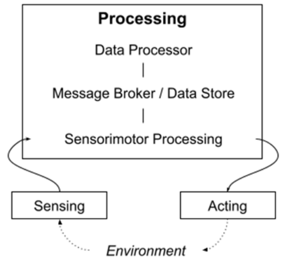

FIGURE 24. Decomposable hierarchy for processing working with sensing and acting.

Morphology: The impact of morphology can be studied by altering parts of the decomposable subsystems. Having multiple sensorimotor systems as decomposable subsystems, adding, removing, or modifying parts of the body is straightforward and transparent with the architecture. For example, adding a new sensor to one of the sensorimotor systems only impacts the subsystem. By freezing and archiving the memory, subsequent changes due to associative learning can be studied repeatedly. Having two identical sensorimotor systems allows even more flexible comparative studies between the two situations.

Integration with cognitive systems: In essence, all cognitive systems are data processors. The general architecture for developmental autonomous behavior is algorithm-agnostic in the sense that, whether the cognitive system is based on artificial neural networks, symbolic processing, or probabilistic reasoning, they would all be implemented in the data processor subsystem. Having multiple data processors becomes particularly useful. For example, by implementing different algorithms in cluster nodes, their behavior can be compared without altering the other subsystems. For example, different policy optimization schemes for reinforcement learning can be implemented in nodes and switching the nodes to observe the performance difference. Multiple levels of associative learning schemes can also be implemented in nodes so that the whole cluster as a unit can develop new behavior from episodic to procedural to conceptual.

Utility: Because of the nature of decomposable subsystems, the general architecture can handle both continuous and discrete systems, allowing a variety of application domains. The sensorimotor system is simply a subsystem, so it does not preclude a physical or non-physical environment.

Ethics: The general architecture allows us to frame the question of ethics as where in the system do we implement

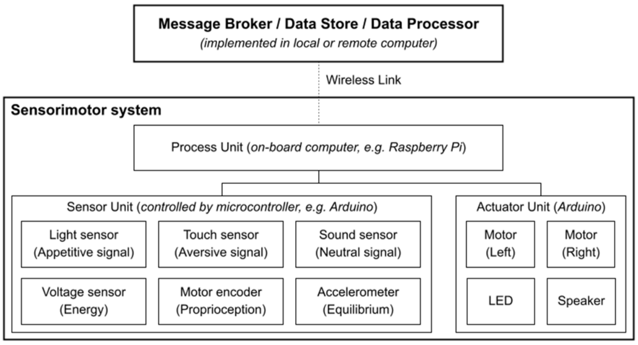

FIGURE 25. Block diagram of a minimally configured developmental autonomous machine.

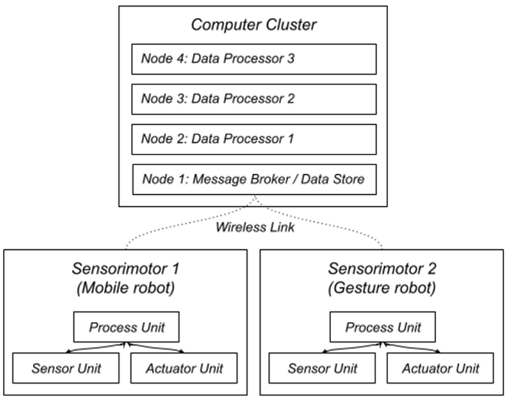

FIGURE 26. Component diagram of the multi-unit developmental autonomous machine.

the mechanism that enforces the ethical guidelines for the machine to obey. Each subsystem brings practical benefits and shortcomings, and the pragmatic analyses of implementational aspects of machine ethics is feasible because of the architecture.

3) A MINIMALLY CONFIGURED MACHINE

Throughout this article, a simple example has been used to demonstrate and exemplify the causality and mechanisms of behavior development. The system is minimal with only a few sensors and actuators, yet sufficiently provides all components necessary to exhibit developmental behavior. Such a system can be built with off-the-shelf components, from sensors and actuators to microcontrollers and microprocessors. The processing capacity of today's microcontrollers and processors is more than adequate to support the basic signal processing and memory operations.

Fig. 25 shows an example of the minimally configured system that demonstrates developmental behavior. The machine has six sensors and four actuators: light, touch, sound, energy, accelerometer, encoders, two motors, LED, and speaker. This setup is considered minimally configured because the number of sensors, actuators, and processors required for developmental behavior is at minimum. The light sensor detects light as an appetitive signal, the touch sensor detects bumps with objects as an aversive signal, and the sound sensor detects sound as a neutral signal. In addition, the robot has a voltage sensor to monitor the battery level, a 3-axis accelerometer to detect balance and motion, and motor encoders to detect motor rotations as a proprioceptor. The mobile robot moves by two motored wheels. It has an ability to emit a light signal by LED and a sound signal by a speaker. These output signals can be used to communicate with humans and other robots. The message broker, data store, and data processor are located remotely and linked to the robot wirelessly.

In case of a need for increased capacities, multiple processors can be stacked together as a computer cluster, as shown in Fig. 26. With this set up, the machine can be easily scaled up or down.

V. CONCLUSION

Evolution is a modification to ontogeny [196]. Until Darwin's Origin of Species was published in 1859, the term "evolution" meant the development of an individual organism, not life in general [197]. Machines are not a living organism but in the hands of human creativity and dexterity, they evolved from simple automatic regulators to complex autonomous robots. With their behavior precisely programmed by humans, machines are expected to execute the given instructions obediently. This notion of predictable behavior is changing because machines are increasingly becoming a learning system. Anticipating the future evolution, this article explored the fundamental question of what it means for machines to be developmental, and why and how a machine could acquire new behavior on its own.

