The Journal of Mind and Behavior
Winter 2021, Volume 42, Number 1
Pages 33–52
ISSN 0271–0137

Cognitive Science Models: An Aristotelian–Thomistic Appraisal

James M. Stedman

University of Texas Health Science Center San Antonio

Cognitive science models emerged in the 1950s with the advent of computers and today have three versions: classical, connectionist, and embodiment (see Dawson, 2013 for an extensive review). Classical and connectionist models focus on the brain as explanatory of all cognitive phenomena but go further to claim that these cognitive phenomena can be replicated by computers, a field now labeled Artificial Intelligence. This essay will explore the claims of the classical and connectionist models and their strengths and weaknesses. I introduce the classical realism of Aristotle and Aquinas (the A–T model) and argue that the A–T model can incorporate classical and connectionist theory and findings, solve most of the weaknesses of both models, and add better causative explanations of higher cognitive phenomena, such as concept formation and thought. The presentation ends with a brief discussion of implications for cognitive psychology.

Keywords: Aristotle, cognitive science, cognitive psychology

The roots of cognitive science stretch back to the post Cartesian idea that it might be possible to construct machines capable of human intellectual behavior; and the advent of computers in the 1950s solidified this conviction: machines could replicate human cognition. Over time, this approach became known as Classical Cognitive Science and was founded on the view that cognition, both human and animal, is computation; computation that can be produced by computers. Critics of this classical approach emerged to propose a theoretical model of cognitive science known as Connectionist Cognitive Science. This model continued to assert that cognition is a computational form of information processing but claimed that cognition is best explained by brain-based neuronal networks.

The author would like to acknowledge Thomas L. Spalding, Ph.D., University of Alberta and Matthew Kostelecky, Ph.D., St Joseph's College at the University of Alberta, for their review of and contributions to this manuscript. Correspondence concerning this article should be addressed to James M. Stedman, Ph.D., Department of Psychiatry, University of Texas Health Science Center, 7703 Floyd Curl Drive, San Antonio, Texas 78229–7792. Email: Stedman@uthscsa.edu

However, the connectionist model continued to claim that computers could model the brain. Most recently, a third model, known as Embodied Cognitive Science, has emerged. This model has abandoned attempts to explain cognition either by the AI model or the connectionist model and asserts that cognition is best understood as an agent's material body sensing and acting in the world (see Dawson, 2013 for an extensive review). However, the embodied model still sees cognition as computational, and was developed via work in cybernetics and robotics. One of the key characteristics of computational accounts is the ability to abstract away from the physical nature of the device doing the computing (and hence, both classic AI and connectionist models are often implemented on standard digital computers, for example).

Cognitive science, though multidisciplinary, served as the foundation of cognitive psychology. Recent texts (Dawson, 2013; Goldstein, 2019; Solso, MacLin, and MacLin, 2014) recognize this fact, although they also show that contemporary cognitive psychology incorporates other models (e.g., Gestalt, neuroscience, social). This essay will focus specifically on the classical and connectionist cognitive science models and consider four goals: (a) presentation of the major claims of the classical and connectionist models; (b) consideration of the strengths and weaknesses of each model; (c) an examination of the Aristotelian–Thomistic (A–T) philosophical position, known as Classical Realism, in relation to the classical and connectionist models; and (d) a brief consideration of the implications for cognitive psychology.

The Classical Model

Dawson (2013), in his extensive review, traces the metaphysical underpinnings of the classical model to Descartes. He points out that cognitive science rejected the mind–substance side of Cartesian dualism but committed fully to the material–substance position. Hence, cognitive science (and cognitive psychology as its derivative) positions itself as a materialist monism and, therefore, must present causal explanations of all cognitive events, including sensation, perception, memory, and higher-order cognitive functions (thinking, reasoning, decision-making, reflecting, intentionality, and the like) based on materialist metaphysics. Brains, human and animal, are material and are obvious candidates for a materialist, mechanistic explanation. However, classical cognitive science claimed something more radical, namely that machines, in and of themselves, could replicate all human and animal cognitive events. Serious pursuit of this claim began as early as 1950 with the advent of the first computers, and early advocates asserted computers would soon demonstrate the ability to replicate all human cognition. In fact, the label "Artificial Intelligence" was coined by John McCarthy and became standard language after the now famous Summer Research Project on Artificial Intelligence held at Dartmouth in 1956 (see McCarthy, Minsky, Shannon, and

Rochester, 2006). Certainly, the accomplishments of AI from then to now are strikingly impressive.

In 1956, psychology was still locked in behaviorism and not interested in revisiting "the mind." A recent cognitive psychology textbook (Goldstein, 2019) points out that the "cognitive revolution" in psychology was slow to develop and cites the publication of the first book devoted to cognitive psychology (Neisser, 1967) as the recognized starting point. Neisser coined the term cognitive psychology and adopted the computational information processing model as the best model for explaining and studying the human and animal mind. The classical information-processing computational model has become vastly complex (see Dawson, 2013, Chapter 3 for a detailed exposition) but retains its basic composition as described by Goldstein (2019): "Information is first received by an 'input processor.' It is then stored in a 'memory unit' before it is processed by an 'arithmetic unit,' which then creates the computer's output" (p. 12). Though this description of a computer remains true, Dawson (2013) points out that cognitive science saw the digital computer at a deeper level, as the grounding for a mechanistic explanation of human and animal cognitive events.

The classical model emerged first and is described by Dawson (2013) as follows:

> The claim that cognition is computation, put in its modern form, is identical to the claim that cognition is information processing. Furthermore, classical cognitive science views such information processing in a particular way: it is identical to that carried out by a physical symbol system, a device like a modern digital computer. As a result, classical cognitive science adopts the representational theory of mind. It assumes the mind contains internal representations (i.e., symbolic expressions) that are in turn manipulated by rules or processes that are part of a mental logic or a (programming) language of thought. Further to this, a control mechanism must be proposed to explain how the cognitive system chooses what operation to carry out. (p. 122)

Two elements of this description of the classical model require unpacking, including: (a) what constitutes an internal representation (as a symbolic expression) and (b) the nature of the control mechanism explaining choice.

