Minimal mind

Alexei A. Sharov
National Institute on Aging, Baltimore, USA

Citation: Sharov, A. A. 2013. Minimal mind. In: L. Swan (Ed.), Origins of Mind, Biosemiotics 8, Springer, New York, London. pp.: 343-360.

Abstract

In contrast to the human standard for mind established by Alan Turing, I search for a “minimal mind”, which is present in animals and even lower-level organisms. Mind is a tool for the classification and modeling of objects. Its origin marks an evolutionary transition from protosemiotic agents, whose signs directly control actions, to eusemiotic agents, whose signs correspond to ideal objects. The hallmark of mind is a holistic perception of objects, which is not reducible to individual features or signals. Mind can support true intentionality of agents because goals become represented by classes or states of objects. Basic components of mind appear in the evolution of protosemiotic agents, thus the emergence of mind was inevitable. The classification capacity of mind may have originated from the ability of organisms to classify states of their own body. Within primary modeling systems, ideal objects are not connected with each other and often tailored for specific functions, whereas in the secondary modeling system, ideal objects are independent from functions and become interconnected via arbitrarily established links. Testing of models can be described by commuting diagrams that integrate measurements, model predictions, object tracking, and actions. Language, which is the tertiary modeling system, supports efficient communication of models between individuals.

1. Introduction

Mind is traditionally considered as a human faculty responsible for conscious experience and intelligent thought. Components of mind include perception, memory, reason, logic, modeling of the world, motivation, emotion, and attention (Premack & Woodruff, 1978). This list can be easily expanded to other kinds of human mental activities. Defects in mental functions (e.g., in logic, attention, or communication) are considered as a loss of mind, partial or complete. In short, mind is a collection of mental functions in humans. However, this definition tells us nothing about the nature of mind. Human mental functions are so diverse that it is difficult to evaluate their relative importance. The only way to identify the most fundamental components of mind is to track its origin in animals, which inevitably leads us to the idea that mind exists beyond humans. Animal mental activities (i.e., “animal cognition”) are definitely more primitive compared to those of the human mind, but they include many common components: perception, memory, modeling of the world, motivation, and attention (Griffin, 1992; Sebeok, 1972). The lack of abstract reasoning in animals indicates that reason is not the most fundamental element of mind, but rather a late addition.

By accepting the existence of mind in animals we commit ourselves to answer many difficult questions. For example, where is the lower evolutionary threshold for mind? Does mind require brain or at least some kind of nervous system? In other words, we enter the quest for the “minimal mind”, which is the topic of this chapter. This evolutionary approach is opposite to Turing’s criterion for machine intelligence, which is based on the ability of a human to distinguish between a computer and a human being based solely on communication with them

(Turing, 1952). To be indistinguishable from a human, a machine should have a “maximal mind” that is functionally equivalent to the human mind. Here I propose that minimal mind is a tool for the classification and modeling of objects and that its origin marks an evolutionary transition from protosemiotic agents, whose signs directly control actions, to eusemiotic agents, whose signs correspond to ideal objects.

2. Agents and functional information

Mind is intrinsically related to life because it is a faculty of living systems. However, according to cybernetics, it can also exist in artificial devices (Nillson, 1998). To present a unified approach to mind, we need first to discuss briefly the nature of life and artifacts. Machine metaphor is often perceived as a misleading simplification of the phenomena of life and mind (Deacon, 2011; Emmeche & Hoffmeyer, 1991). The motivation to separate life and mind from machines comes from the fact that simple machines are manufactured and programmed by humans, whereas organisms are self-produced and develop from eggs into their definite shape (Swan & Howard, 2012). Also, machines change their state following deterministic rules rather than internal goals and values. But, despite these differences, the progress in understanding life and mind seems to lie in bridging the gap between life and artifacts rather than in building a wall between them. In particular, biological evolution can be seen as a sequence of inventions of various instruments that are needed to perform living functions (Dennett, 1995). Cellular processes are based on molecular machines that copy sequences of nucleic acids, synthesize proteins, modify them and assemble them into new molecular machines. Thus, components of organisms are manufactured, and living systems are indeed artifacts (Barbieri, 2003). Although man-made machines lack some features of living organisms, this deficiency should be attributed to our insufficient knowledge and experience. Humans only just began learning how to make self-programmable and self-repairable mechanisms, whereas living cells mastered these skills billions of years ago.

One of the heuristics of systems methodology is “functionalism”, which assumes that systems should be compared based solely on their functions rather than their material composition. This idea was initially proposed as a foundation for “relational biology” (Rashevsky, 1938; Rosen, 1970), and later was formulated as “functional isomorphism” (Putnam, 1975). If an artificial system performs the same (or similar) functions as a living organism, then there is good reason to call it “alive”. However, it would be confusing to apply the term “living organism” to artificial devices. Instead, it is better to use the term “agent” which fits equally well to living organisms and artificial devices. Agents should not be viewed only as externally-programmed devices, as is commonly done in cybernetics. Although all agents carry external programs, the majority of agents, including all living organisms, also have self-generated programs. An agent is a system with spontaneous activity that selects actions to pursue its goals. Goals are considered in a broad sense, including both achievable events (e.g., capturing a resource, reproduction), and sustained values (e.g., energy balance). Some goals are externally programmed by parental agents or higher-level agents, and other goals emerge within agents. Note that mind is not necessarily present in agents. Simple agents can automatically perform goal-directed activities based on a program.

In the field of artificial intelligence, ideas of functionalism are often misinterpreted as a primacy of the digital program over the body/hardware and environment. Internet-based programs like the virtual world of “Second Life” may convince people that their functionality can be fully digitized in the future. However, programs are not universal but instead tailored for specific bodies and environments, and therefore, can be exchanged without loss of functionality only between similar agents in similar environments. Thus, “digital immortality” is a myth (Swan & Howard, 2012). Self-producing agents have many body-specific functions associated with metabolism, assembly of subagents, growth, development, and reproduction. Obviously, these functions cannot be realized in a qualitatively different body. But functional methodology works even in this case because the body can support a large number of alternative activities, and it needs information to organize and control these activities. In summary, agents require *both* specific material organization (body) and functional information to control their actions.