This article explained that in principle, it would be possible to construct a developmental autonomous machine. By establishing the theoretical foundation, this article provided a general and practical framework to analyze and synthesize machines that exhibit developmental behavior. As machines increasingly become autonomous, and as we become more and more dependent on them, it is critically important that we deepen our understanding of their behavior and implications. Machines will use their frames, sensors, actuators, memory, and processing powers to develop new skills. Machines will acquire new skills by exploring and exploiting the environment that humans place them in. Underlying all this, value systems play the key role that drives behavior development. Ultimately, they are all in the hands of humans because we are the ones who create machines. We must fully understand the principles and implications.

ACKNOWLEDGMENT

The author would like to thank the editor and reviewers for taking the necessary time and effort to review the manuscript and for their valuable comments.

REFERENCES

[1] H. Simon, *The Sciences of the Artificial*. Cambridge, MA, USA: MIT Press, 1969, p. 53.

[2] D. T. Thomsen and L. Joy, "The stray," in *Westworld, Season 1, Episode 3*. New York, NY, USA: HBO Entertainment, Oct. 2016.

[3] T. Froese, N. Virgo, and E. Izquierdo, "Autonomy: A review and a reappraisal," in *Advances in Artificial Life* (Lecture Notes in Computer Science), vol. 4648, F. Almeida e Costa, L. M. Rocha, E. Costa, I. Harvey, and A. Coutinho, Eds. Berlin, Germany: Springer, 2007, pp. 455–464.
[4] C. P. Ezenkwu and A. Starkey, "Machine autonomy: Definition, approaches, challenges and research gaps," in *Intelligent Computing* (Advances in Intelligent Systems and Computing), vol. 997, K. Arai, R. Bhatia, and S. Kapoor, Eds. Berlin, Germany: Springer, 2019.
[5] Merriam-Webster. (Oct. 1, 2022). *Autonomy*. [Online]. Available: https://www.merriam-webster.com/dictionary/autonomy
[6] H. M. Huang, "Autonomy levels for unmanned systems (ALFUS) framework volume I: Terminology version 2.0," Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Tech. Rep. NISTSP 1011-I-2.0, 2008.
[7] A. Gibbons, "Complex behavior arose at dawn of humans," *Science*, vol. 359, no. 6381, pp. 1200–1201, Mar. 2018.
[8] P. Agudelo-Romero, P. Carbonell, M. A. Perez-Amador, and S. F. Elena, "Virus adaptation by manipulation of host's gene expression," *PLoS ONE*, vol. 3, no. 6, p. e2397, Jun. 2008.
[9] L. P. Villarreal, "Are viruses alive?" *Sci. Amer.*, vol. 291, no. 6, pp. 100–105, Dec. 2004.
[10] M. Heidegger, *An Introduction to Metaphysics*, New Haven, U.K.: Yale Univ. Press, 1959, pp. 7–8.
[11] A. Witherall, "The fundamental question," *J. Philos. Res.*, vol. 26, pp. 53–87, Jan. 2001.
[12] R. W. Sharples, *Peripatetic Philosophy, 200 BC to AD 200: An Introduction and Collection of Sources in Translation* (Cambridge Source Books in Post-Hellenistic Philosophy). Cambridge, NY, USA: Cambridge Univ. Press, 2010.
[13] C. Shields, "Aristotle," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., Spring 2022 Ed. Stanford, CA, USA: Stanford Univ., 2022. Accessed: Oct. 1, 2022. [Online]. Available: https://plato.stanford.edu/archives/spr2022/entries/aristotle/
[14] D. P. S. Bhawuk, *Spirituality and Indian Psychology*, New York, NY, USA: Springer, 2011, p. 172.
[15] S. H. Rizvi, "Avicenna (Ibn Sina) (c. 980—1037)," in *The Internet Encyclopedia of Philosophy*, 2006. Accessed: Jan. 20, 2022. [Online]. Available: https://iep.utm.edu/ and https://iep.utm.edu/avicenna-ibn-sina/
[16] G. Hatfield, "René descartes," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., Summer 2018 ed. Stanford, CA, USA: Stanford Univ., 2018. Accessed: Jan. 18, 2020. [Online]. Available: https://plato.stanford.edu/archives/sum2018/entries/descartes/
[17] R. Descartes, "Objections to the meditations and descartes's replies," in *Early Modern Philosophy*, J. Bennett, Ed. 2007. Accessed: Jan. 20, 2020. [Online]. Available: http://www.earlymoderntexts.com and http://www.earlymoderntexts.com/assets/pdfs/descartes1642_2.pdf
[18] P. Markie and M. Folescu, "Rationalism vs. empiricism," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., Fall 2021 ed. Stanford, CA, USA: Stanford Univ., 2021. Accessed: Oct. 1, 2022. [Online]. Available: https://plato.stanford.edu/archives/fall2021/entries/rationalism-empiricism/
[19] J. Locke, *An Essay Concerning Human Understanding*, London, U.K.: Thomas Basset, 1690.
[20] W. Uzgalis, "John Locke," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., Spring 2019 ed. Stanford, CA, USA: Stanford Univ., 2018. Accessed: Jan. 20, 2020. [Online]. Available: https://plato.stanford.edu/archives/spr2019/entries/locke/
[21] I. Kant, *Critique of Pure Reason*, vol. 1781, M. Weigelt and M. Müller, Eds. London, U.K.: Penguin Classics. Jan. 2008.