Paul Thagard (2019) points out that mental representations are "analogous to computer data structures" and are composed of "logical propositions, rules, concepts, images and analogies" (p. 7) which are manipulated by computational information processing systems. Hence, the mental representation, at its most basic level, is a pattern of on/off (0s and 1s) electrical circuits, described by Dawson (2013) as symbolic structures having semantic content (expressions of meaning); a relationship to the external world, known as designation (Newell, 1980); and a capacity to generate output. This sequence, as it applies to higher-order cognition, has become known as the sense–think–act cycle (Pfeifer and Scheier, 1999). Philosophers of mind (Dennett, 1978; Fodor, 1968) defend this conceptualization of the mental representation as composed of increasingly simple sub-functions, finally decomposable into the yes/no functions of electrical circuits and, therefore, replicable by a machine. Dennett (1978) has described this process thusly: "One discharges fancy homunculi from one's scheme by organizing armies of such idiots to do the work" (p. 124).

What of the control mechanism explaining choice? Russell and Norvig (2009) describe AI generally as a field devoted to building intelligent agents. At its most basic level, the agent receives environmental input via sensors, then "chooses" action as dictated by the programmed condition–action rules, then produces an action which is delivered to the environment. This theoretical model, developed by Brooks (1991), consists of five levels, with each level adding more internal processes (Critic, Learning Element, Problem Generator, Performance Element, Communication) until reaching human capacity. At levels II and III computers can "think" ahead to future steps in proscribed environments that include clearly defined rules (e.g., win chess competitions) and, if loaded with declarative knowledge, can perform deductions to reach environmental goals (win at *Jeopardy*). Computers operate according to condition–action rules (choice rules) provided by external programming; however, more current AI involves computers programmed with simple learning rules which come to rely on programming not directly built in by programmers.

Strengths and Weaknesses of the Classical Model

A major, and perhaps obvious, strength of the classical model is that it broke the grip behaviorism had on psychology and reintroduced "the black box" into theory and research. A few of these advances in cognitive psychology are the following:

1. Perception and attention: perceptual scan (Cattell, 1986); iconic storage (Neisser, 1967); sensation and perception (Broadbent, 1958); visual attention and perception (Treisman, 1988).

2. Memory: long-term memory models (Shiffrin and Atkinson, 1969); working memory (Baddley, 1986); episodic and semantic memory (Tulving and Donaldson, 1972).

3. Higher-order cognition: ability to "plan ahead" as in chess (Bringsjord and Govindarajulu, 2020,p. 4) or answer informational questions as in *Jeopardy* (Strzalkowski and Harabagiu, 2006); development of planning algorithms (Bringsjord and Govindarajulu, 2020, p. 8); development of analogy solving algorithms (Bringsjord, 2011); higher-order cognitive models of mind, such as Theory–Theory (Gopnik and Meltzoff, 1997).

Over time, the various disciplines came to recognize difficulties with the basic claim of the classical model, namely that all human and animal cognitive phenomena could be produced by computers. Much of this criticism originated from philosophers but also came from experimental results pointing toward issues with the model. Some of the more prominent difficulties, stated briefly, are as follows:

1. Intentionality. The symbols incorporated in computer information processing must relate to the external world. This is known in the classical model as the problem of representation (Cummings, 1989) or the symbol grounding problem (Harnad, 1990). Dawson (2013) points out that the classical model has not solved intentionality but he does assert that a solution exists: "The physical symbol system hypothesis does not propose a solution, but merely assumes that such a solution exists. This assumption is plausible to the extent that computers serve as existence proofs that designation is possible" (p. 8). This argument appears to claim that the success of computer replication of human cognitive events proves that there must be a solution. Many critics of the classical model see this argument as simply begging the question.

2. The logicism hypothesis. The classical model claims that thinking occurs by performing logical operations, operations that are identical to information processing by computers. However, a number of studies have shown that humans often do not think according to logical principles (Hastie, 2000). Hence, the issue is, if human thinking is identical to logical operations, how would humans ever think illogically? Of course, there have been attempts to solve this problem but not one has fully succeeded.

3. Computer vs. brain. Connectionist critics point out that the machine claim of the classical model must eventually match mechanisms in the human (and animal) brain. Connectionists assert that the match fails. For example, computers have few very fast components involved in information processing; brains have many (neurons) that process at much slower speeds (von Neumann, 1958).

4. Semantics (meaning) and qualia. Classical theory claims that the symbols manipulated by computers also contain meaning. As seen above, some philosophers of mind (Dennett, 1978) claim that meaning is assembled from primitive elements, themselves having no meaning. Many have challenged this claim (Robinson, 2008; Searle, 1984). Additionally, property dualists, such as Chalmers (1996), have asserted that qualia, defined as subjective experiences unique to each individual, cannot emerge from either a machine or a brain.

5. The Chinese Room. Proof of the central claim of the classical theory has been divided into weak and strong. The weak claim, devised by Turing (1950), is that the central assertion of the classical model could be established if a human listener cannot distinguish a message presented by another human versus a computer. This proof, long accepted by many in the AI community, has been criticized as too weak (Searle, 1997). In their recent review of AI, Bringsjord and Govindarajulu (2020) characterized the strong equivalence approach thusly: "'Strong" AI seeks to create artificial persons that have all the mental powers we have, including phenomenal consciousness" (p. 35). Searle (1984) opposed both strong and weak equivalence in a thought experiment known famously as the Chinese Room argument. Briefly (and the reader is encouraged to examine this argument in full), Searle's argument is as follows: fluent Chinese speakers create meaningful messages in Chinese that

are then entered into a box. In the box is a non-Chinese speaker who does not understand the messages at all but who has a codebook showing how the messages should be outputted to Chinese speakers outside the box who will receive the messages and understand their meaning. According to Searle, this is all that a computer does. This argument has never been successfully defeated.

