OPEN ACCESS

COGNITIO
Revista de Filosofia
Centro de Estudos de Pragmatismo
São Paulo, v. 26, n. 1, p. 1-13, jan.-dez. 2025
e-ISSN: 2316-5278

https://doi.org/10.23925/2316-5278.2025v26i1:e72521

Straying from the pragmatic ground of natural intelligence

Afastando-se da base pragmática da inteligência natural

Nathan Houser[*]
nhouser@iupui.edu

Abstract: We are in the vestibule of an era when Artificial Intelligence seems destined not only to manage, but to govern, the flow and use of information. Because the impact on human life and culture will surely be profound, it is imperative that the conception of intelligence that guides AI is conformable to human intelligence. Neuroscientist and AI developer, Jeffrey Hawkins, is convinced that AI is on the wrong path with its reliance on large language (big data) models, and that to achieve Artificial General Intelligence (AGI), the near-universal goal of AI developers, a biologically-inspired approach is needed. He has developed a promising new natural theory of intelligence based on the functioning of the human neocortex, but he dismisses old brain processes as mainly unnecessary and counter-productive for post-Darwinian AI. The naturalist conception of intelligence developed by the classical pragmatists a century and a half ago, provides a richer theoretical framework of intelligence that situates intelligent agents in interactive environments, the pragmatic ground of natural intelligence. Peirce, in particular, with his life-long quest to deeply understand how knowledge can be gleaned from experience, developed theories of perception, belief formation, semiosis, and cognitive logic, which taken together constitute the basis for a comprehensive theory of intelligence that incorporates instinctive and emotional intelligence and that allows for a conceptual space where thought can be focused on complex theoretical, normative, and aesthetic concerns and interests. This pragmatic account of intelligence, of mentality in general, is necessary for understanding the limitations and risks of AI as it rushes to develop AGI.

Keywords: Existential graphs. Hawkins. Natural intelligence. Neocortex. Peirce. Pragmatic ground. Semiosis.

Recebido em: 22/02/2025.
Aprovado em: 13/04/2025.
Publicado em: 14/07/2025.

Resumo: *Estamos no vestíbulo de uma era em que a Inteligência Artificial parece destinada não apenas a gerenciar, mas a governar o fluxo e o uso da informação. Como o impacto sobre a vida e a cultura humanas certamente será profundo, é imperativo que a concepção de inteligência que orienta a IA seja compatível com a inteligência humana. O neurocientista e desenvolvedor de IA Jeffrey Hawkins está convencido de que a IA está trilhando um caminho errado ao se basear em modelos de linguagem de grande escala (big data), e que, para alcançar a Inteligência Artificial Geral (AGI), objetivo quase universal dos desenvolvedores de IA, é necessário adotar uma abordagem inspirada biologicamente. Ele desenvolveu uma nova e promissora teoria natural da inteligência baseada no funcionamento do neocórtex humano, mas descarta os processos do cérebro mais antigo como em grande parte desnecessários e contraproducentes para uma IA pós-darwiniana. A concepção naturalista de inteligência desenvolvida pelos pragmatistas clássicos há um século e meio oferece um arcabouço teórico mais rico, que situa os agentes inteligentes em ambientes interativos, o terreno pragmático da inteligência natural. Peirce, em particular, com sua busca ao longo da vida para compreender profundamente como o conhecimento*

* Indiana University Bloomington.

1 This paper was presented on 4 November 2024 at the 23rd International Meeting on Pragmatism at the Pontifical Catholic University of São Paulo. I am indebted to Thomas L. Short for several in-depth discussions about the substance of this paper and to Ivo Ibri for his penetrating comments as the official respondent following its presentation.

pode ser extraído da experiência, desenvolveu teorias da percepção, formação de crenças, semiose e lógica cognitiva, que, tomadas em conjunto, constituem a base para uma teoria abrangente da inteligência. Essa teoria incorpora formas instintivas e emocionais de inteligência e permite um espaço conceitual no qual o pensamento pode se concentrar em preocupações e interesses teóricos, normativos e estéticos complexos. Essa concepção pragmatista da inteligência, da mentalidade em geral, é necessária para compreender as limitações e os riscos da IA em sua corrida rumo à AGI.