Agents are always produced by other agents of comparable or higher functional complexity Sharov, 2006). This statement is an informational equivalent of the gradualism principle in the theory of evolution (Sharov, 2009b). The reason why agents cannot self-assemble spontaneously is that they carry substantial functional complexity. Long evolutionary (or learning) timelines are required to develop each new function via trail and error; therefore, simultaneous and fast emergence of numerous novel functions is unlikely. The origin of life does not contradict the principle of gradualism because primordial agents were extremely simple and started from single functions (Sharov, 2009a). The production of artificial agents by humans also satisfies the principle of gradualism because humans have a higher level of functional complexity than any human-made devices. Methods of agent manufacturing may include assembly from a set of parts, as well as self-organization and self-development. Although the majority of human-made agents is assembled, some of them use elements of development. For example, satellites can unfold and re-assemble in space after launch. Self-assembly is a common approach in nanotechnology and in synthetic organisms.

3. Functional information

Agents are unusual material objects whose dynamics cannot be effectively described by physics, although they do not contradict physics. Instead, a semiotic description appears more meaningful: agents carry functional information, which is a collection of signs that encode and control their functions. The adjective ‘functional’ helps to distinguish functional information from quantitative approaches developed by Shannon and Kolmogorov (Shannon, 1948; Kolmogorov, 1965). Although signs are material objects, they have functions within agents that are not directly associated with their physical properties.

Semiotics stems from the work of Charles Sanders Peirce, who defined a sign as a triadic relationship between a sign vehicle, object, and interpretant, which is a product of an interpretive process or a content of interpretation (Peirce, 1998). However, not all agents can associate signs with content or meaning . Thus, I prefer a more generic definition of signs as objects that are used by agents to encode and control their functions (Sharov, 2010). Most signaling processes that take place within the cells of living organisms do not invoke ideal representations, but they encode and/or control cellular functions, and thus have a semiotic nature. Peirce deemphasized the role of agents in informational processes and did not consider the agent or organism as a component of the triadic sign relationship. He thought that meanings belonged to nature rather than to agents. For example, he wrote about nature’s ability to acquire habits, which is consistent with his philosophy of objective idealism. Similar views were expressed by Jesper Hoffmeyer who assumed ‘minding nature’ (Hoffmeyer, 2010). In contrast, I view signs only in connection with agents who use them, and see no reason to consider nature an agent. Although it may be hard to refute claims that the universe or Gaia are superorganisms (Lovelock 1979), I take a conservative approach and use the notion of ‘agent’ only for those systems that clearly show a reproducible goal-directed activity and carry functional information to organize this activity (Sharov, 2010).

Functional information is inseparable from agents who use it. Living organisms are products of their genome, which controls their development and growth. In contrast, cybernetics often distinguishes information (software) from computational devices (hardware). The distinction of software and hardware is meaningful only for slave-agents like computers, which are produced and externally programmed by humans. A computer is similar to a ribosome in a living cell, because ribosomes are manufactured and externally programmed to make proteins. Programmed agents are often viewed as non-semiotic systems (Barbieri, 2008). However, this idea appears confusing because the execution of a program is a part of the semiotic activity of all agents, and agency is not possible without it. We humans are programmed genetically by our ancestors, behaviorally by our parents, and culturally by our society. These programs support our identity as a *Homo sapiens* species, as well as our race, sex, nationality, personality, and a whole range of physical and mental abilities. In addition to external programs, humans and most other organisms develop their own programs. When we learn new behaviors and skills we convert them into programs that can be executed automatically or with minimal intervention from our consciousness. These self-generated programs comprise our personal identity. Our freedom comprises only a tiny fraction of our functional behavior. In fact, freedom would be destructive if it were not well balanced with programmed functions that can correct mistakes. But evolution would not be possible if all agents were 100% externally programmed, and non-evolving agents would perish in changing environments. Thus, the role of fully-programmed agents is limited to supportive functions for other agents that are able to evolve and learn.

The meaning of functional information is grounded in a communication system, which is a set of compatible communicating agents (Sharov, 2009c). For example, the genome alone does not mean anything; it has meaning only in relation to the organisms that use it. An egg can be viewed as a minimal interpreter of the genome (Hoffmeyer, 1997). Although the structure of an egg is encoded by the genome, a real egg is needed to interpret the genome correctly. Thus, heredity is based on a combination of [genome + egg] rather than on the genome alone. This leads us to the idea that functional information is not universal but has its meaning only in relation to a certain communication system. Even a single agent is involved in a continuous self-communication through memory, and therefore can be viewed as a communication system. Memory is a message sent by an agent to its own future state, and its purpose is to preserve the agent’s ability to perform certain functions. Heredity is an extended self-communication, or inter-generation memory (Sharov, 2010). Other communication systems include multiple agents that exchange signals or messages. The most common example of such horizontal communication in living organisms is sexual reproduction, where the egg encounters an unfamiliar paternal genetic sequence. Agents from different communication systems do not exchange functional information on a regular basis because their interpretation modules are not fully compatible. For example, most interspecies hybrids in mammals are nonviable or sterile as a result of misinterpretation of the paternal genome. Communication systems often have a hierarchical structure. For example, species are partitioned into populations, which in turn are partitioned into colonies or families. Subagents within organisms (e.g., cells) make their own communication systems. Communication is often asymmetric when one kind of agent manipulates the functional information of another kind of agent. For example, agents can (re)program their sub-agents or offspring agents. Asymmetric communication often occurs between interacting organisms of different species (e.g., parasites reprogram their hosts, or preys mislead predators via mimicry and behavioral tricks). Because communication systems are multi-scale and interdependent, evolution happens at multiple levels simultaneously.