[22] M. Rohlf, "Immanuel Kant," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., Fall 2020 ed. Stanford, CA, USA: Stanford Univ., 2020. Accessed: Oct. 1, 2021. [Online]. Available: https://plato.stanford.edu/archives/fall2020/entries/kant/
[23] L. Campbell, "Kant, autonomy and bioethics," *Ethics, Med. Public Health*, vol. 3, no. 3, pp. 381–392, Jul. 2017.
[24] M. Meulders, *Helmholtz: From Enlightenment to Neuroscience*, L. Garey, Ed. Cambridge, MA, USA: MIT Press, 2012.
[25] L. Patton, "Signs, toy models, and the a priori: From Helmholtz to Wittgenstein," *Stud. Hist. Philosophy Sci. A*, vol. 40, no. 3, pp. 281–289, 2009.
[26] L. Patton, "Hermann von Helmholtz," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., Winter 2018 ed. Stanford, CA, USA: Stanford Univ., 2018. Accessed: Jan. 20, 2020. [Online]. Available: https://plato.stanford.edu/archives/win2018/entries/hermann-helmholtz/
[27] G. E. Hinton and R. S. Zemel, "Autoencoders, minimum description length, and Helmholtz free energy," in *Advances in Neural Information Processing Systems*, vol. 6, J. D. Cowan, G. Tesauro, and J. Alspector, Eds. San Mateo, CA, USA: Morgan Kaufmann, 1994.
[28] P. Dayan, G. E. Hinton, R. Neal, and R. S. Zemel, "The Helmholtz machine," *Neural Comput.*, vol. 7, pp. 1022–1037, Sep. 1995.
[29] K. Friston, "The free-energy principle: A rough guide to the brain?" *Trends Cognit. Sci.*, vol. 13, no. 7, pp. 293–301, Jul. 2009.
[30] T. Parr, G. Pezzulo, and K. J. Friston, *The Free Energy Principle in Mind, Brain, and Behavior*, Cambridge, MA, USA: MIT Press, 2022.
[31] E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, and A. J. Hudspeth, *Principles of Neural Science*, 5th ed. New York, NY, USA: McGraw-Hill, 2012.
[32] A. M. Turing, "Intelligent machinery," Nat. Phys. Lab., Teddington, U.K., Tech. Rep., 1948.
[33] N. Wiener, *Cybernetics: Or Control and Communication in the Animal and the Machine*, Paris, France: Hermann & Cie, 1948.
[34] W. G. Walter, "An imitation of life," *Sci. Amer.*, vol. 182, no. 5, pp. 42–45, 1950.
[35] W. G. Walter, "A machine that learns," *Sci. Amer.*, vol. 185, no. 2, pp. 60–63, Aug. 1951.
[36] L. F. Barrett, *How Emotions Are Made: The Secret Life of the Brain*. Boston, MA, USA: Houghton Mifflin Harcourt, 2017.
[37] R. C. Arkin, *Behavior-Based Robotics*. Cambridge, MA, USA: MIT Press, 1998.
[38] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, 1948.
[39] D. O. Hebb, *The Organization of Behavior*. New York, NY, USA: Wiley, 1949.
[40] J. McCarthy, M. Minsky, N. Rochester, and C. E. Shannon. (1955). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. Accessed: Jan. 18, 2018. [Online]. Available: http://raysolomonoff.com/dartmouth/boxa/dart564props.pdf
[41] J. Moor, "The Dartmouth college artificial intelligence conference: The next fifty years," *AI Magazine*, vol. 27, no. 4, pp. 87–89, 2006.
[42] A. Newell, *Unified Theories of Cognition*, Cambridge MA, USA: Harvard, Univ. Press, 1990.
[43] J. R. Anderson, "ACT: A simple theory of complex cognition," *Amer. Psychol.*, vol. 51, no. 4, pp. 355–365, Apr. 1996.
[44] D. E. Kieras and D. E. Meyer, "An overview of the EPIC architecture for cognition and performance with application to human-computer interaction," *Hum.-Comput. Interact.*, vol. 12, no. 4, pp. 391–438, Dec. 1997.
[45] P. Langley, "Cognitive architectures and the construction of intelligent agents," in *Proc. AAAI Workshop Intell. Agent Archit.*, Stanford, CA, USA, 2004, p. 82.
[46] A. Newell, "Physical symbol systems," *Cogn. Sci.*, vol. 4, no. 2, pp. 135–183, 1980.
[47] F. Rosenblatt, "The perceptron a perceiving and recognizing automaton," Cornell Aeronaut. Lab., Buffalo, New York, NY, USA, Tech. Rep., 85-60-1, 1957.
[48] M. Minsky and S. Papert, *Perceptrons*. Cambridge, MA, USA: MIT Press, 1969.
[49] B. Kuipers, E. A. Feigenbaum, P. E. Hart, and N. J. Nilsson, "Shakey: From conception to history," *AI Mag.*, vol. 38, no. 1, pp. 88–103, Mar. 2017.
[50] SRI International. (2020). *75 Years of Innovation: Shakey the Robot*. Accessed: Jan. 20, 2022. [Online]. Available: https://medium.com/dish/75-years-of-innovation-shakey-the-robot-385af2311ec8
[51] W. R. Ashby, *An Introduction to Cybernetics*. London, U.K.: Chapman, 1956.
[52] W. R. Ashby, "Requisite variety and its implications for the control of complex systems," *Cybernetica*, vol. I, no. 2, pp. 83–99, 1958.
[53] R. C. Conant and W. R. Ashby, "Every good regulator of a system must be a model of that system," *Int. J. Syst. Sci.*, vol. 1, pp. 89–97, Oct. 1970.
[54] A. H. Klopf, "Brain function and adaptive systems—A heterostatic theory," Air Force Cambridge Res. Laboratories, Bedford, MA, USA, Tech. Rep., AFCRL-72-0164, 1972.
[55] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.
[56] J. H. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor, Michigan: Univ. Michigan Press, 1975.