Palavras-chave: *Base pragmática. Grafos existenciais. Hawkins. Inteligência natural. Neocórtex. Peirce. Semiose.*

Hardly a day goes by without our being informed by one media outlet or another of the blistering advance of Artificial Intelligence. Typically, the news is positive about what has been achieved and what lies ahead; sometimes it is cautionary. Both can be justified. One thing seems clear: we are on the verge of a profound revolution in civilization, a revolution at least on the scale of the first information revolution that followed the introduction and wide-spread use of the printing press or the Industrial Revolution that supplanted the dominant agrarian and craft economies. We are in the vestibule of an era when Artificial Intelligence seems destined to not only manage but actually govern the flow and use of information. Massive amounts of money, resources, and research are being expended to make this happen. It is no longer a question of when the AI revolution will arrive; it is already here. What we don't yet know is how dominant AI will become or what its impact will be on human life and culture.

Many of the notable achievements of AI in robotics and automation, in data management, in healthcare diagnostics, in music and the arts, and so on, are impressive but it is the remarkable progress in natural language processing resulting in sophisticated language translation programs and large language systems which provide unprecedented human-computer interaction capabilities that has been most convincing, at least to the general public, that the intelligence of AI systems is real. Chatbots, especially, have captured our imagination because they easily impart the illusion of being thoughtfully, almost consciously, engaged with us. The rapid development and widespread use of chatbots like ChatGPT, Gemini, and Copilot, with their proficiency in processing natural language and providing comprehensive and usually accurate answers to most queries, has fostered the notion that not only have we crossed the threshold into the era of Artificial Intelligence but that achieving Artificial General Intelligence (AGI), when human intelligence is matched or surpassed, is just around the corner. Demis Hassabis, the CEO of Google's DeepMind, is convinced that AGI will be achieved within a decade. And we may be closer than we think to that great watershed moment that computer scientist and science fiction author, Vernor Vinge, has dubbed "technological singularity", when control of technology and technological growth shifts irreversibly to AI.

It is easy to fall into this way of thinking. In one of my chatbot exchanges, I asked ChatGPT if pragmatism is worth pursuing as an area of study and was assured that it is. ChatGPT gave several reasons why pragmatism is definitely worth pursuing: it has practical application and interdisciplinary relevance; it takes a problem-solving approach; it has both historical significance and contemporary relevance; and it emphasizes cognitive flexibility. After elaborating on these reasons, ChatGPT advised that "if these aspects resonate with [my] interests and goals, studying pragmatism could be a rewarding and intellectually enriching pursuit". Were I a student seriously pondering whether or not to pursue pragmatism as an area of study, ChatGPT's assurance might have triggered a release of dopamine and a confident decision to enroll in a course on pragmatism instead of one on analytic philosophy. And I would be sanguine in knowing that my artificial colleague, ChatGPT, or some other helpful chatbot, would be on hand to provide impressively smart answers to challenging questions – questions of the sort that philosophy professors might ask students to write essays on.

But as remarkable as the latest chatbots and advanced AI systems are, are they really intelligent, as AI boosters confidently declare them to be? Granted, there are usually caveats. Often even the more advanced large language model chatbots are said to be only narrowly or weakly intelligent but, nevertheless, they are confidently described as providing intelligent interactions between people and

machine. Other advanced AI systems, such as Tesla Autopilot or DeepMind's Alpha Zero, which taught itself to play Go, chess, and shogi at superhuman levels, are enthusiastically presumed to be intelligent but within a limited range of application. The prevailing assumption is that intelligence can be attributed to an AI system if it can perform tasks or operations requiring intelligence if performed by humans. But if that is our criterion, we would have to ascribe at least some level of intelligence to calculators and ordinary laptop computers. At most, we can say that the *humans would have to be intelligent to do it* requirement for ascribing intelligence to AI systems is a necessary but not a sufficient condition for genuine intelligence. Be that as it may, if an irreversible course has been set on the way to Artificial General Intelligence, with the expectation of relinquishing agency to AGI systems for many vital functions now controlled by humans, and if the criterion of intelligence being employed is *humans would have to be intelligent to do this,* then it is paramount that the people in charge of designing AGI systems have a deep understanding of what human intelligence is. It is far from clear that they do. It is not clear that anyone does. After trying without much success to isolate the common thread in the variety of ways intelligence has been defined, I decided that insofar as it is taken as the standard against which to gauge the success of AI, we are more in need of an appropriate theory of intelligence than an overarching definition.