4. Emergence of mind from elementary signaling processes

Mind is not a necessary component of agents. Bacteria are examples of mindless agents that operate via elementary signaling processes such as DNA replication, transcription, translation, and molecular sensing. They do not perceive or classify objects in the outside world as humans do; instead they detect signals that directly control their actions. Direct control, however, may include multiple steps of signal transfer as well as logical gates. Following Prodi, I call this primitive level of semiosis “protosemiosis” (Prodi, 1988). Protosemiosis does not include classification or modeling of objects; it is ‘know-how’ without ‘know-what’. Because molecular signaling is so different from higher levels of semiosis, Eco excluded it from consideration in semiotics (Eco, 1976). However, the analysis of molecular signs in bacteria helps us to understand the origin and nature of signs in animals and humans; thus, protosemiosis should not be dismissed. Protosigns (i.e., signs used in protosemiosis) do not correspond to any object, which may seem confusing because our brains are trained to think in terms of objects. Although we associate a triplet of nucleotides in the mRNA with an amino acid as an object, a cell does not have a holistic internal representation of amino acid; thus, it is not an object for a cell. Instead, a triplet of nucleotides in the mRNA is associated with an action of tRNA and ribosome, which together append an amino acid to the growing protein chain.

Mind represents a higher level of information processing compared to protosemiosis because it includes classification and modeling of objects and situations (e.g., food items, partner agents, and enemies). These classifications and models represent the ‘knowledge’ an agent has about itself and its environment, which are Innenwelt and Umwelt following the terminology of Uexküll (Uexküll, 1982). I proposed calling this new level of semiosis “eusemiosis” (Sharov, 2012). Information processing in eusemiosis can no longer be tracked as a sequence of signal exchanges between components. Instead, it goes through multiple semi-redundant pathways, whose involvement may change from one instance to another, but invariantly converge on the same result. Thus, attractor domains are more important for understanding the dynamics of mind than individual signaling pathways. Classification of objects can be viewed as a three-step process. The first step is immediate perception, when various receptors send their signals to the mind, and these signals collectively reset the mind to a new state (or position in a phase space). The second step is the internal dynamics of mind which starts with the new state of mind and then converges to one of the attractors. This process is equivalent to recognition or classification. Each attractor represents a discrete meaningful category (e.g., fruit or predator), which I call ‘ideal object’. In contrast to real objects that are components of the outside world, ideal objects exist within the mind and serve as tools for classifying real objects. Finally, at the third step, the ideal object acts as a checkpoint to initiate some other function (physical or mental).

Ideal objects do not belong to a different parallel universe as claimed by Popper (1999). Instead they are tools used by agents to perceive and manipulate the real world. Following the "law of the instrument" attributed to Mark Twain, to a man with a hammer, everything looks like a nail. Thus, ideal objects within mind determine how the outside world is perceived and changed. Ideal objects are implemented as functional subunits within complex material systems, for example, as specific patterns of neuron activity or "brain-objects" (Swan & Goldberg, 2010). But the material implementation of ideal objects is flexible whereas the function is stable. Similarly, computer programs are functionally stable despite the fact that they are loaded each time into a different portion of physical memory and executed by a different processor (if available).

'Object' is one of the most complex and abstract notions in human thought. However, we should not transfer all this complexity to simple agents like worms or shells. For example, we usually distinguish between objects and their attributes, where attributes are generic (e.g., whiteness) and can be applied to various classes of objects. Although we cannot directly assess the minds of simple agents, it is unlikely that they can contemplate generic attributes. Simple agents distinguish between classes of objects but they do it unconsciously without considering attributes as independent entities. Humans can think of hypothetical ideal objects (e.g., unicorns), which include certain combinations of abstract attributes. Obviously, simple agents are not able to do that. Another difference is that humans can recognize individual objects whereas simple agents cannot distinguish objects within the same functional category. Learning and modeling capacities of mind have progressed substantially in evolution (see below), and we should not expect that simple agents have the same flexibility in connecting and manipulating of ideal objects as humans do.

Mind is a necessary tool for intentional behavior, which I consider a higher level of goal-directed activity. In contrast to protosemiotic agents, mind-equipped agents have holistic representations of their goals, which are perceived as ideal objects and integrate a large set of sensorial data. For example, immune cells of eukaryotic organisms can recognize a viral infection by the shape of the viral proteins as well as by specific features of viral nucleic acids, and launch a defense response by producing interferon, antibodies, and cytokines. Memory T-cells keep information on the properties of viral proteins acquired during the previous exposure to the same virus.

Goals may emerge internally within agents, however they can be also programmed externally. For example, instinctive behavior of organisms is programmed genetically by ancestors. In this case, ideal objects develop somehow together with the growing brain. External programming of goals is also typical for artificial minds in robotic devices equipped with automated image processing modules (Cariani, 1998, 2011). For example, a self-guided missile is programmed to classify objects into targets and non-targets and to follow the target. Agents with externally-programmed mind can support a given static set of functions in agents, but they lack adaptability, and would not be able to keep competitive advantage in changing environments. Thus, autonomous agents need adaptive minds capable of improving existing ideal objects and creating new ones via learning. Mind can generate new behaviors by creating novel attractors in the field of perception states and linking them with specific actions. If such behaviors prove useful, they can become habits and contribute to the success of agents. Requirement of learning does not imply that mind-carrying agents learn constantly. Minds may persist and function successfully in a non-learning state for a long time. Most artificial minds are static replicas of some portion of the dynamic human mind. But minds cannot improve without learning.