[57] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Boston, MA, USA: Addison Wesley, 1989.

[58] J. R. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Evolution*. Cambridge, MA, USA: MIT Press, 1992.

[59] I. Rechenberg, *Evolutionsstrategie: Optimierung Technischer Systeme Nach Prinzipien Der Biologischen Evolution*. Stuttgart, Germany: Frommann-Holzboog Verlag, 1973.

[60] H. P. Schwefel, *Numerical Optimization of Computer Models*. Hoboken, NJ, USA: Wiley, 1981.

[61] L. J. Fogel, A. J. Owens, and M. J. Walsh, *Artificial Intelligence Through Simulated Evolution*. Hoboken, NJ, USA: Wiley, 1966.

[62] G. M. Edelman and V. B. Mountcastle, *The Mindful Brain: Cortical Organization and the Group-Selective Theory of Higher Brain Function*. Cambridge, MA, USA: MIT Press, 1978.

[63] G. N. Reeke, O. Sporns, and G. M. Edelman, "Synthetic neural modeling: The 'Darwin' series of recognition automata," *Proc. IEEE*, vol. 78, no. 9, pp. 1498–1530, Sep. 1990.

[64] G. M. Edelman, "Learning in and from brain-based devices," *Science*, vol. 318, no. 5853, pp. 1103–1105, Nov. 2007.

[65] J. L. Krichmar, "Design principles for biologically inspired cognitive robotics," *Biologically Inspired Cognit. Archit.*, vol. 1, pp. 73–81, Jul. 2012.

[66] A. G. Barto, R. S. Sutton, and C. W. Anderson, "Neuronlike elements that can solve difficult learning control problems," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-13, no. 5, pp. 835–846, Sep./Oct. 1983.

[67] R. Brooks, "A robust layered control system for a mobile robot," *IEEE J. Robot. Autom.*, vol. JRA-2, no. 1, pp. 14–23, Mar. 1986.

[68] E. L. Thorndike, *The Elements of Psychology*. New York, NY, USA: A. G. Seiler, 1905.

[69] I. P. Pavlov, *Conditioned Reflexes*. London, U.K.: Routledge and Kegan Paul, 1927.

[70] H. C. Blodgett, "The effect of the introduction of reward upon the maze performance of rats," *Univ. California Publications Psychol.*, vol. 4, pp. 113–134, Jan. 1929.

[71] B. F. Skinner, *The Behavior of Organisms: An Experimental Analysis*. Cambridge, MA, USA: B. F. Skinner Foundation, 1938.

[72] J. Piaget, *Origins of Intelligence in the Child*. London, U.K.: Routledge & Kegan Paul, 1936.

[73] C. L. Hull, "Mind, mechanism, and adaptive behavior," *Psychol. Rev.*, vol. 44, no. 1, pp. 1–32, 1937.

[74] W. J. Brogden, "Sensory pre-conditioning," *J. Experim. Psychol.*, vol. 25, no. 4, pp. 323–332, Sep. 1939.

[75] K. F. Muenzinger, "Vicarious trial and error at a point of choice: I. A general survey of its relation to learning efficiency," *Pedagogical Seminary J. Genetic Psychol.*, vol. 53, no. 1, pp. 75–86, Sep. 1938.

[76] E. C. Tolman, "The determiners of behavior at a choice point," *Psychol. Rev.*, vol. 45, no. 1, pp. 1–41, 1938.

[77] G. A. Miller, "The cognitive revolution: A historical perspective," *Trends Cognit. Sci.*, vol. 7, no. 3, pp. 141–144, Mar. 2003.

[78] S. A. McLeod. (2015). *Cognitive Psychology in Simply Psychology*. Accessed: Jan. 20, 2022. [Online]. Available: https://www.simplypsychology.org/cognitive.html

[79] S. A. McLeod. (2017). *Behaviorist Approach in Simply Psychology*. Accessed: Jan. 20, 2020. [Online]. Available: https://www.simplypsychology.org/behaviorism.html

[80] G. E. Hinton, S. Osindero, and Y. W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 54–1527, 2006.

[81] C. Monroy. (2022). *What is Warehouse Robotics?* Accessed: Jun. 30, 2022. [Online]. Available: https://6river.com/what-is-warehouse-robotics/

[82] M. L. Anderson, "Embodied cognition: A field guide," *Artif. Intell.*, vol. 149, no. 1, pp. 91–130, Sep. 2003.

[83] D. Vernon, G. Metta, and G. Sandini, "A survey of artificial cognitive systems: Implications for the autonomous development of mental capabilities in computational agents," *IEEE Trans. Evol. Comput.*, vol. 11, no. 2, pp. 151–180, Apr. 2007.

[84] T. Ziemke, "What's that thing called embodiment?" in *Proc. 25th Annu. Meeting Cogn. Sci. Soc.*, 2003, pp. 1–7.

[85] H. De Garis, C. Shuo, B. Goertzel, and L. Ruiting, "A world survey of artificial brain projects, Part I: Large-scale brain simulations," *Neurocomputing*, vol. 74, nos. 1–3, pp. 3–29, Dec. 2010.

[86] B. Goertzel, R. Lian, I. Arel, H. De Garis, and S. Chen, "A world survey of artificial brain projects, Part II: Biologically inspired cognitive architectures," *Neurocomputing*, vol. 74, nos. 1–3, pp. 30–49, Dec. 2010.

[87] J. S. Lowe and P. G. Anderson, *Musculoskeletal System*, 4th ed. Amsterdam, The Netherlands: Elsevier, 2015.

[88] M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini, "Developmental robotics: A survey," *Connection Sci.*, vol. 15, no. 4, pp. 151–190, Dec. 2003.