Ever since Darwin shook things up with his theory of evolution, we have increasingly come to realize how interconnected everything is and that interaction, adaptation, and growth are ubiquitous. This lesson was taken to heart by the early pragmatists almost at once following the publication of *The Origin of Species.* Twenty-two years ago, at the 5th International Meeting on Pragmatism (IMP 5), I characterized pragmatism as the first teleological philosophy to grow out of the Darwinian conclusion that human intelligence is a thoroughly natural development and that human thought and language are means by which we mediate our past and present experiences with future expectations.[2] I believe the classical pragmatists, with their naturalist conception of intelligence, have much to contribute toward a general theory helpful for understanding some of the limitations of the conception of intelligence steering AI. In a few minutes I'll offer what I believe are key features of intelligence from the standpoint of pragmatic naturalism and will suggest that Peirce provides the basis for a formalization of the pragmatic general theory. But first I want to tell you about some interesting research from inside the AI community.

I was well along in preparing this address when I learned of a recent book entitled *A Thousand Brains* that promised a new theory of intelligence intended to advance AI research (Hawkins, 2022). Its author is neuroscientist Jeffrey Hawkins, a pioneer in the field of hand-held computing and co-founder of Numenta, a company focused both on investigating the neurophysiology of human intelligence and on developing machine intelligence. What caught my attention were two things in particular: first, Hawkins' candid statement of his belief that today's AI systems are not really intelligent and that AI will never achieve AGI following the path it is on (Hawkins, 2022, p. 145) and, second, his assumption that AI needs to be grounded in a theory of natural intelligence (Hawkins, 2022, p. 117). These were the leading ideas I had been following. Hawkins argues that the pursuit to build intelligent machines should begin with learning how intelligence arises naturally in brains and should then figure out how to apply those principles in developing AI. In following that course of research, Hawkins studied the evolution of the human brain from the earliest spinal nervous system that controls basic movements to the brain we have today, with its added layers to control increasingly complex behaviors. Generally, the human brain is regarded as having three main layers with the earliest being the "reptilian" or the "instinctual brain". The next layer to develop was the "limbic brain" also called the "emotional or feeling brain". Hawkins refers to these layers, together, as the Old Brain, which controls our more primitive behaviors. The newest part of the brain, which controls more sophisticated behaviors, is the neocortex, the "rational or thinking brain". Hawkins regards the neocortex as the "organ of intelligence".

2 "Pragmatism and the Loss of Innocence" (Houser, 2003). Presented: 5th International Meeting on Pragmatism, 7 November 2002.

In his book, Hawkins developed his theory of the function of the neocortex in some detail, explaining how its approximately 150,000 nearly identical cortical columns of neurons build predictive models from sensory input and situate those models in reference frames (like three-dimensional grids). Objects are modeled from differing perspectives and modalities in multiple columns providing distributed information that is coordinated and integrated across columns.[3] According to Hawkins, this modeling function "is the basis of our predictions, perceptions, and actions" (Hawkins, 2022, p. 34).

I will not attempt to summarize the neuroscience supporting Hawkins' theory, but I can recount some key principles he ascertained from his research. He learned that intelligence emerges from the interaction between sensing and movement in the world. He learned that the brain primarily learns through unsupervised exposure to the world, not through explicit training. He learned that the basic function of the cortical columns is to predict future sensory inputs based on past experiences. This prediction function is fundamental for learning and memory and for updating our model of the world. He learned that the neocortex not only processes sensory inputs but also integrates them with motor actions, enabling the brain to learn about the world through movement and interaction. He learned that the brain uses a common mechanism, reference frames, for processing spatial and conceptual information. And he discovered that learning from sensory inputs and experience is ongoing continuously. The neocortex is essentially a prediction system that generates habits from sensory experience. Hawkins is confident that for AI systems to be truly intelligent they must be aligned more closely with these principles of natural intelligence. He says that AI's dismissal of biologically-inspired approaches and its overreliance on deep learning have led it on a path that will never reach AGI, and he claims that his thousand brains theory "suggests that the future of machine intelligence is going to be substantially different from what most AI practitioners are thinking today" (Hawkins, 2022, p. 114).