Dawson (2013) states that the classical model is still preeminent in cognitive science and this claim is no doubt true considering all of the disciplines involved in cognitive theory and research, particularly applied to AI. However, because of the critiques cited above, many occurring quite early in the history of the classical model, the connectionist model arose as an alternative.

The Connectionist Model

Buckner and Garson (2019) describe the connectionist model as follows:

> Connectionism is a movement in cognitive science that hopes to explain intellectual abilities using artificial neural networks (also known as "neural networks" or "neural nets"). Neural networks are simplified models of the brain composed of large numbers of units (the analogs of neurons) together with weights that measure the strength of connections between these units. These weights simulate the effect of the synapses that link one neuron to another. Experiments on models of this kind have demonstrated an ability to learn such skills as face recognition, reading, and the detection of simple grammatical structure. (p. 1)

They point out that these networks consist of many units joined together to produce patterns of connections. These include input units receiving information to be processed; output units which generate actions/results; and hidden units which occur between the input and output units. As related to the brain analogy, Buckner and Gerson (2019) say this, "If a neural net were to model the whole human nervous system, the input units would be analogous to the sensory neurons, the output units to the motor neurons, and the hidden units to all other neurons" (p. 1).

The weights mentioned above occur first at the input level and have activation values related to features of the environment. These are essentially numerous weighted stimuli transmitted to the hidden units (analogous to neuronal structures). The hidden unit generates its own activation weight as a signal that passes on to other hidden units or to output units. Activation patterns in the neural network relate directly to the weighted strength of connections between all units. The connectionist model attempts to use computers to simulate the brain by building computer-based "neural networks" that can learn.

For example, using a procedure called backpropagation, computers can be trained to distinguish male and female faces. The input training set would have many faces of males and females broken down into many variables at the pixel level, such as shape and other male–female facial characteristics. The output unit consists of two categories, male and female faces. Input units are compared to output facial categories and the input unit's weights are adjusted until features matching the female are increased and the male features decreased. This procedure continues until correct category selection occurs at the output unit. This process generally takes numerous rounds (hundreds) of input weight adjustment learning. This learning model has been used for many tasks which connectionists believe to be models of human cognition, including nets that read English text (Sejnowski and Rosenberg, 1987), handwriting recognition (Hofstadter and McGraw. 1995), as well as nets that recognize human voices and can type language, and such.

The hidden units of computer simulated neural networks consist of layers and require a model to explain interactions between hidden units at more superficial levels and those at deeper layers, similar to neuronal structures in the brain. Technical advances have led to the newest iteration of connectionism, known as deep learning, in which learning occurs with many layers of nodes (hidden units) between input and output. Development of the graphic processing unit (GPU) has allowed large numbers of parallel processors with the computational power to train large networks (as many as five to several hundred). As stated by Buckner and Garson (2019), "The key is that patterns detected at a given layer may be used by the subsequent layers to repeatedly create more and more complex discriminations" (p. 19). These advances have also improved computer simulation of learning to cope with "nuisance parameters," variables which interfere with correct sorting decisions. This results in a system's ability to detect "hidden" similarities at deeper levels. Hence, deep learning has produced models appearing closer to human brain function.

As with the classical model, the connectionist model must account for representation and meaning. Regarding representations, the connectionist model posits that each representation is "a pattern of activity across all units" with "no principled way to distinguish between simple and complex representations" (Buckner and Garson, 2019, p. 9). The units here are the hidden units that cluster together as the result of leaning, analogous to brain-based increases in synaptic densities that occur during human and animal learning. This process can be simulated, to a degree, by deep learning mechanisms. Hence, like the classical model, the connectionist representation is a material entity, is manipulated according to information processing, and is not based on symbols but on clusters of hidden units.

Buckner and Garson (2019) state the following about how clustered brain states (the representations described above) could have meaning: "The idea is that similarities and differences between activation patterns along different dimensions of neural activity record semantic information. So the similarity properties of neural activations provide intrinsic properties that determine meaning" (p. 14). Representations, which carry these activation patterns, presumably derived from external environmental sources and assembled in the hidden units based on similarity, is the connectionist theory of meaning.

Strengths and Weaknesses of the Connectionist Model

The connectionist model offers several advantages over the classical model. Some of these strengths are as follows:

1. The connectionist model manages the intentionality/symbol grounding problem of the classical model by incorporating, at least theoretically, the human (and animal) sensory and CNS biological structure. Following Locke, the connectionist model asserts that cognition begins in the senses and the senses are directly connected to the externally experienced world. Of course, the connectionist model claims that the biological model can be replicated by computer information processing that simulates brain-based neural networks and that claim is subject to challenge.

2. Being grounded, at least in theory, in the sensory–brain mechanism, the connectionist model can incorporate some findings from neuroscience. In fact, connectionists often cite neuroscience theory and findings as support for their model.

3. As mentioned above, connectionists solve both the speed of processing and the complexity of organization issues plaguing the classical model by asserting the neural network theory which connectionists view as grounded in the 100 billion neurons and two trillion synaptic interconnections of the "hidden units" of the human brain.

4. Formulations of higher-order cognition begin with concept formation, the basic building block for thinking, reasoning, planning, etc. Rogers and McClelland (2004) present an elaborate model depicting composition of the "concept of canary." There are four levels: concept (*canary* as opposed to *rose*, *sunfish*, etc.); representation (the activation of units with varying connection weights described above) and of relation units (units indicating relationship such as *is*, *is a*, *can*, etc.); hidden units which further process the concept *canary*; and finally properties (consisting of abstract and highly abstract descriptors such as *living thing*, *plant*, *pretty*, *yellow*, *wings*, *roots*, etc. and their relations). This sort of structure, through a series of weighted interconnections, produces the concept of *canary* (as opposed to *oak tree* or *bear*) and is learned. Goldstein (2019) describes this learning as follows: "The answer to 'a canary is a …' is represented in the network by activation of the property units plus the pattern of activation of the network's representations and hidden units. However, according to connectionism, a connectionist network has to be trained in order for a result to occur. This training involves adjusting the network's connection weights" (p. 262). He goes on to describe how this might occur in young children. Once connectionism establishes an explanation for the formation of concepts, higher-order cognition (reasoning, planning, deciding, thinking, creativity and so forth) flows from use of concepts.