[89] J. Weng, "Developmental robotics: Theory and experiments," *Int. J. Humanoid Robot.*, vol. 1, no. 2, pp. 199–236, Jun. 2004.

[90] M. Asada, K. Hosoda, Y. Kuniyoshi, H. Ishiguro, T. Inui, Y. Yoshikawa, M. Ogino, and C. Yoshida, "Cognitive developmental robotics: A survey," *IEEE Trans. Auton. Mental Develop.*, vol. 1, no. 1, pp. 12–34, May 2009.

[91] P.-Y. Oudeyer, "On the impact of robotics in behavioral and cognitive sciences: From insect navigation to human cognitive development," *IEEE Trans. Auto. Mental Develop.*, vol. 2, no. 1, pp. 2–16, Mar. 2010.

[92] Y. Jin and Y. Meng, "Morphogenetic robotics: An emerging new field in developmental robotics," *IEEE Trans. Syst., Man, Cybern., C*, vol. 41, no. 2, pp. 145–160, Mar. 2011.

[93] A. Cangelosi and M. Schlesinger, *Developmental Robotics: From Babies to Robots*. Cambridge, MA, USA: MIT Press, 2015.

[94] H. Min, C. Yi, R. Luo, J. Zhu, and S. Bi, "Affordance research in developmental robotics: A survey," *IEEE Trans. Cognit. Develop. Syst.*, vol. 8, no. 4, pp. 237–255, Dec. 2016.

[95] A. Stoytchev, "Some basic principles of developmental robotics," *IEEE Trans. Auto. Mental Develop.*, vol. 1, no. 2, pp. 122–130, Aug. 2009.

[96] J. Krichmar and G. M. Edelman, "Principles underlying the construction of brain-based devices," in *Proc. AISB*, 2006, pp. 37–42.

[97] J. Krichmar and T. Hwu, "Design principles for neurorobotics," *Frontiers Neurorobotics*, vol. 16, May 2022, Art. no. 882518.

[98] R. Pfeifer, F. Iida, and J. Bongard, "New robotics: Design principles for intelligent systems," *Artif. Life*, vol. 11, nos. 1–2, pp. 99–120, Jan. 2005.

[99] R. M. Ryan and E. L. Deci, "Intrinsic and extrinsic motivations: Classic definitions and new directions," *Contemp. Educ. Psychol.*, vol. 25, no. 1, pp. 54–67, 2000.

[100] P. Y. Oudeyer and F. Kaplan, "What is intrinsic motivation? A typology of computational approaches," *Frontiers Neurorobot.*, vol. 1, pp. 1–14, Nov. 2007, Art. no. 6.

[101] T. J. Wiltshire, S. F. Warta, D. Barber, and S. M. Fiore, "Enabling robotic social intelligence by engineering human social-cognitive mechanisms," *Cognit. Syst. Res.*, vol. 43, pp. 190–207, Jun. 2017.

[102] E. N. Chiang and P. S. Wrightson, *Intelligent Human-Machine Collaboration: Summary of a Workshop*. Washington, DC, USA: The National Academies Press, National Research Council, 2012.

[103] N. Tinbergen, "On aims and methods of ethology," *Zeitschrift Fur Tierpsychologie*, vol. 20, no. 4, pp. 410–433, Apr. 2010.

[104] B. A. Conley, "Mayr and Tinbergen: Disentangling and integrating," *Biol. Philosophy*, vol. 35, no. 1, pp. 1–23, Feb. 2020.

[105] E. Mayr, "Cause and effect in biology: Kinds of causes, predictability, and teleology are viewed by a practicing biologist," *Science*, vol. 134, no. 3489, pp. 1501–1506, 1961.

[106] T. C. Scott-Phillips, T. E. Dickins, and S. A. West, "Evolutionary theory and the ultimate-proximate distinction in the human behavioral sciences," *Perspect. Psychol. Sci.*, vol. 6, no. 1, pp. 38–47, Jan. 2011.

[107] Merriam-Webster. *Cognition*. Accessed: Oct. 1, 2022. [Online]. Available: https://www.merriam-webster.com/dictionary/cognition

[108] S. E. G. Lea and B. Osthaus, "In what sense are dogs special? Canine cognition in comparative context," *Learn. Behav.*, vol. 46, no. 4, pp. 335–363, Dec. 2018.

[109] W. Huitt and J. Hummel, "Piaget's theory of cognitive development," in *Educational Psychology Interactive*. Valdosta, GA, USA: Valdosta State Univ, May 2020. [Online]. Available: http://www.edpsycinteractive.org/topics/cognition/piaget.html

[110] M. F. Mascolo and K. W. Fischer, "Dynamic development of thinking, feeling, and acting," in *Handbook of Child Psychology and Developmental Science: Theory and Method*, W. F. Overton, P. C. M. Molenaar, and R. M. Lerner, Eds. Hoboken, NJ, USA: Wiley, 2015, pp. 113–161.

[111] J. Rasmussen, "Skills, rules, and knowledge; Signals, signs, and symbols, and other distinctions in human performance models," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-13, no. 3, pp. 257–266, May 1983.