Hawkins' theory is a significant attempt to understand and apply principles of natural intelligence to AI. My apprehension with the rapid rise of AI technology has been fueled by what I believe to be it disconnect with natural intelligence, so it was a relief to discover that, from within the AI community, someone was whistling a similar tune. But Hawkins' naturalism is markedly limited. While he champions the application of the principles of neocortex functionality as the basis for AI systems, he believes that most of the cognitive operations of the rest of the brain can and should be left behind. One of the infelicities of natural brain architecture, Hawkins believes, is that the neocortex has to signal the Old Brain to execute its instructions. This is an undesirable legacy from the brain's evolutionary past. The Old Brain evolved to govern behavior instinctively and by emotional triggers – intrinsic goals and values – which are absent in the neocortex which merely models the world, anticipates experience, and processes sensory input algorithmically. The natural programing of the Old Brain prioritized biological survival and protection which can trigger harmful and unpredictable behaviors responsive to strong urges and emotions. Hawkins believes that these reactive behaviors are dangerous and unnecessary given a sufficiently evolved neocortex, or in AI, systems that model the neocortex. So, Hawkins recommends that AI be concentrated on the principles of information processing derived from the study of the neocortex and let go the more primitive and less predictable processes of the Old Brain. In the absence of instincts and emotions, overarching goals and ethical guidelines can be programed in master algorithms and AI can become post-Darwinian, focused on higher-level cognitive functions including perception, reasoning, memory, language, and abstract thought. Intelligence will no "longer be held captive by the Darwinian processes that got us all here"

3 At this point in presenting this paper, I projected a slide showing five photographs of my teacup in different settings to illustrate how individual cortical columns build their separate predictive models of the world. In this simple example, five columns have modeled my teacup in different reference frames featuring variant characteristics – different positions, visual appearance, temperature, touch, etc. Given how often I use my cup, probably it is actually modeled it thousands of my cortical columns so that information about my tea cup is distributed across many columns. The variant predictive models from different reference frames are integrated, or merged, to generate a coherent general conception of my tea cup.

(Hawkins, 2022, p. 201). In his review of Hawkins' book, Bill Gates puts it this way: "We will eventually be able to create machines that replicate the logical, rational neocortex without having to wrap it around an Old Brain that's an "ignorant brute" wired for fear, greed, jealousy, and other human sins" (Gates, 2021).

A few minutes ago, I remarked that pragmatism was born in Darwin's wake as his influence swept over the American continent and I indicated that pragmatism was the first major post-Darwinian philosophical naturalism. The tenet of pragmatic naturalism that concerns me tonight, the pragmatic ground I refer to in my title, is the assumption that intelligence, and in particular, human intelligence, is a product of evolution – a natural outcome of the lengthy mutual adaptation of our species with our environmental circumstances. The crucial link between natural adaptations, including human intelligence, and environmental circumstances is experience. For the pragmatist, pre-conceptual cognitive experience of the kind processed by the so-called "Old Brain" is essential for intellectual development and a key factor necessary for robust intelligence operative in a complex ever-changing world. Moreover, the pragmatist does not believe that one's intelligence is only stored in the brain as does Hawkins, who holds that "everything we know is stored in the connections between neurons" (Hawkins, 2022, p. 37). While it is true that intelligence must be embodied, it is essentially functional, not physical. It is both apparent and fortunate that the founders of pragmatism were able to store some of their intelligence in the books and articles they wrote, and it is unwarranted to suppose that they did not continue to have easy access to that intelligence by consulting their writings, just as we do. So, pragmatists regard intelligence, and mind in general, as logically, or functionally, independent of its embodiment, whether that embodiment is a brain, an institution or cultural practice, an instinct, or an artifact.

I believe the account of intelligence, of mentality in general, found in pragmatism is necessary for understanding where AI is heading, especially as it seeks to develop AGI. As most of you here are well aware, after so many years of meetings on pragmatism, pragmatists do not all sing the same tune and certainly not in the same voice, but all are naturalists, as I have characterized naturalism, and all stand on the pragmatic ground I have described. Of the early pragmatists, John Dewey is usually regarded as the most influential early proponent of philosophical naturalism and his 1925 *Experience and Nature* was a key text for the naturalist movement that was prominent in America during the early 20th century. But that wave of American naturalism died out with the rise of analytic philosophy (Bernstein, 2020, p. 6).