Despite the advantages noted above, connectionism also faces serious challenges, some of which are as follows:

1. Critics challenge both the classical and connectionist models regarding consciousness and semantics. The issue regarding consciousness involves explaining how something entirely material, such as computers and the brain, can be conscious. A number of philosophers of mind challenge all or parts of this claim (Chalmers, 1996; Dreyfus, 1992; Feser, 2019, p. 442; Robinson, 2008; Searle, 1984, 2014) and some defend the position that material reality can appear conscious but really is not. This position is known as eliminative materialism (Churchland, 2007; Dennett, 1979).

2. Connectionist critics point out that backpropagation learning methods require very extensive repetitions to achieve weight adjustment in the network, whereas humans can learn from a single example. For instance, children can learn the name of a two-wheeled vehicle in one or two trials (Lake, Wojciech, Fergus, and Todd, 2015).

3. Deep learning has not solved the problem of constructing new abstractions above the level of input vocabulary (Bringsjord and Govindarajulu, 2020, p.13); the problem of generalization to inputs from outside the original training set (Zhang, Bengio, Hardt, and Recht, 2016); or the problem of supplying convincing explanations of the mechanisms of deep learning related to "adversarial examples" which can "fool" other nets trained in the same manner (Goodfellow, Pouger–Abadie, Mirza, Xu, Wade–Farley, Ozair, Courville, and Bengio, 2014).

4. There have been challenges regarding the similarity between the connectionist network model and the actual biological CNS, and Shimansky (2009) has argued that there is little neuroscience evidence to support important connectionist learning rules.

5. Despite connectionism's claim to solve the problem of grounding/intentionality, the model offers no direct relationship between the senses and CNS. In fact, inputs are arranged in a structured way by the person building the model. So, already, the model is working on "interpreted representations" coming from the programmer rather than raw sensory data from the environment.

In sum, the classical and connectionist models share a commitment to a materialist metaphysics and a general commitment to the view that information processing underpins human and animal cognition. Both acknowledge that, ultimately, cognitive processing begins in the senses, is processed in the brain, and finishes with various forms of output (the sense–think–act sequence). Both models have strengths and weakness, some shared. However, the classical model appears more committed to demonstrating that, though the brain must be factored in at some point, cognition is entirely replicable by machine. The connectionist model is more committed to showing that computers can operate like the brain, but the brain is central. Computers merely replicate the brain's mechanisms.

An Aristotelian–Thomistic (A–T) Appraisal

Dawson (2013) traces the metaphysical roots of classical cognitive science to the substance dualism of Descartes and points out that the formulators of the classical model rejected the rationalist, immaterial position and accepted the materialist position. He traces connectionist cognitive science to Locke's empiricism, which was a reaction to Descartes' rationalism. I propose the classical realism of Aristotle and Aquinas as a metaphysical alternative to the classical and connectionist models. However, before discussing the A–T cognitive model, we must start with the basics of A–T metaphysics because these principles underpin the A–T account of cognition. These principles are the hylomorphic theory, the theory of act and potency, and the A–T view of substance.

Aristotle's general metaphysical theory claims that all reality can be explained by means of the four causes: efficient, material, formal, and final. The efficient cause is the agency generating a particular thing. The material cause is the stuff/matter from which something is generated. The formal cause is that which organizes matter to have the structure and characteristics it has. The final cause is the purpose or endpoint of the object, its function. A classic example of the four causes is Aristotle's commentary on sculpting a statue. The efficient cause is the sculptor and his sculpting; the material cause is the block of marble; the formal cause is the image to be sculpted, e.g., a statue of Hermes which specifies the "what it is" of the statue; and the final cause is the purpose, to create and exhibit art. The A–T analysis of human and animal cognition occurs within this framework.

The concepts of act and potency are also major metaphysical principles of the A–T model and go hand in hand with the casual principles presented above. A dictionary definition first: act is the intrinsic principle, which confers a definite perfection on a being, hence, a form. Potency is the capacity to be acted on or changed; the capacity to receive (a form), to be acted on, to be modified. An example may help clarify these concepts. Even today, we say that an athlete seems to have the potential to be great but her potential needs to be perfected through hard and extensive training. Thus, the potential of the athlete is actuated as positive change caused by the training.

Another principle of the A–T theory is the distinction between substantial and accidental forms. Substantial form is essential to the thing and accidental form is a quality of (inhering in) the substantial form but is not essential to it. In the statue example, its substantial form would be that of a statue or a depiction of someone or something. Accidental forms are such qualities as color, height, type of material, and such. It is important to point out that the A–T model does posit a moderate form of dualism; however, this is not the full substance dualism of Descartes. Unlike with Descartes' view, the human being is fundamentally unified in the A–T account, which asserts that humans are one substance only but possess faculties or powers, the intellect and will, which are immaterial.

The External Senses

Aristotle recognized five external senses and asserted that all knowledge begins in them. This claim immediately deals with the classical problem of grounding; humans and animals interact directly with the environment by means of the senses. The connectionist model presents somewhat less of an issue because connectionists presuppose a sensory connection to the external world. The problem for them arises when they try to simulate this process via computer based information processing.

Little will be said about the details of the physics, chemistry, anatomy, and neuroscience involved in external senses; these are givens. These factors explain "how" the external senses are "activated" and proceed into the brain for processing. Two words here are important. The "how" indicates that physics, etc. explain the mechanisms activating each sense. The term "activated" indicates that these various (and different) stimuli move the sense organs from passive to active. This process is a way of describing the metaphysics of potency and act discussed above. An obvious example occurs when I shut my eyes. Reflected light ceases and I do not see; however, my eyes remain in potency to shift into act (activation) as soon as I open my eyes and stimulation commences again.