[112] J. Rasmussen, "The role of hierarchical knowledge representation in decision making and system management," *IEEE Trans. Syst., Man Cybern.*, vol. SMC-15, no. 2, pp. 234–243, Mar./Apr. 1985.
[113] A. N. Meltzoff and M. K. Moore, "Imitation, memory, and the representation of persons," *Infant Behav. Develop.*, vol. 17, no. 1, pp. 83–99, Jan. 1994.
[114] A. N. Meltzoff and M. K. Moore, "Explaining facial imitation: A theoretical model," *Early Develop. Parenting*, vol. 6, nos. 3–4, pp. 179–192, Sep. 1997.
[115] Wiktionary. *Episode*. Accessed: Oct. 1, 2022. [Online]. Available: https://en.wiktionary.org/wiki/episode
[116] D. M. Wolpert and M. Kawato, "Multiple paired forward and inverse models for motor control," *Neural Netw.*, vol. 11, nos. 7–8, pp. 1317–1329, Oct. 1998.
[117] M. Kawato and A. Cortese, "From internal models toward metacognitive AI," *Biol. Cybern.*, vol. 115, no. 5, pp. 415–430, Oct. 2021.
[118] G. Kim, J. Laurens, T. A. Yakusheva, and P. M. Blazquez, "The macaque cerebellar flocculus outputs a forward model of eye movement," *Frontiers Integrative Neurosci.*, vol. 13, p. 12, Apr. 2019.
[119] Online Etymology Dictionary. *Procedure*. Accessed: Oct. 1, 2022. [Online]. Available: https://www.etymonline.com/word/procedure
[120] W. Starr, "Counterfactuals," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., Summer 2021 ed. Stanford, CA, USA: Stanford Univ., 2021. Accessed: Jan. 20, 2022. [Online]. Available: https://plato.stanford.edu/archives/sum2021/entries/counterfactuals/
[121] J. Pearl and D. Mackenzie, *The Book of Why: The New Science of Cause and Effect*. New York, NY, USA: Basic Books, 2018.
[122] Online Etymology Dictionary. *Concept*. Accessed: Oct. 1, 2022. [Online]. Available: https://www.etymonline.com/word/concept
[123] Wiktionary. *Concipio*. Accessed: Oct. 1, 2022. [Online]. Available: https://en.wiktionary.org/wiki/concipio#Latin
[124] Merriam-Webster. *Concept*. Accessed: Oct. 1, 2022. [Online]. Available: https://www.merriam-webster.com/dictionary/concept
[125] Online Etymology Dictionary. *Social*. Accessed: Oct. 1, 2022. [Online]. Available: https://www.etymonline.com/word/social
[126] Wiktionary. *Social*. Accessed: Oct. 1, 2022. [Online]. Available: https://en.wiktionary.org/wiki/social
[127] Wiktionary. *Anticipate*. Accessed: Oct. 1, 2022. [Online]. Available: https://en.wiktionary.org/wiki/anticipate
[128] Wiktionary. *Predict*. Accessed: Oct. 1, 2022. [Online]. Available: https://en.wiktionary.org/wiki/predict
[129] Wiktionary. *Prepare*. Accessed: Oct. 1, 2022. [Online]. Available: https://en.wiktionary.org/wiki/prepare
[130] Wiktionary. *Prospect*. Accessed: Oct. 1, 2022. [Online]. Available: https://en.wiktionary.org/wiki/prospect
[131] R. Rosen, "Anticipatory systems," in *Anticipatory Systems, IFSR International Series on Systems Science and Engineering*, vol. 1. New York, NY, USA: Springer, 2012.
[132] G. Pezzulo, "Coordinating with the future: The anticipatory nature of representation," *Minds Mach.*, vol. 18, pp. 179–225, Jan. 2008.
[133] M. Butz, O. Sigaud, and P. Gerard, "Internal models and anticipations in adaptive learning systems," *Anticipatory Behavior in Adaptive Learning Systems* (Lecture Notes in Computer Science), vol. 2684. Berlin, Germany: Springer, 2002.
[134] K. Cuevas and A. Giles, "Transitions in the temporal parameters of sensory preconditioning during infancy," *Develop. Psychobiology*, vol. 58, no. 7, pp. 794–807, Nov. 2016.
[135] J. O'Keefe and L. Nadel, *The Hippocampus as a Cognitive Map*. New York, NY, USA: Oxford, 1978.
[136] D. Hu and A. Amsel, "A simple test of the vicarious trial-and-error hypothesis of hippocampal function," *Proc. Nat. Acad. Sci. USA*, vol. 92, no. 12, pp. 5506–5509, 1995.
[137] K. Pearson and J. Gordon, "Spinal reflexes," in *Principles of Neural Science*, E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, A. J. Hudspeth, and S. Mack, Eds., 4th ed. New York, NY, USA: McGraw-Hill, 2012.
[138] C. S. Sherrington, *The Integrative Action of the Nervous System*. New Haven, U.K.: Yale Univ. Press, 1906.
[139] E. P. Gardner and K. O. Johnson, "Sensory coding," in *Principles of Neural Science*, E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, A. J. Hudspeth, and S. Mack, Eds., 4th ed. New York, NY, USA: McGraw-Hill, 2012.