A second wave of naturalism, one that is still very active today, may be said to have begun with Quine's *Two Dogmas of Empiricism* (Quine, 1951). It is noteworthy that Quine cited Peirce's behavioral theory of the formation and role of belief as a major advance toward naturalizing philosophy (Quine, 1981, 36-37). Peirce and Quine each held that the evolution of intelligence begins with experience and they both embraced the peripatetic maxim: nothing is in the intellect that was not first in the senses.[4] Quine called the maxim the watchword of epistemology (Quine, 1990, p. 12). But Peirce and Quine did not have the same understanding of what it means to be "in the senses." What it meant for Quine was "our meager contacts" with the physical world, that is, "the mere impacts of rays and particles on our surfaces and a few odds and ends such as the strain of walking uphill" (Quine, 1995, p. 16). What Peirce understood by "in the senses" was "in a perceptual judgment" and he held that the intellect consists of cognitions of all kinds including general symbols (EP 2:227). That kept Peirce outside of the wave of naturalism that runs from Quine through Dennett because of resistance to Peirce's metaphysical realism – his view that reality is not limited to what exists and, especially, his acceptance of the reality of generals and abstract entities and his belief that they are consequential.

4 The first recorded expression of the Peripatetic maxim (or axiom) is found in Thomas Aquinas' *Questiones Disputatae de Veritate*, question 2, article 3, argument 19: *Nihil est in intellectu quod non sit prius in sensu.* Peirce rendered the maxim as: *Nihil est in intellectu quod non prius fuerit in sensu* (EP 2:226).

Peirce rejected the Lockean idea that, at birth, the human mind is a blank slate to be filled in only by our experience and supposed that our brains have evolved to embody important innate non-conceptual intelligence, as with other animals and even insects. Comparing Hawkins' view with that of Peirce, he seems to agree with Peirce that humans come supplied with instinctive intelligence, thanks to the Old Brain, but he regards most Old Brain cognitive programming as burdensome and an impediment to rational behavior.[5] So, except for a few "behavioral primitives", programmed to manage sensory intake and control basic movements, Hawkins recommends that Old Brain "intelligence" be left behind when designing AI systems. In recommending that artificial intelligence be modeled primarily on the neocortex, the rational brain, he is more-or-less adopting Locke's blank slate theory for AI.

I believe that the theory of intelligence that can be drawn from Peirce's philosophy of mind (more specifically, from his writings on experience, perception, semiotics, and normative logic) has particular relevance and importance for AI. I have included Peirce as a pragmatic naturalist along with most other pragmatists because Peirce shared, even with pragmatists with whom he disagreed on technicalities, what he referred to as "our pragmatist faith" (EP 2:421).[6] But Peirce insisted that his narrower form of pragmatism, which he sometimes called "pragmaticism" to distinguish it from the more general doctrine, was better suited to serve its original purpose of providing a method for clarifying ideas and concepts. In his later years he restricted the purview of pragmaticism to "intellectual concepts".

According to Peirce, an important feature of his narrower pragmatism was that it could be proved. I know that Peirce's proclaimed proof of pragmatism has been the subject of several presentations at these meetings, over the years, and I am aware that there is disagreement about both his purpose, as well as his success, in trying to prove pragmatism.[7] I have come to believe that Peirce believed that pragmatism, as he understood it, was closely allied with the nature and function of conceptual, or symbol-based, cognition and that his interest in proving pragmatism was related to his quest to understand the development and growth of conceptual intelligence – the intelligence Hawkins attributes to the neocortex. After formulating different proofs of pragmatism based on his theory of perception and his theory of signs, Peirce gravitated to his Existential Graphs (EG) believing his logical graphs might facilitate the weightiest and most revealing proof. Peirce believed that a special feature of EG was that it represented thought in action: he called it "a moving-picture of thought" (CP 4.11). I believe Peirce was pioneering what we now call cognitive science and that he achieved something with his graphs that has relevance for AI development.