The statement “minds cannot improve without learning” is correct if applied to individual agents; however limited improvements of minds are possible in lineages of self-reproducing non-learning agents via genetic selection. Mutations may cause the appearance of new attractors in the dynamic state of non-learning minds or new links between ideal objects and actions. If these heritable representations help agents to perform some functions, the agents will reproduce and disseminate new behaviors within the population. This process, however, is slow and inefficient because of several problems. First, genetic selection can hardly produce any results in such highly redundant systems as minds because most changes of individual elements have no effects on the behavior. In other words, the fitness landscape is almost flat. Second, mind is a complex and well-tuned system; thus any heritable change to individual elements that does have a phenotype is likely to be disruptive. Third, the functionality of mind has to be assessed in each situation separately, because it may work in some cases but not in others. Genetic selection depends mostly on the worst outcome from a single life-threatening situation, and thus it is ineffective for improving the performance of mind in individual situations. But despite these problems it is conceivable that limited improvements on mind can be achieved by genetic selection. This helps us to explain how most primitive non-learning minds appeared in the evolution of protosemiotic agents. Moreover, simple learning algorithms may emerge in the evolution of mind solely via genetic selection, making minds adaptable and partially independent from the genetic selection (see below). But genetic mechanisms are still important for the functionality of mind even in humans because the architecture of the brain is heritable.

5. Components of minimal mind can emerge within protosemiotic agents

Because the emergence of mind is a qualitative change in organisms, it is difficult to understand the intermediate steps of this process. Here I argue that all necessary components of mind, which include semi-redundant signaling pathways, stable attractors, and adaptive learning, can emerge at the protosemiotic level. Moreover, these components emerge not as parts of mind (which does not exist yet), but as tools that increase the efficiency of other simpler functions.

Redundancy of signaling pathways may seem to be a waste of valuable resources; however, it appears beneficial for agents in the long run. First, redundancy ensures the reliability of signaling. If one pathway is blocked (e.g., as a result of injury, stress, or infection), then normal functions can be restored via alternate pathways. Each cell has multiple copies of all kinds of membrane-bound receptors because cells cannot predict the direction of incoming signals and thus distribute receptors around the whole surface. Second, redundant signaling pathways may generate novel combinatorial signals. For example, one photoreceptor can only distinguish different intensities of light, but multiple photoreceptors can identify the direction of light and even distinguish shapes. Third, redundant signaling pathways increase the adaptability of agents because some of them may start controlling novel functions in subsequent evolution.

Stable attractors are common to most auto-regulated systems, including simple devices with a negative feedback (e.g., a Watts pressure regulator). Stability is necessary for all living organisms to maintain vital functions at optimal rates. Any function that escapes regulation may become harmful and lead to disease or death. However, simple stability in the form of steady states is usually not sufficient for living organisms. Reproduction, growth, and the development of organisms require more complex regulation pathways that combine stability with change in a form of limit cycles, branching trajectories, and even chaotic attractors (Waddington, 1968).

Genetic mechanisms are not suitable for learning because the sequence of nucleotides in the DNA is not rewritable (although limited editing is possible). In contrast, simple autocatalytic networks can switch between two stable states (“on” and “off”) and serve as a dynamic memory for the cell. Moreover, such networks can support primitive learning (e.g., sensitization and habituation) as well as associative learning as follows from a simple model of two interacting genes (Ginsburg & Jablonka, 2009). In this model, genes *A* and *B* are activated by different signals S_a and S_b, and the product P_a of gene *A* has 3 functions: (1) it induces a specific phenotype or physiological response; (2) it stimulates temporarily the expression of gene *A* so that the gene remains active for some time after the initial signal S_a; and (3) it makes the expression of gene *A* dependent on the product P_b of gene *B*. If gene *A* is silent, then signal S_b activates gene *B* but its activity does not produce any phenotype. However, if signal S_b comes shortly after signal S_a, then the product P_b will activate gene *A* and produce a phenotype. This network belongs to the protosemiotic level because it is based on fixed interactions between few components.

Because all components of minimal mind can appear within protosemiotic agents, the emergence of mind seems inevitable. But there is still a problem of how to combine these components. In particular, the level of complexity of signaling pathways should increase to support a large number of dynamically interconnected attractors, and learning mechanisms have to change from individual ad-hoc assembled simple signaling networks to streamlined ideal objects. In other words, agents have to develop standard building blocks for constructing ideal objects and equip these blocks with flexible control means. It appears that epigenetic mechanisms can convert DNA segments into standard building blocks of mind, as discussed in the following section.

6. Epigenetic regulation may have supported the emergence of minimal mind

It is difficult to pinpoint the emergence of mind on the evolutionary tree of life. However, it is certain that mind appeared in eukaryotic organisms with well-developed epigenetic regulation. Epigenetic mechanisms include various changes in cells that are long-lasting but do not involve alterations of the DNA sequence. I will consider only those epigenetic mechanisms that are mediated by chromatin structure because they are likely to have facilitated the emergence of mind. Chromatin consists of DNA assembled together with histones, which are specific proteins that support the stability of DNA and regulate its accessibility to transcription factors. Histones can be modified in many ways (e.g., acetylated, methylated, phosphorylated, or ubiquitinated) by molecular agents and these modifications affect the way histones bind to each other and interact with DNA and other proteins. Some modifications convert chromatin to a highly condensed state (heterochromatin), other modifications support loose chromatin structure (euchromatin), which allows binding of transcription factors and subsequent activation of mRNA synthesis (Jeanteur, 2005). Molecular agents can both read and edit histone marks. In particular, they can modify newly-recruited histones after DNA replication in agreement with marks on the partially retained parental histones (Jeanteur, 2005). As a result, chromatin states survive cell division and are transferred to both daughter cells. Thus, chromatin-based memory signs can reliably carry rewritable information through cell lineages and control differentiation of embryos (Markoš &

Švorcová, 2009). The chromatin state depends not only on histone marks but also on other proteins that establish links between distal DNA segments, as well as links between chromatin and nuclear envelopes. These proteins, which include insulators, mediators, cohesions, and lamins, create and maintain a complex 3-dimensional structure of the chromatin (Millau & Gaudreau, 2011). Distal links create new neighborhoods and change the context for chromatin assembly.