[140] S. G. Lisberger, "The neural basis for learning of simple motor skills," *Science*, vol. 242, no. 4879, pp. 728–735, Nov. 1988.
[141] N. H. Frijda, *The Emotions*. Cambridge, U.K.: Cambridge Univ. Press, 1986, p. 207.
[142] C. M. Brendl and E. T. Higgins, "Principles of judging valence: What makes events positive or negative?" in *Advances in Experimental Social Psychology*, vol. 28. Cambridge, MA, USA: Academic Press, 1996, pp. 95–160.
[143] D. L. Shacter and A. D. Wagner, "Learning and memory," in *Principles of Neural Science*, E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, A. J. Hudspeth, and S. Mack, Eds., 4th ed. New York, NY, USA: McGraw-Hill, 2012.
[144] J. E. Steinmetz and D. H. Lindquist, "Neuronal basis of learning," in *Handbook of Neuroscience for the Behavioral Sciences*, vol. 1, G. G. Berntson and J. T. Cacioppo, Eds. Hoboken, NJ, USA: Wiley, 2009, pp. 507–527.
[145] R. F. Thompson, "Habituation: A history," *Neurobiol. Learn. Memory*, vol. 92, no. 2, pp. 127–134, Sep. 2009.
[146] E. R. Kandel and A. J. Hudspeth, "The brain and behavior," in *Principles of Neural Science*, E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, A. J. Hudspeth, and S. Mack, Eds., 4th ed. New York, NY, USA: McGraw-Hill, 2012.
[147] R. A. Rescorla, "Pavlovian conditioning: It's not what you think it is," *Amer. Psychologist*, vol. 43, no. 3, p. 151, 1988.
[148] C. R. Gallistel and P. D. Balsam, "Time to rethink the neural mechanisms of learning and memory," *Neurobiol. Learn. memory*, vol. 108, pp. 136–144, 2014.
[149] M. Moscovitch, R. Cabeza, G. Winocur, and L. Nadel, "Episodic memory and beyond: The hippocampus and neocortex in transformation," *Annu. Rev. Psychol.*, vol. 67, no. 1, pp. 105–134, Jan. 2016.
[150] T. W. Buchanan, "Retrieval of emotional memories," *Psychol. Bull.*, vol. 133, no. 5, p. 761, 2007.
[151] G. Richter-Levin and I. Akirav, "Amygdala-hippocampus dynamic interaction in relation to memory," *Mol. Neurobiol.*, vol. 22, no. 1, pp. 11–20, 2000.
[152] Y. Yang and J.-Z. Wang, "From structure to behavior in basolateral amygdala-hippocampus circuits," *Frontiers Neural Circuits*, vol. 11, p. 86, Oct. 2017.
[153] A. Pritzel, B. Uria, S. Srinivasan, A. Puigdomènech, O. Vinyals, D. Hassabis, D. Wierstra, and C. Blundell, "Neural episodic control," 2017, *arXiv:1703.01988*.
[154] R. A. Rescorla, "Probability of shock in the presence and absence of CS in fear conditioning," *J. Comparative Physiological Psychol.*, vol. 66, no. 1, pp. 1–5, 1968.
[155] W. Schultz, P. Dayan, and P. R. Montague, "A neural substrate of prediction and reward," *Science*, vol. 275, no. 5306, pp. 1593–1599, 1997.
[156] J. C. Horvitz, "Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events," *Neuroscience*, vol. 96, no. 4, pp. 651–656, Mar. 2000.
[157] S. Kakade and P. Dayan, "Dopamine: Generalization and bonuses," *Neural Netw.*, vol. 15, nos. 4–6, pp. 549–559, Jun. 2002.
[158] R. Ventura, C. Morrone, and S. Puglisi-Allegra, "Prefrontal/accumbal catecholamine system determines motivational salience attribution to both reward- and aversion-related stimuli," in *Proc. Nat. Acad. Sci. USA*, vol. 104, no. 12, pp. 5181–5186, 2007.
[159] J. M. Wenzel, N. A. Rauscher, J. F. Cheer, and E. B. Oleson, "A role for phasic dopamine release within the nucleus accumbens in encoding aversion: A review of the neurochemical literature," *ACS Chem. Neurosci.*, vol. 6, no. 1, pp. 16–26, Jan. 2015.
[160] T. Wichmann and M. R. DeLONG, "Pathophysiology of Parkinson's disease: The MPTP primate model of the human disorder," *Ann. New York Acad. Sci.*, vol. 991, no. 1, pp. 199–213, Jan. 2006.
[161] A. Aubret, L. Matignon, and S. Hassas, "A survey on intrinsic motivation in reinforcement learning," 2019, *arXiv:1908.06976*.
[162] R. Bellman, "The theory of dynamic programming," *Bull. Amer. Math. Soc.*, vol. 60, no. 6, pp. 503–516, 1954.
[163] C. Watkins, "Learning from delayed rewards," Ph.D. thesis, King's College, Cambridge Univ., 1989.
[164] N. Sajid, P. J. Ball, T. Parr, and K. J. Friston, "Active inference: Demystified and compared," *Neural Comput.*, vol. 33, no. 3, pp. 674–712, 2021.
[165] B. Milner, L. R. Squire, and E. R. Kandel, "Cognitive neuroscience and the study of memory," *Neuron*, vol. 20, no. 3, pp. 445–468, 1998.