I will not try to demonstrate this; partly because it would take too much time but, mainly, because I am long out of practice working with EG.[8] I will just point out that EG is a complete and consistent system of logic, with a diagrammatic syntax, equal in power and functionality to more familiar linear, or string, logic systems. It seems to model natural reasoning, as I believe Peirce intended, and is therefore relevant for understanding his theory of intelligence. Five years ago, Ahti Pietarinen began publishing a multi-volume edition of Peirce's work on EG which he predicts will better fit the needs of cognitive science than predicate-calculus-oriented string logics (Peirce, 2020-2025). He refers to EG as "the logic of the future". This is not idle speculation. John Sowa, with his system of Conceptual Graphs, based on Peirce's EG, has made significant contributions to the development of AI and, in particular, knowledge representation and natural language processing.[9] EG continues to be of interest to cognitive science in these areas and it is of increasing interest for research on visual reasoning and cognitive development.

---

5 Hawkins also regards most Old Brain programming as detrimental to the transition of our species from a biological species to a technological species, a transition he hopes for (Hawkins, 2022, p. 197).

6 I have elaborated on this in Houser (2010a) and Houser (2010b)

7 For example, see Hookway (2005; 2008; 2011), and Pietarinen (2008).

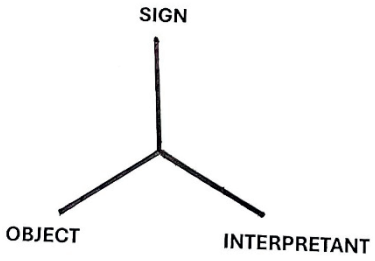
8 See my introduction to EP2 (pp. xxvi–xxxvii) for a brief discussion of Peirce's planned proof using EG.

9 Sowa continues to work on developing suggestions Peirce made to extend the modal application of EG in ways anticipatory of current applications in cognitive science and AI.

But EG is a conceptual logic addressing the realm of intelligence associated with the neocortex.[10] Conceptual logic does not deal with the incipient formation of the concepts it works with even though that constitutes a major component of our cognitive activity. To get a more complete account of Peirce's general theory of intelligence we must take into account his theory of perception and his theory of signs.

A complete reconstruction of Peirce's theory would consider the biological evolution of intelligence as well as the formal, or programmatic, structure of intelligence. Peirce's account of how intelligence develops and propagates is based on the Darwinian thesis that the function of natural intelligence is to learn from current experience in order to predict future experience and adjust behavior accordingly. Peirce's early account of learning from experience was his well-known doubt/belief theory of inquiry, arguably the basis of his pragmatism. He held that thought (or inquiry) is motivated by states of doubt (states of irritability resulting from environmental exigencies) and ends when those states of irritability have been resolved by suitable actions (generating or contributing to natural beliefs). Peirce says that the function of thought is to produce belief and that belief is a rule of action – a habit. Of course, learning from experience, even just having experience, depends on an actual link-up between our brain and what is experienced. Sensation is our operative link with the world we inhabit, and perception is our cognitive connection to what is experienced. Peirce's account of the progression from sensory impingements (proximal stimuli) to percepts, and from percepts, via abduction, to perceptual judgments, and thus, conceptualization, is fundamental for his complete account of human intelligence,[11] but probably what is most important for understanding Peirce's general theory of intelligence is his theory of signs.

Many of you have seen, perhaps ad nauseam, the simple diagram representing Peirce's basic sign relation:



This simple irreducible triad represents the basic component of intelligence. It is not difficult to imagine, if somewhat vaguely, how a percept, as a sign, can relate an object that is sensed, with a concept, the interpretant, by virtue of a perceptual judgment. Although this fixed diagram represents the basic sign relation, it is important to remember that signs are not static objects but are processes or routines for deriving information (knowledge) from experience to guide behavior and to develop or elaborate knowledge previously drawn. Consequently, it is more accurate to think of the basic sign diagram as a diagram of semiosis, or sign action. Semiosis with respect to its objects occurs in stages which led Peirce to consider objects either as *dynamic*, the external object that determines the sign, or as *immediate*, the object as the sign represents it. With respect to interpretants, Peirce distinguished them as either *immediate*, as the sign represents them, *dynamic*, the actual effect produced by the sign, or *final*, the habit that exhausts the function of the sign. Drawing on these six constituents of semiosis, Peirce was able to identify sixty-six classes of signs – sixty-six basic forms of cognition. (I'm not about to start charting out the sixty-six forms, which many of you are already familiar with. In fact, the study