Epigenetic mechanisms are important for the origin and function of mind because: (1) they support a practically unlimited number of attractors, (2) these attractors can be utilized as rewritable memory signs, and (3) chromatin attractors can become interconnected via products of co-localized genes. Chromatin structure is repaired after mild perturbations by special molecular agents that edit histone marks. These repair mechanisms ensure the stability of attractors in the field of chromatin states. However, strong perturbations may cross the boundary between attractors and chromatin would converge to another stable (or quasi-stable) state, which means over-writing the chromatin memory. Specific states of chromatin are spatially associated with certain genes, and these genes become activated or repressed depending on the chromatin state. Active genes produce proteins (e.g., transcription factors) which may regulate chromatin state at other genome locations. Association of chromatin with DNA is not sequence-specific, which gives organisms the flexibility to establish regulatory links between any subsets of genes.

The combination of these three features of chromatin can support adaptive learning at the cellular level. As a toy model, consider a gene that can be activated via multiple regulatory modules in its promoter. Initially the chromatin is loose at all regulatory modules, and therefore, DNA is accessible to transcription factors. Eventually, a successful action of a cell (e.g., capturing food) may become a “memory triggering event”, which forces the chromatin to condense at all regulatory modules except for the one that was functional at the time of the event. Then, as the cell encounters a similar pattern of signaling next time, only one regulatory module would become active - the one that previously mediated a successful action. Modification of chromatin (i.e., opening or closing) is controlled by the production of certain transcription factors that move from the cytoplasm to the nucleus and find specific DNA patterns where they bind. But how can transcription factors differentiate between active and non-active regulatory modules, so that only non-active modules become closed? This kind of context-dependent activity is possible thanks to the interaction between multiple transcription factors that are located close enough along the DNA sequence. For example, binding of the P300 protein to the regulatory module indicates on-going activity of this module (Visel et al., 2009), and transcription factors may have opposite effects on the chromatin depending on whether they are bound to DNA alone or in combination with P300. This kind of mechanism may support associative learning at the initial steps of the emergence of mind. An important component of this mechanism is the ability of an agent to classify its own states as ‘success’ or ‘failure’, and activate memory in the case of success.

The importance of chromatin is supported by the fact that mechanisms of learning and memory in the nervous system include DNA methylation and histone acetylation (Levenson & Sweatt, 2005; Miller & Sweatt, 2007). However, it is plausible that mind appeared even before the emergence of the nervous system. For example, unicellular ciliates have elements of non-associative learning (Wood, 1992), and even associative learning (Armus, Montgomery, & Gurney, 2006). Plants, fungi, sponges, and other multicellular organisms without nervous systems are all likely to anticipate and learn, although their responses are much slower than in animals (Ginsburg & Jablonka, 2009; Krampen, 1981). It is reasonable to assume that mind functions were initially based on intracellular mechanisms and only later they were augmented via communication between cells. Then a multicellular brain should be viewed as a community of cellular “brains” represented by the nuclei of neurons. The idea that cellular semiosis is the basis for the functionality of the brain has been recently proposed by Baslow (Baslow, 2011). The human brain consists of one hundred billion neurons and each neuron has thousands of synaptic links with other neurons. Synapses of single neurons are all specialized in various functions; some of them are active, while others are repressed. Thus, a neuron has to ‘know its synapses” because otherwise signals coming in from different synapses would be mixed up. In addition, neurons have to distinguish temporal patterns of signals coming from each synapse (Baslow, 2011). Individual neurons need at least minimal mind capacity to classify these complex inputs.

Baslow proposed that the “operating system” of neurons is based on metabolism (Baslow, 2011). Although active metabolism is indeed required for the functioning of neurons, it does not seem to be specific for mind and cannot explain how cells learn to recognize and process new signaling patterns. The cellular level of mind is more likely to be controlled by epigenetic regulatory mechanisms in the nucleus. In multicellular organisms, however, many additional processes are involved in learning and memory, such as the establishment of synaptic connections between neurons and the specialization of neural sub-networks for controlling specific behaviors.

Mind appears as a new top-level regulator of organism functions, but it does not replace already existing hardwired protosemiotic networks. Many low-level functions do not require complex regulation; they are well controlled by direct signaling, and replacing them with a learning mechanism would be costly and inefficient. However, some hard-programmed processes like embryo development may acquire partial guidance from the minds of individual cells or from the brain. Neurons establish functional feedback regulation of growing organs, where non-functional cells or cell parts (e.g., synapses) are eliminated (Edelman, 1988). In other words, cells attempt to find a ‘job’ in the body that fits to an available functional niche and the cell’s pre-history. If a job is not found, then the cell goes into apoptosis.

7. The first object classified by minimal mind was the body

The initial task of mind was to classify those objects that are most important for the life of an organism. Because an agent’s body is most intimately linked with a large number of functions, we can hypothesize that the body was the first object to be classified by mind. The purpose of classifying body states is to assign priorities to various functions, such as the search for food, defense from enemies, and reproduction. Functions of protosemiotic agents are directly controlled by internal and external signs and therefore priorities are fixed by a heritable signaling network. In contrast, agents with mind can learn to distinguish body states and adjust the priority of functions based on previous experience.

Of the two components of mind, Innenwelt (classifications and models of self) and Umwelt (classifications and models of external objects), Innenwelt is primary and Umwelt is secondary. Simple agents do not distinguish between internal and external sensations. It requires additional complexity for agents to realize that there are external objects beyond signals that come from receptors. The main difference between ‘internal’ and ‘external’ worlds is a higher predictability of the internal world and a lower predictability of the external world. Thus, it is reasonable to presume that Umwelt emerged as a less predictable portion of a former Innenwelt. This evolutionary approach to the differentiation of ‘external’ from ‘internal’ is profoundly different from cybernetics, where the boundary between the system and environment is defined a priori.