A second A–T principle of metaphysics is at work here, namely the four-cause analysis presented above. The physics, chemistry, anatomy, and neuroscience describe efficient causes involved in activating the senses. The principles of matter and form are also at play. Aristotle asserted that all material reality is a combination of matter and form, with matter being the capacity to receive a particular form. Applied to the senses, this metaphysical principle means that each sense organ must have the potential (potency) to receive a particular form in order to activate (to move it from potency to act). Hence, the eye's material reality is constructed to receive a particular form of sensory stimulation. Once received, this combination results in the process of a particular instance of vision. It should be noted that the same sensory stimulation will not activate hearing. Though there is some overlap in sensory stimulation able to be "read" by each sense, in general each sense requires a particular combination of matter and form; the sense organ must have the capacity (material potency) to "read" the incoming "data." In the A–T model, the final cause is the end or purpose of the sense: vision to see; touch to feel and so forth.

The last sentence contains some terms that could easily occur in a classical or connectionist description of the sense organs. This illustrates the point that, at this level of discussion, there is much overlap between the A–T model and the classical/connectionist models. The major differences occur at the metaphysical level of explanation. Casual analysis, especially efficient causal analysis of the mechanisms in interaction, is quite similar.

The Internal Senses

Dawson (2013) discusses the idea of "reverse engineering" as used in AI. The process starts with outcomes, and engineers reconstruct the internal workings of a device from output inward. The ancient Greeks had their own version of reverse engineering, and Aristotle was the master of this process. His analysis of human behavior, without the benefit of psychological studies, imaging, or detailed neuroscience, led him to postulate "internal senses" involved in processing external sense input: the common sense, and the imagination (memory). Later medieval philosophers, Averros, Avicenna, and Aquinas, reformulated Aristotle to include four internal senses: the common sense, which receives and arranges all sense data; the imagination that combines and reassembles images; the estimative faculty, which gauges the dangerousness of the sensed object; and the memory, which retains the sensory level images or representations for later use. Shields (2003) points out that the A–T analysis of the internal senses flows from the hylomorphic theory and the Aristotelian theory of change from potency to act.

The internal senses are similar to the hypothetical constructs of modern psychology. Psychological constructs have two properties, one referring to "entities" such as known brain structures and hypothesized "entity-like" processes discussed below, and a second referring to the function of the construct and its interaction with other constructs. As we will see, Aristotle and later commentators followed a similar line of reasoning.

The common sense receives and arranges all input from the senses. Aristotle and later commentators located the common sense, as a biological entity, in the brain and determined its functions to be (a) integration of the external senses; (b) discrimination between the separate senses according to their proper object, meaning visual to visual stimuli, auditory to sound stimuli, etc.; (c) unifying those separate sensations into a single sensory perception; and (d) modification of currently sensed objects by "sense memory" based on past experience with that object. In modern cognitive science and psychology, this is known as perceptual binding and there is a vast psychological and neuroscience literature devoted to explaining the mechanisms involved in producing perceptual binding. This literature involves studies of brain structures (entities) as well as studies of theoretical constructs and their interactions (functions) related to sensory integration. An example would be John's (2002) hypothesized construct, described as perceptual frames with each sensory frame evaluated in the context of the previous frame and memory.

In terms of the relationship of the common sense to the classical model, it appears that the programmer, architecture, and algorithm provide the perceptual binding. As related to connectionism, presumably the "hidden elements" create the binding. However, both the classical model and connectionism focus on recreating this cognitive event in machines; although, of the two, connectionism is more directly brain focused. In fact, in modern cognitive psychology texts, one often finds connectionist oriented psychological topics followed by sections with labels like "Cognitive Neuroscience Support," which cite direct investigation of the brain as supportive to psychological theories and findings. This pattern repeats across the other internal senses.

The imagination receives the perceptually bound sensation and performs subsequent operations, including (a) retaining and recalling sensation organized by perceptual binding; (b) in humans producing the ability to combine percepts never directly sensed to be imagined (a unicorn); and (c) most importantly, generating the phantasm. The phantasm is sensory image or representation that is the product of perceptual binding and combinations of previous sensory perceptions. Learning and recall enrich the phantasm, which becomes more complex over time and experience. Nevertheless, it should be noted that, in the A–T model, the phantasm is at the sensory level as opposed to the conceptual level.

Both the classical and connectionist models recognize that a construct, such as the imagination, occurs in the brains of humans and, to a lesser degree, in animals. To a certain degree, computers can simulate this assembly of data and can produce a collected data set (a representation) similar to the phantasm. Hence, the A–T model and the classical and connectionist models are compatible at this level.

Aristotle recognized that both humans and animals have memory, which stores phantasms but is not simply a static storehouse. Spalding, Stedman, Gagné, and Kostelecky (2019) describe A–T memory as follows:

> It is important to remember that memory works in concert with all the other internal sense powers such that any phantasm might be stored, whether it is the more or less direct production of an external sensed object, or a combination of other sensory information from the common sense, or whether it is a phantasm (created by the common sense) of an imaginary thing. (p. 40)

Of course, memory has always been the central construct of the classical and connectionist models. In fact, both models assert that what is processed in information processing are sets of assembled data stored in RAM. Computer simulation of this process merely replicates what happens in the brain.

Little will be said about the estimative sense, except to note that "reverse engineering" makes it obvious that such a process exists. Perhaps a parallel to computer simulation occurs when defensive software in computers recognizes and fends off attacks by hackers.

At the levels of the external and internal senses, as noted in previous considerations (Spalding et al., 2019; Stedman, 2013), the A–T model, the classical model, and the connectionist model overlap to a high degree. The vast psychological, classical, connectionist, and neuroscience literature can be interpreted as fleshing out details unavailable to Aristotle and later philosophers. At the level of philosophical analysis, the A–T model is compatible with the classical and connectionist models, particularly in terms of analysis of efficient causal mechanisms.