10 Peirce apparently envisioned an expansion of EG to treat non-conceptual intelligence. In discussing his use of tinctures to distinguish different modal universes on which graphs could be scribed, Don D. Roberts explained that the tinctures were designed with more than formal logic in mind; they were meant to provide a structure in terms of which Peirce could apply his categories to propositions and inferences, to hypotheses, questions, and commands, to "all that ever could be present to the mind in any way or any sense" (Ms 499(s)) – and to do this with the same system that handled the formal logic as well (Roberts, 1973, p. 100 [the imbedded quotation is from Peirce]).

11 See Houser 2024 for some treatment of Peirce's account of the progression from proximal stimuli to perception.

of this taxonomy has probably been more intensively studied here in São Paulo than anywhere else).[12] These sixty-six forms of cognition constitute a key component of Peirce's theory of intelligence; they map the cognitive space in which intelligence functions. It is notable that Peirce believed that semiosis progresses toward the formation of habits, which he called "the most characteristic property of the nervous system" (EP 1: 264) and which Hawkins regards as essential for intelligence.

I want to emphasize one last thing about Peirce's theory of signs: the way he differentiated interpretants according to their ontological natures. The interpretants effectuated by signs may be emotional, energetic, or logical, corresponding to Peirce's categorial triad of feeling, action, and thought. A close study of Peirce's sixty-six classes of signs reveals that most of the forms that semiosis can take involve emotional or energetic interpretants. With reference to the fragment of brain science that Hawkins provided, it seems likely that emotional and energetic interpretants, and the habits they produce, are processed by the Old Brain, and that logical interpretants, which we can regard as conceptual thought, are processed in the neocortex. However, for Peirce, as for pragmatists across the board, even the most developed conceptual intelligence is fundamentally interlaced with emotional and energetic intelligence.

Peirce's semiotic theory addresses the full scope of intelligence, largely from the standpoint of its formal or structural characteristics. His Existential Graphs treat conceptual intelligence as an operational system of rational inference – especially relevant to language-based or symbol-based reasoning. Together, along with his theory of experience and perception, they constitute a comprehensive treatment of intelligence that takes full account of its non-conceptual base. Although Peirce's sign theory and Existential Graphs were unique to him among the early pragmatists, I believe the general theory implicit in his comprehensive treatment of intelligence applies to pragmatic naturalism more broadly.

This account of general intelligence aligns closely with Hawkins' views on the development of intelligence as the brain evolved from a primitive cluster of neurons into the layered brain we have today, but unlike Hawkins, pragmatic naturalists do not believe that the higher-level cognitive functions of the neocortex can be severed from the programming of the brain's evolutionary past.[13] Although Hawkins' research convinced him that experience is a prerequisite for the development of intelligence, aligning him with pragmatists, by narrowing his conception of intelligence to the operations of the neocortex he forsook its connection with direct experience: the forms of experience Peirce associated with firstness (feeling) and secondness (reactive consciousness). His conception of intelligence is similar to that of traditional cognitive science, which regards all mental processes as computational, contrary to the emerging view of embodied cognition theorists who hold that an adequate model of intelligence must take account of the role of noncomputational (and nonconceptual) interactions with the environment.[14]

Why, one wonders, does Hawkins, while insisting that to achieve Artificial General Intelligence AI research should adopt a theory of natural intelligence modeled on the brain's neural processing, adopt a limited theory based only on the functioning of the neocortex, thus shifting the balance in decision-making heavily to purely rational or logical processing. For one thing, I don't believe anyone has any idea how to design AI architecture to model a comprehensive theory of intelligence like Peirce's. In recent years, Terrence Deacon has developed a plausible theory of the evolution of semiosis-based intelligence[15] that is informed by contemporary neuroscience but, like Peirce's more formal theory, it

12 For a more thorough examination of Peirce's sign classes see Houser (2023).

13 Hawkins recognizes that some of the functions of the Old Brain, in particular, functions that create goals and priorities that regulate or influence the execution of neocortex command signals, continue to be necessary for operative human intelligence as well as for the achievement of AGI. Although learning and reinforcement protocols can be programmed to give rise to the evolution of normative behaviors of AGI systems in consequence of experiential interactions, the normative base for AI cannot feasibly be the bare-bones survival protocols of primitive biological organisms. The normative base for AGI will have to be embedded ethical frameworks and regulations and constraints encoded by AI designers and programmers.