The capacity of mind to classify and model objects is closely related to the ability of agents to track objects. In particular, agents can rely on the assumption that objects keep their properties over time. For example, a predator that is chasing an object identified previously as prey does not need to repeat identification over and over again. Similarly, modeling appears most beneficial if the agent keeps track of the predicted object. Thus, tracking of objects by agents augments the utility of classification and modeling. The advantage of body as the first classified and modeled object is that it is always accessible, and thus, agents do not need additional skills for object tracking.

8. Modeling functions of mind

Modeling, which can be defined as prediction or anticipation of something unperceived, is the second major function of mind after the classification of objects. Elements of modeling are present in any classification, because ideal objects are already models. Recognition of an object is based on the anticipated combination of traits, as follows from the extensively explored area of image recognition. Some of these models are fixed whereas others include parameters that are adjusted to increase the likelihood of a match between the model and sensorial data (Perlovsky, Deming, & Ilin, 2011). For example, distance to the object can be used as a parameter which affects the size and resolution of the image as well as its position relative to other objects. These simple models belong to the primary modeling system, where ideal objects are not connected and therefore not used for prediction or anticipation of something different than what is perceived. Some of them are pure sensations, and others are integral sensation-actions. As an example of sensation-action, consider a moth that by instinct starts laying eggs after recognizing its host plant.

Advanced models that establish relationships between ideal objects belong to the secondary modeling system (Sebeok, 1987). For example, if a bird attempts to eat a wasp and gets stung, then it connects the ideal object of a wasp with pain. As a result, this bird will not attempt to eat anything that looks like a wasp because the image of a wasp reminds it of pain. It was suggested that the secondary modeling system is handled by the interpretive component of the brain, whereas cybernetic and instinctive components handle the primary modeling (Barbieri, 2011). The secondary modeling system establishes links between various ideal objects, and therefore allows agents to develop flexible relationships between signs and functions. The origin of the secondary modeling system can be associated with the emergence of powerful sense organs that provided animals with more information than was needed for immediate functions. As a result, the classification of objects became more detailed and partially independent from their utility. Using a combination of a large number of traits, animals are able to recognize individual objects, associate them with each other, and make a mental map of their living space. Individual objects are then united into functionally relevant classes. Animals also can use abstract ideal objects that correspond to individual traits (e.g., color, shape, or weight) of real objects. Dynamic models associate the current state of an object with future states of the same object. They are used by predators to predict the movement of their prey. Association models predict the presence of one object from the observation of another kind of object. For example, animals associate smoke with forest fires and attempt to escape to a safe location.

One of the recent approaches to model-building is dynamic logic (Perlovsky et al., 2011). The idea is to maximize the likelihood of matching between the set of models with adjustable parameters to the set of empirical data. Each model corresponds to a potential object, which can be added or deleted in the process of optimization. The accuracy of comparison between object-models increases and model parameters are adjusted as optimization progresses. This approach explains two important aspects of modeling. First, detection of objects is not possible without models because models specify what we are looking for. And second, objects can be measured using optimal parameters of object-models (although this is not the only way to measure objects). Because the data are referenced by space and time, models include motion equations and yield plausible trajectories of object-models. However, all object-models identified with this method are primary ideal objects (i.e., they belong to the primary modeling system). Connections between primary objects have to be established at a higher level of the hierarchy of objects (Perlovsky et al., 2011).

Models are the main subject of Peirce's semiotics, where the perceived object is a sign vehicle that brings into attention the interpretant, or associated ideal object. The primary modeling system operates with icons, which are associated with isolated ideal objects (sensations or sensation-actions), whereas the secondary modeling system also includes indexes which are the links between ideal objects (Sebeok & Danesi, 2000). Peirce, however, viewed sign relationships as components of the world rather than models developed by agents. He believed that models were embedded in the world. The danger of this philosophy (i.e., objective idealism) is that it easily leads to dogmatism as models become overly trusted. But how can we evaluate the relationship between a model and reality? Models can be used in two ways: they can be trusted and they can be tested. When a bird does not attempt to catch wasps after being stung, it trusts the model of a wasp. However, not all models generate reproducible results, and therefore models need to be tested and modified if necessary.

9. Testing models

Model testing is one of the most important activities in science, and it has direct implications for epistemology (Cariani, 2011; Popper, 1999; Rosen, 1991; Turchin, 1977). Animals also test models, but they do not run experiments for the sake of testing hypotheses as humans do. Instead, they evaluate the success rates of their behavioral strategies and establish preferences for more successful behaviors. In this way, predators learn how to chase and capture prey, and birds learn how to attract the attention of predators away from their nests.

Model testing is a complex procedure that determines if predictions generated by the model match the real world. In the simplest case, an agent measures the initial state of the object, and the obtained results are used as input for the model. Then the output of the model is compared to the measurement of the final state of the object, and if they match, the test is considered successful (Cariani, 2011; Rosen, 1991; Turchin, 1977). To formalize model testing, we need to generalize our terms. First, the expression "initial state of the object" implies that agents have a method for tracking objects. In particular, each object O is associated with the final object $G(O)$, where G is the tracking function. Second, objects are characterized either quantitatively by measurements or qualitatively by the identification of individual features or by classifying whole objects. In result, each object O becomes associated with some ideal object $M(O)$ in mind, which is interpreted as a measurement of that object. In general, agents use multiple measurement methods M_1, M_2, ... $M_{n,}$ which are applicable in different situations. Similarly, in science, we use different measurement devices and sensors to characterize objects. Finally, the model is a map, F, between ideal objects in mind. For example, a dynamic model associates initial measurements of an object with measurements of its final state. Then successful model testing can be represented by a commuting diagram (Fig. 1), where measurement of the final state of the object, $M_2(G(O))$, matches to the model output from the measurement of the initial state of the object used as input, $F(M_1(O))$. Two measurement methods M_1 and M_2 may be the same, but in the general case they are different. If the equation $M_2(G(O)) = F(M_1(O))$ is true for all available objects, then the model F is universal relative to measurement methods M_1 and M_2 and tracking method G.

$$\begin{tikzcd}
M_1(O) \arrow[r, "F"] & F(M_1(O)) = M_2(G(O)) \\
O \arrow[u, "M_1"'] \arrow[r, "G"] & G(O) \arrow[u, "M_2"']
\end{tikzcd}$$

Figure 1. Commuting diagram of model testing. M_2 and M_2 are measurement methods for the initial object O, and final object, $G(O)$, respectively; G is the object tracking function, and F is the map between ideal objects in the model.