Finally, the A–T model addresses weakness of the classical and connectionist models at this level. First the classical model. As previously noted, the A–T model deals with the classical model's problem of intentionality by grounding all cognitive phenomena in the external senses. The A–T model agrees with the connectionist view that the classical model fails to match up well with brain structure and processes. The A–T model solves that problem by dealing with the entire CNS.

The A–T model also addresses issues with the connectionist model. The problem of backpropagation, the objection that humans learn without numerous repetitions, does not occur in the A–T model. First, the A–T model deals directly with human leaning, not machine learning. Second, the A–T model deals with the objection that deep learning does not create new abstractions by postulating a philosophical paradigm to explain all abstraction. More will be said about this process in the next section. Finally, controversy regarding the similarity between the connectionist computer-based paradigm and the biological CNS does not occur because the A–T model deals directly with the sensory-to-CNS pathway.

Higher-Order Cognition

Higher-order cognition encapsulates a number of cognitive events: concept formation, thinking, reasoning, problem solving, decision-making, logic, meaning creation, theory of mind intelligence, and even consciousness. All of these cognitive phenomena start with concept formation and proceed from that base. In fact, the cognitive events can all be reduced to concepts and the relations between and among concepts; thus, we will consider only concept formation and reasoning. But first, how do the classical, connectionist, and A–T models account for these realities?

The classical account, as stated above, follows Thagard's (2019) description of mental representations as "analogous to computer data structures" containing "logical propositions, rules, concepts, images and analogies" (p. 7). These are arranged and rearranged by the computational processing system of the computer according to the classical model and, by analogy, in the brain. The classical account also claims that these representations carry meaning. Recall Dennet's (1978) claim that these data structures start as yes/no "idiots" that somehow become realities with rules, concepts, logical propositions and such. Perhaps conceptual representations in the classical account are actually supplied by meaning makers outside the computer and passed into the computer "Chinese Room" style. This notion receives support from Russell and Norvig (2009) who note that machine learning has not progressed on the important problem of constructing new representations at levels of abstraction higher than the input vocabulary.

Once concepts are loaded, computers can manipulate them in thought-like ways. For example, once concepts are loaded, computers can perform feats of cognition, such as syllogistic reasoning, problem solving, and decision-making.

However, these cognitive events are limited in scope, in that input comes from abstract concepts developed by outsiders, are passed into programs that control processes of assembly, and then output occurs.

In their review, Bringsjord and Govindarajulu (2020) note that AI can muster some support for "weak" AI. Under the "weak" criterion, both classical and connectionist models can pass the basic Turing test but not the total Turing test in which "a machine must muster more than linguistic indistinguishability: it must pass for a human in all behaviors — throwing a baseball, eating, teaching a class, etc." (p. 35). Bringsjord and Govindarajlulu go on to assert that AI has not come close to meeting the "strong" criterion, that is, production of machines that have all the mental powers of humans.

The connectionist model describes concept formation as neural networks, based in the brain, but replicable in computer-based "neural networks" produced by learning. As pointed out earlier, connectionists claim specific neural networks become concept-like as repeated deep learning trials modify weights in the hidden units. Critiques of this conceptualization, presented earlier, will not be repeated here. I will present a description of how this learning, known as back propagation, occurs according to connectionist theory (Goldstein, 2019):

> To explain the idea behind activation and back propagation, let's consider a behavioral example. A young child is watching a robin sitting on a bench, when suddenly the robin flies away. This simple observation, which strengthens the association between "robin" and "can fly," would involve activation. But, if the child were to see a canary and say "robin," the child's parent might correct her and say "That is a canary" and "Robins have red breasts." The information provided by the parent is similar to the idea of feedback provided by backpropagation. Thus, a child's learning about concepts begins with little information and some incorrect ideas, which are slowly modified in response to observation of the environment and to feedback from others. Similarly, the connectionist networks' learning about concepts begins with incorrect connection weights, which are slowly modified in response to error signals. In this way, the network learns that things that look like birds can fly, things that look like fish can swim, and things that look like trees are places where robins and other birds might perch. (p. 262)

Many psychological discussions of concept formation endorse this type of paradigm and present schematics that move from the particular canary to modules representing highly abstract concepts such as live, mammals, animals, man, energy, and such. A famous connectionist study (McClelland and Rogers, 2003) found that a computer, through backpropagation, could learn to differentiate daisy from canary after 250 trials and daisy, canary, and rose after 2,500 trials (one wonders how many trials it might take to differentiate energy vs. truth). This is in contrast to the previously cited study by Lake, Wojciech, Fergus, and Todd (2015) showing how children learn names of two-wheeled vehicles in one or a few trials. This critique, and others mentioned earlier, all cast doubt on the connectionist model as a complete explanation of concept formation. However, these critiques cast doubt on the computer simulation portion of the connectionist model, and less so on its claims about brain function.

As with the classical model, reasoning within the connectionist model involves various manipulations of concepts, some highly abstract and some very particular relating to persons, places, or things. The classical syllogism illustrates this point: all men are mortal. John is a man. Therefore, John is mortal. The classical model views all of these components as "computer data structures," loaded into the computer and manipulated. The model makes no distinction between the more abstract terms and the particular, John. The connectionist model regards all of these components as the products of learning, synaptic connections in the brain and deep learning structures in computer simulation of the brain. Unlike the classical model, the connectionist model does recognize differences in levels of abstraction, as in the canary example. The basic claim is that reasoning of all types assembles these abstractions (all men and mortal) and connects them with particulars (John) and other new and old abstractions (see Stedman, Hancock, Sweetman, 2009 for further discussion). Of course, both the abstractions (all men and mortal), the particular (John), and their connections are explained by the same processes as described above in the canary example.