[166] V. Gallese, "Mirror neurons and the social nature of language: The neural exploitation hypothesis," *Social Neurosci.*, vol. 3, nos. 3–4, pp. 317–333, Sep. 2008.

[167] S. Hurley, "The shared circuits model (SCM): how control, mirroring, and simulation can enable imitation, deliberation, and mindreading," *Behav. brain Sci.*, vol. 31, no. 1, pp. 1–58, 2008.

[168] S. Dehaene and L. Cohen, "Cultural recycling of cortical maps," *Neuron*, vol. 56, no. 2, pp. 384–398, Oct. 2007.

[169] M. L. Anderson, "The massive redeployment hypothesis and the functional topography of the brain," *Phil. Psychol.*, vol. 20, no. 2, pp. 143–174, Apr. 2007.

[170] D. Crystal, *English as a Global Language*. Cambridge, U.K.: Cambridge Univ. Press, 1997.

[171] K. J. Holyoak and R. G. Morrison, *The Cambridge Handbook of Thinking and Reasoning*, Cambridge, U.K.: Cambridge Univ. Press, 2005.

[172] G. L. Murphy, *The Big Book of Concepts*. Cambridge, MA, USA: MIT Press, 2002.

[173] R. M. Balzer, "EXDAMS—Extendable debugging and monitoring system," RAND Corp., Santa Monica, CA, USA, Tech. Rep. RM-5772-ARPA, 1969.

[174] J. C. King, "A new approach to program testing," in *Proc. Int. Conf. Reliable Softw.*, Apr. 1975, pp. 228–233.

[175] R. S. Boyer, B. Elspas, and K. N. Levitt, "SELECT—A formal system for testing and debugging programs by symbolic execution," in *Proc. Int. Conf. Reliable Softw.*, 1975, pp. 234–245.

[176] L. A. Clarke, "A program testing system," in *Proc. Annu. Conf.*, 1976, pp. 488–491.

[177] S. K. Cha, T. Avgerinos, A. Rebert, and D. Brumley, "Unleashing mayhem on binary code," in *Proc. IEEE Symp. Secur. Privacy*, May 2012, pp. 380–394.

[178] D. Kuts, "Towards symbolic pointers reasoning in dynamic symbolic execution," 2021, *arXiv:2109.03698*.

[179] T. Kriete, D. C. Noelle, J. D. Cohen, and R. C. O'Reilly, "Indirection and symbol-like processing in the prefrontal cortex and basal ganglia," *Proc. Nat. Acad. Sci. USA*, vol. 110, no. 41, pp. 16390–16395, Oct. 2013.

[180] C. Zhang, M. Raghu, J. Kleinberg, and S. Bengio, "Pointer value retrieval: A new benchmark for understanding the limits of neural network generalization," 2021, *arXiv:2107.12580*.

[181] P. Blouw, E. Solodkin, P. Thagard, and C. Eliasmith, "Concepts as semantic pointers: A framework and computational model," *Cognit. Sci.*, vol. 40, no. 5, pp. 1128–1162, Jul. 2016.

[182] J. McWhorter, *What Language Is: And What It Isn't and What It Could Be*. New York, NY, USA: Penguin Random House, 2012.

[183] S. Kugele, A. Petrovska, and I. Gerostathopoulos, "Towards a taxonomy of autonomous systems," in *Software Architecture* (Lecture Notes in Computer Science), vol. 12857, S. Biffl, E. Navarro, W. Lowe, M. Sirjani, R. Mirandola, and D. Weyns, Eds. Cham, Switzerland: Springer, 2021.

[184] N. Noori, "The symbolic working memory: Memory accommodations for schematic processing of symbolic information," in *Proc. Annu. Conf. Cogn. Sci. Soc.*, Aug. 2015, pp. 1–6.

[185] I. Asimov, *I, Robot*. New York, NY, USA: Del Rey Books, Apr. 2008.

[186] D. C. Ong, J. Zaki, and N. D. Goodman, "Computational models of emotion inference in theory of mind: A review and roadmap," *Topics Cognit. Sci.*, vol. 11, no. 2, pp. 338–357, Apr. 2019.

[187] D. Sznycer and A. Cohen, "How pride works," *Evol. Hum. Sci.*, vol. 3, p. e10, Jan. 2021.

[188] A. LaVarco, N. Ahmad, Q. Archer, M. Pardillo, R. N. Castaneda, A. Minervini, and J. P. Keenan, "Self-conscious emotions and the right fronto-temporal and right temporal parietal junction," *Brain Sci.*, vol. 12, no. 2, p. 138, Jan. 2022.

[189] *Degrees of Protection Provided by Enclosures (IP Code), 2.2 Ed, International Electrotechnical Commission*, document IEC 60529, p. 21, 2013.

[190] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin, "An integrated theory of the mind," *Psychol. Rev.*, vol. 111, no. 4, p. 1036, 2004.

[191] M. Wilson, "Six views of embodied cognition," *Psychonomic Bull. Rev.*, vol. 9, no. 4, pp. 625–636, 2002.

[192] N. D. Leitan and L. Chaffey, "Embodied cognition and its applications: A brief review," *Sensoria, A J. Mind, Brain Culture*, vol. 10, no. 1, pp. 3–10, 2014.

[193] B. Z. Mahon, "The burden of embodied cognition," *Can. J. Exp. Psychol./Revue Canadienne de Psychologie Expérimentale*, vol. 69, no. 2, pp. 172–178, 2015.

[194] R. P. Bonasso, R. J. Firby, E. Gat, D. Kortenkamp, D. P. Miller, and M. G. Slack, "Experiences with an architecture for intelligent, reactive agents," *J. Exp. Theor. Artif. Intell.*, vol. 9, pp. 2–3, pp. 237–256, 1997.

[195] D. Kortenkamp and R. Simmons, "Robotic systems architectures and programming," in *Springer Handbook of Robotics*, B. Siciliano O. Khatib, Eds. Berlin, Germany: Springer, 2008, pp. 187–206.

[196] A. V. Martynov, "Ontogenetic systematics: The synthesis of taxonomy, phylogenetics, and evolutionary developmental biology," *Paleontological J.*, vol. 46, no. 8, pp. 833–864, Dec. 2012.

[197] P. J. Bowler, "The changing meaning of 'evolution,'" *J. Hist. Ideas*, vol. 36, pp. 95–114, Jan. 1975.

SATORU ISAKA (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in systems science from the University of California at San Diego, La Jolla, CA, USA, in 1984, 1986, and 1989, respectively.

From 1990 to 1998, he was a Research Scientist at OMRON in factory automation. From 1998 to 2004, he worked with a series of venture startups as the Chief Technology Officer and the Chief Scientist in customer service automation, natural language processing, and speech interaction. In 2004, he founded Vision Del Mar, LLC., to focus on research in the field of public health and machine learning. His research topic was on adaptive control systems theory and artificial intelligence. His current research interests include embedded networked control systems and their developmental behavior.