Commuting diagrams, similar to Fig. 1, were proposed previously (Cariani, 2011), but function G was interpreted as objective natural dynamics of the world. In contrast, I associate function G with an agent's ability to track or manipulate objects. An example of non-trivial object tracking is the association of the "morning star" with the "evening star" (i.e., planet Venus) on the basis of the model of planetary movement. This example illustrates that all four components of the model relation (F, G, M_1, M_2) are interdependent epistemic tools, and one component may help us to improve another component.

Cariani suggested that the manipulation of objects is the reverse of measurement, and therefore he changed the direction of function M_1 (Cariani, 2011). This approach, however, implies that real objects are created from ideal objects without any matter. In contrast, I suggest associating the manipulation of objects with various tracking functions G. Some G-functions may represent a passive experiment, where objects are mapped to their natural future state, whereas other G-functions represent active experiments where objects are mapped into their products after specific manipulations. If we want to construct meta-models that describe multiple methods of object manipulation, then each method i should be linked with a corresponding model F_i and object tracking method G_i.

Commuting diagrams of model testing capture a very important aspect of epistemology: the equivalence is achieved in the domain of ideal objects rather than in the domain of real objects.

Thus, different models may equally well capture the same process or relationship in the real world. The second conclusion is that models are always tested together with measurement methods and tracking methods, which are usually ignored in physics. As a result, agents from one communication system cannot take advantage of models developed within another communication system if measurement methods and tracking methods do not match.

According to the critical rationalism of Popper, a model, whose predictions are wrong, should be removed from the domain of science (Popper, 1999). However, this rarely happens; instead, model components (F, G, M_1, M_2) are adjusted to make the diagram in Fig. 1 commuting. Popper condemned this practice because it makes hypotheses non-falsifiable. However, Popper's argument does not make sense from the evolutionary point of view. If animals rejected any model that once had generated a wrong result, then they would soon run out of models and fail to perform their functions. Any model is a product of evolution and learning and integrates long-term experience of agents. It is better to have a non-accurate or non-universal model than no model at all. This explains why models are so persistent both in biological evolution and in human culture.

10. Model transfer between individuals

Most models used by animals are not communicated to other individuals. Thus, each animal has to develop its own models based on trial and error as well as heritable predispositions. However, social interactions may facilitate the development of models in young animals. For example, animals may copy the behavior of their parents and eventually acquire their models in a faster way than by pure trial and error. However, efficient communication of models is possible only by language, which corresponds to the cultural level of semiosis, following the terminology of Kull (Kull, 2009). In language, signs do not only correspond to ideal objects, they also replicate the structure of relationships between ideal objects in the model. Thus, language itself becomes the modeling environment called the tertiary modeling system (Sebeok & Danesi, 2000). Language is based on symbols which are signs whose meanings are established by convention within the communication system. Then, a message with two (or more) interconnected symbols is interpreted as a link between corresponding ideal objects within the model. Thus, the tertiary modeling system is based on symbols (Sebeok & Danesi, 2000).

In conclusion, minimal mind is a tool used by agents to classify and model the objects. Classification ends up at the ideal object, which serves as a checkpoint to initiate certain physical or mental functions. Mind is projected to appear within eukaryotic cells with well-developed epigenetic regulation because these mechanisms can convert DNA segments into standard information-processing modules with multiple attractor domains and flexible control. Classification and modeling of objects had started from the body of agent and then expanded to external objects. Modeling functions of mind progressed from primary models that simply support classification of objects to secondary models that interconnect ideal objects, and finally, to tertiary models that can be communicated to other agents.

References

Armus, H. L., Montgomery, A. R., & Gurney, R. L. (2006). Discrimination learning and extinction in paramecia (P. caudatum). *Psychol Rep, 98*(3), 705-711.

Barbieri, M. (2003). *The organic codes: an introduction to semantic biology*. Cambridge, UK. ; New York: Cambridge University Press.

Barbieri, M. (2008). Biosemiotics: a new understanding of life. *Die Naturwissenschaften, 95*(7), 577-599.

Barbieri, M. (2011). Origin and evolution of the brain. *Biosemiotics, 4*(3), 369-399.

Baslow, M. H. (2011). Biosemiosis and the cellular basis of mind. How the oxidation of glucose by individual neurons in brain
results in meaningful communications and in the emergence of “mind”. *Biosemiotics, 4*(1), 39-53.

Cariani, P. (1998). Towards an evolutionary semiotics: the emergence of new sign-functions in organisms and devices. In G. V. d. Vijver, S. Salthe & M. Delpos (Eds.), *Evolutionary systems* (pp. 359-377). Dordrecht, Holland: Kluwer.

Cariani, P. (2011). The semiotics of cybernetic percept-action systems. *International Journal of Signs and Semiotic Systems, 1*(1), 1-17.

Deacon, T. W. (2011). *Incomplete nature: How mind emerged from matter*. New York: W. W. Norton and Company.

Dennett, D. C. (1995). *Darwin's dangerous idea: evolution and the meanings of life*. New York: Simon & Schuster.

Eco, U. (1976). *A theory of semiotics*. Bloomington: Indiana University Press.