The A–T model is a dualism; however, this is not the Cartesian dualism most psychologists and neuroscientists know. Feser (2005, 2009) refers to A–T dualism as hylomorphic dualism because it is founded on Aristotle's principles: the four causes, substance, and act and potency. Whereas Descartes postulated a reality of two separate substances, mind and body, hylomorphic dualism asserts that humans have one unified substance composed of matter and form. Recall that form is that which makes something what it is, such as the statue of Hermes. According to the A–T model, the human form has special powers, the intellect and the will, which mark a distinction between animals and humans.

Aristotle examined the object or product of the human intellect and deduced that the intellect abstracts a universal concept from particulars presented in the phantasm. This universal concept is the sensory composite of matter and form held in the brain and described by Shields (2003) as like a blueprint, applicable to all particulars covered by the concept. For example, humans can experience a particular dog, elephant, and fish and, via the process of abstraction, can understand that all fall under the universal concept, animal.

Reasoning, according to hylomorphic dualism, operates as an interplay between concepts, known by the intellect, and particulars, known by the external and internal senses. Reasoning takes place at several levels, concept to concept, concept to particular, and particular to particular. Again, the syllogism example: all men are mortal (concept to concept), John is a man (particular to concept), and James is also a man like John (particular to particular). Note that the processes meet all the requirements of the sense–think–act sequence important to the classical and connectionist models. Also, for those seeking additional information, note that elaborate presentations of this model occur from ancient (Aristotle, De Anima, trans. 1907) to contemporary sources (Feser, 2005, 2009, 2019; Madden, 2013; Spalding et al., 2019).

Hylomorphic dualism postulates a three-step process of concept formation and judgement. The first two steps have been described previously, namely, the formation of the phantasm and the process of abstraction. The final step, judgement, involves movement from the universal concept back to the particular. This involves confirming that the particular sensory composite of matter and form is a match to the universal concept (this process is known as conversion to the phantasm in the A–T model).

Conclusions

The first thing to conclude, perhaps surprisingly, is that there is much overlap between the classical and connectionist models and the A–T model. The sections on the external and internal senses describe these common elements in some detail. These same processes overlap with judgment, as the A–T model descends back into brain mechanisms. However, note that, in all of the sense–think–act sequence, the A–T model would employ a four cause, hylomorphic explanation, whereas explanation by the classical and connectionist models would rely on mechanistic, efficient cause oriented principles. The A–T model advocates hylomorphic dualism and the classical and connectionist models do not accept any form of dualism.

The A–T model offers solutions to problems plaguing the classical and connectionist models, as follows: (a) A–T solves the intentionality problem, a serious issue for the classical model and a continuing problem for connectionism, by dealing directly with the sensory system perception of the world (see Spalding, Stedman, Hancock, and Gagné, 2014, for more discussion); (b) A–T solves the computer–brain speed issues by dealing directly with the brain without any attempt to demonstrate equivalence of computers and the brain; (c) A–T offers a solution to the semantic or meaning controversy by reasserting a psychology that predates the Cartesian mind–body doctrine. Of course, the pre-Cartesian A–T psychology was challenged, specifically by the claims of Nominalism; (d) A–T incorporates the connectionist focus on learning. The A–T model would agree that learning enriches the external and especially the internal senses leading to more elaborate phantasms, which, though still an assembly of sensory data, allows for more elaborate abstractions of universal concepts; (e) A–T can incorporate all findings of neuroscience but would deny that these findings offer the entire explanation of cognition; and (f) the A–T model offers a better explanation of higher-order concept formation and thinking than the classical model's primitive electrical circuits or the connectionist "canary" learning model, although both can be incorporated into the A–T format. In sum, the A–T model is a viable alternative to the classical and connectionist models but is, in fact, compatible with both the classical and connectionist models at least at the brain level.

So far there has been little discussion of cognitive psychology specifically. That was purposeful because, as outlined in the goals, our primary interest was in the classical and connectionist models. However, cognitive psychology is tightly bound to some elements of the classical model and a great deal to connectionist theory. Hence, any weakness in the classical model and connectionism will carry over into cognitive psychology. I propose that cognitive psychology might profit from a closer look at A–T as a metaphysical underpinning (see Spalding et al., 2019 for more discussion).

References

Aristotle (1907). *De Anima* [R.D. Hicks, Trans.]. Cambridge: At the University Press. (Original date of publication unknown).

Baddley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Bringsjord, S. (2011). Psychometric artificial intelligence. *Journal of Experimental and Theoretical Intelligence*, *23*, 271–277.

Bringsjord, S., and Govindarajulu, N. S. (2020). Artificial intelligence. In *The Stanford encyclopedia of philosophy*. Retrieved from https: //plato.stanford.edu/entries/ artificial intelligence.

Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon Press.

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence, 47,* 139–159.

Buckner, C., and Garson, J., (2019). Connectionism. In *The Stanford encyclopedia of philosophy*. Retrieved from https://. stanford.edu. entries/ connectionism.

Cattell, J. M. (1986). The time it takes to see and name objects. *Mind*, *11*, 63–65.

Chalmers D. J. (1996). *The conscious mind: In search of a fundamental theory*. Oxford: Oxford University Press.

Cummings, R. (1989). *Meaning and mental representation*. Cambridge, Massachusetts: MIT Press.

Churchland, P. E. (2007). *Neurophilosophy at work*. Cambridge: Cambridge University Press.

Dawson M. R. W. (2013). *Mind, body, world: Foundations of cognitive science*. Edmonton, Alberta: Athabasca University Press.

Dennett, D. C. (1978). *Brainstorms*. Cambridge, Massachusetts: MIT Press.

Dennett, D. C. (1979). Artificial intelligence as philosophy and psychology. In M. Ringle (Ed.), *Philosophical perspectives in artificial intelligence* (pp. 57–80). Atlantic Highlands, New Jersey: Humanities Press.

Dreyfus, H. (1992). *What computers still can't do*. Cambridge, Massachusetts: MIT Press.

Feser, E. (2005*). Philosophy of mind: A beginner's guide*. London: Oneworld Publications.

Feser, E. (2009) *Aquinas: A beginner's guide*. London: Oneworld Publications.

Feser, E. (2019). *Aristotle's revenge: The metaphysical foundations of physical and biological science*. Neunkirchen–Seelschied, Germany: Editions Scholasticae.

Fodor, J. A. (1968). Appeal to tacit knowledge in psychological explanation. *Journal of Philosophy*, *65*, 627–640.

Goldstein, E. B. (2019). *Cognitive psychology: Connecting mind, research, and everyday experience (*fifth edition). Boston: Cengage Learning.

Goodfellow, J. J., Pouger–Abadie, J., Mirza, M., Xu, B., Wade–Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). *Generative adversarial nets*. Proceedings of the 27th International Conference on Neural Information Processing Systems [NIPS '14]. Cambridge, Massachusetts: MIT Press.

Gopnik, A., and Meltzoff, A.N. (1997). *Words, thoughts, and theories*. Cambridge, Massachusetts: MIT Press.

Harnad, S. (1990). The symbol grounding problem. *Physica D, 42,* 335–346.

Hastie, R. (2000). Problems for judgement and decision making. *Annual Review of Psychology*, *52,* 653–683.

Hofstadter. D., and McGraw, G. (1995) Letter spirit: Esthetic perception and creative play in the rich microcosm of the Roman alphabet. In D. Hofstadter (Ed.), *Fluid concepts and creative analogies: Computer models of the fundamental mechanisms of thought* (pp. 407–488). Cambridge, Massachusetts: MIT Press.

John, E. R. (2002). The neurophysics of consciousness. *Brain Research Reviews, 39*, 1–28.

Lake, B. M., Wojciech, Z., Fergus, R., and Todd, M. (2015). *Deep neural networks predict category typicality ratings for images.* Proceedings of the 37th Annual Cognitive Science Society. Pasadena, California.

Madden, J. D. (2013). *Mind, matter and nature: A Thomistic proposal for the philosophy of mind.* Washington, D.C.: Catholic University of America Press.

McClelland. J. L., and Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, *4*, 310–322.

McCarthy, J., Minsky, M., Shannon, C. E., and Rochester, N., (2006). A proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31,1955. *AI Magazine, 27*, 12.

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton–Century–Croft.

Newell, A. (1980). Physical symbol systems, *Cognitive Science, 4,* 135–183.

Pfeifer, R., and Scheier, C. (1999). *Understanding intelligence*. Cambridge, Massachusetts: MIT Press.

Robinson, D. N. (2008). *Consciousness and mental life*. New York: Columbia University Press.

Rogers, T. T., and McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, Massachusetts: MIT Press.

Russell, S., and Norvig, P. (2009). *Artificial intelligence: A modern approach (*third edition). Saddle River, New Jersey: Prentice Hall.

Searle, J. (1984). Minds, brains, and science. Cambridge, Massachusetts: Harvard University Press.

Searle, J. (1997). *The mystery of consciousness*. New York: New York Review Books.

Searle, J. (2014). What your computer can't know. *New York Review of Books*, *October 9 issue.*

Sejnowski, T. J., and Rosenberg, C. R. (1987). Parallel networks that learn to pronounce English texts. *Complex Systems*, *1*, 145–168.

Shields, C. (2003). Aristotle's psychology. In *The Stanford encyclopedia of philosophy*. Retrieved from http:/plato.stanford,edu.enteries/Aristotle-psychology.

Shiffrin, R. M., and Atkinson, R. C. (1969). Storage and retrieval processing in long-term memory. *Psychological Review*, *76*, 179–193.

Shimansky. Y. P. (2009). Biologically plausible learning in neural networks: A lesson from bacterial chemotaxis. *Biological Cybernetics*, *101*, 379–385.

Solso, R. L., MacLin, O. H., and MacLin, M. K. (2014) *Cognitive psychology (*eighth edition*)*. Essex, England: Pearson Education Limited.

Spalding, T. L., Stedman, J. M., Gagné, C. L., and Kostelechy, M. (2019). *The human person: What Aristotle and Thomas Aquinas offer modern psychology*. Cham, Switzerland: Springer Nature Switzerland AG.

Spalding, T. L., Stedman, J. M., Hancock, C., and Gagné, C. L. (2014). Intentionality and the Aristotelian–Thomistic view of concepts. *Journal of Mind and Behavior*, *35*, 245–262.

Stedman, J. (2013). Aristotle and modern cognitive psychology and neuroscience: An analysis of similarities and differences. *Journal of Mind and Behavior, 34,* 121–132.

Stedman, J., Hancock, C., and Sweetman, B. (2009). Cognitive psychology, neuroscience, and the problem of abstraction. *Contemporary Philosophy,* 29, 28–36.

Strzalkowski, T., and Harabagiu, M. S. (Eds). (2006). *Advances in open domain question answering*. [In the Series Text, Speech, and Language Technology, Volume 32]. Dordrecht, The Netherlands: Springer Verlag.

Thagard, P. (2019). Cognitive science, In *The Stanford encyclopedia of philosophy.* Retrieved from https://plato.stanford.edu/spr2019/cognitive science.

Treisman, A. (1988). Features and objects: The fourteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*, *40A*, 201–248.

Tulving, E., and Donaldson, W. (Eds.). (1972). *Organization of memory*. New York: Academic Press.

Turing, A. (1950). Computing machinery and intelligence, *Mind*, *59*, 433–460.

von Neumann, J. (1958). *The computer and the brain*. New Haven, Connecticut: Yale University Press.

Zhang, C., Samy B., Hardt, M., Recht B., and Vinyals O. (2016). *Understanding deep learning requires rethinking generalization.* [Cornell University] International Conference of Learning Representations. Rerieved from http/://arXlv.org/abs/1611.03530.