Edelman, G. M. (1988). *Topobiology: An introduction to molecular embryology*. New York: Basic Books.

Emmeche, C., & Hoffmeyer, J. (1991). From language to nature - the semiotic metaphor in biology. *Semiotica, 84*(1/2), 1-42.

Ginsburg, S., & Jablonka, E. (2009). Epigenetic learning in non-neural organisms. *J Biosci, 34*(4), 633-646.

Griffin, D. R. (1992). *Animal minds*. Chicago: University of Chicago Press.

Hoffmeyer, J. (1997). Biosemiotics: Towards a new synthesis in biology. *European Journal for Semiotic Studies, 9*(2), 355-376.

Hoffmeyer, J. (2010). Semiotics of nature. In P. Cobley (Ed.), *The Routledge Companion to Semiotics* (pp. 29-42). London, New York: Routledge.

Jeanteur, P. (2005). *Epigenetics and chromatin*. Berlin: Springer.

Kolmogorov, A. N. (1965). Three approaches to the quantitative definition of information. *Problems of Information Transmission, 1*(1), 1-7.

Krampen, M. (1981). Phytosemiotics. *Semiotica, 36*(3/4), 187-209.

Kull, K. (2009). Vegetative, animal, and cultural semiosis: the semiotic threshold zones. *Cognitive Semiotics, 4*, 8-27.

Levenson, J. M., & Sweatt, J. D. (2005). Epigenetic mechanisms in memory formation. *Nat Rev Neurosci, 6*(2), 108-118.

Lovelock, J. E. (1979). Gaia. A new look at life on Earth. Oxford: Oxford University Press.

Markoš, A., & Švorcová, J. (2009). Recorded versus organic memory: Interaction of two worlds as demonstrated by the chromatin dynamics. *Biosemiotics, 2*(2), 131–149.

Millau, J. F., & Gaudreau, L. (2011). CTCF, cohesin, and histone variants: connecting the genome. *Biochem Cell Biol, 89*(5), 505-513.

Miller, C. A., & Sweatt, J. D. (2007). Covalent modification of DNA regulates memory formation. *Neuron, 53*(6), 857-869.

Nillson, N. J. (1998). *Artificial Intelligence: A New Synthesis*. San Francisco: Morgan Kaufmann Publishers.

Peirce, C. S. (1998). *The Essential Peirce: Selected Philosophical Writings* (Vol. 2). Indiana: Indiana University Press.

Perlovsky, L., Deming, R., & Ilin, R. (2011). *Emotional cognitive neural algorithms with engineering applications. Dynamic logic: from vague to crisp.* (Vol. 371). Warsaw: Polish Academy of Sciences.

Popper, K. (1999). *All life is problem solving*. London: Routledge.

Premack, D. G., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*, 515-526.

Prodi, G. (1988). Material bases of signification. *Semiotica, 69*(3/4), 191-241.

Putnam, H. (1975). *Mind, Language and Reality* (Vol. 2). Cambridge: Cambridhe University Press.

Rashevsky, N. (1938). *Mathematical Biophysics*. Chicago: University of Chicago Press.

Rosen, R. (1970). *Dynamical system theory in biology*. New York: Wiley-Interscience.

Rosen, R. (1991). *Life itself : a comprehensive inquiry into the nature, origin, and fabrication of life*. New York: Columbia University Press.

Sebeok, T. (1987). Language: How primary a modeling system? In J. Deely (Ed.), *Semiotics 1987* (pp. 15-27). Lanham, MD: University Press of America.

Sebeok, T. A. (1972). *Perspectives in zoosemiotics*. The Hague,: Mouton.

Sebeok, T. A., & Danesi, M. (2000). *The forms of meaning. Modeling systems theory and semiotic analysis*. New York: Mouton de Gruyter.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal, 27*, 379–423, 623–656.

Sharov, A. (2010). Functional information: Towards synthesis of biosemiotics and cybernetics. *Entropy, 12*(5), 1050-1070.

Sharov, A. A. (2012). The origin of mind. In T. Maran, K. Lindström, R. Magnus, & M. Tønnensen (Eds.), Semiotics in the wild (pp. 63–69). Tartu: University of Tartu.

Sharov, A. A. (2006). Genome increase as a clock for the origin and evolution of life. *Biol. Direct, 1*, 17.

Sharov, A. A. (2009a). Coenzyme autocatalytic network on the surface of oil microspheres as a model for the origin of life. *Int J Mol Sci, 10*(4), 1838-1852.

Sharov, A. A. (2009b). Genetic gradualism and the extraterrestrial origin of life. *Journal of Cosmology, 5*, 833-842.

Sharov, A. A. (2009c). Role of utility and inference in the evolution of functional information. *Biosemiotics, 2*(1), 101-115.

Swan, L. S., & Goldberg, L. J. (2010). How is meaning grounded in the organism? *Biosemiotics, 3*(2), 131-146.

Swan, L. S., & Howard, J. (2012). Digital Immortality: Self or 01001001? *International Journal of Machine Consciousness, 3*(2), in press.

Turchin, V. F. (1977). *The phenomenon of science*. New York: Columbia University Press.

Turing, A. (1952). Can automatic calculating machines be said to think? In B. J. Copeland (Ed.), *The essential Turing: The ideas that gave birth to the computer age* (pp. 487-506). Oxford: Oxford University Press.

Uexküll, J. v. (1982). The theory of meaning. *Semiotica, 42*(1), 25-82.
Visel, A., Blow, M. J., Li, Z., Zhang, T., Akiyama, J. A., Holt, A., et al. (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature, 457*(7231), 854-858.
Waddington, C. H. (1968). Towards a theoretical biology. *Nature, 218*(5141), 525-527.
Wood, D. C. (1992). Learning and adaptive plasticity in unicellular organisms. In L. R. Squire (Ed.), *Encyclopedia of learning and memory* (pp. 623–624). New York: Macmillan.