14 Hawkins does stress the importance, even necessity, of physical mobility for intelligent systems but his focus is on the acquisition of data, not on the consciousness of movement or the experience of secondness.

15 Terrence W. Deacon, *Incomplete Nature: How Mind Emerged from Matter* (New York: W. W. Norton & Co., 2012).

is thought to be too comprehensive and complex to be of practical application for AI – and, besides, comprehensive theories like Peirce's and Deacon's present programming challenges that are beyond the foreseeable capabilities of current AI technologies. So, Hawkins' decision to limit his theory of natural intelligence to the neurophysics of the neocortex and, roughly speaking, the logic of conceptual intelligence, was partly to narrow the focus of AI research to higher-level cognitive functions that are more easily programmable. But mainly, I believe, he is convinced that the more primitive and less predictable processes of the Old Brain have served their evolutionary purpose and have become a hindrance to the efficient functioning of the rational neocortex. So even if it were possible and feasible to program a master algorithm for general intelligence with the comprehensiveness of Peirce's theory, it would not be Hawkins' theory of choice because it would recapitulate the non-conceptual reactive and emotional triggers that add the elements of non-rationality and unpredictability that Hawkins believes are undesirable for AI. What Hawkins aims for, and what AI is striving for in general, are high-power logic-driven non-empathic systems that can outperform humans in processing algorithmically representable information – the sort of information processing that occurs in the neocortex. Instead of appealing to the Darwinian evolved Old Brain to execute its orders, as the neocortex must do, Hawkins' robots will depend on AI programmers for normative rules and constraints.

Consider what will be the mental contents of a Hawkins robot that is fully loaded with a massive store of information, lightning-speed processors (perhaps quantum processors), and with high order mechatronic smart sensors for in-time situational and environmental input. At most, we can say that its mental contents will consist of data (information, loosely speaking) and algorithms. The mind of the Hawkins robot will be empty of feeling, of affect of any kind; it will only be an information processor and will have no contents of consciousness. To regard this robot as truly intelligent is to dissociate from intelligence the vast range of nonconceptual experience for which feeling and movement, as forms of consciousness and expression, are central. One thinks of the arts, of painting, music, dance, for example, and even of poetry, which, though symbolic in form, can hardly be said to be 'understood' absent aesthetic affect.[16] No wonder AI engineers have not been able to figure out how to program a computer with the overall intelligence, the everyday knowledge, of a five-year-old child (Hawkins, 2022, p. 121-123); how could they when their designed minds are incapable of curiosity or excitement or a sense of wonder; when their robots can't feel hunger or desire and have no capacity for pleasure and pain?

A limited conception of intelligence as unconscious information processing makes sense for AI research related to products and services of the sort, we are becoming familiar with – who wants emotionally reactive EV's or chatbots? But to apply Hawkins' narrow post-Darwinian conception of intelligence to the design of AGI systems, systems intended to match or surpass human intelligence, is seriously risky. AGI systems are expected to have responsibility for vital tasks that involve moral or ethical considerations, but without the instinctive and emotional guidance of Old Brain intelligence (the pragmatic ground of natural intelligence). Hawkins believes that adequate ethical frameworks and social values can be programmed into AGI algorithms to mitigate the risks, but those guidelines and values will not have the deep-rooted influence of the normative values ingrained in the characters of naturally evolved humans. It will be a derivative normativity, reflecting the values of the AI programmers but just another set of algorithms from the standpoint of the AI system. It seems evident that AGI systems programmed to strictly follow algorithms, without consciousness, conscience, and empathy, are functionally sociopathic. What could go wrong?

Probably worst case is the risk to civilization of technological singularity. It is all too easy to imagine that after singularity, should that day ever arrive, when virtually complete technological control has passed to AGI systems, a purely rational decision will be made to abandon moral and ethical

16 Ivo Ibri has emphasized in many of his writings the fundamental importance of *the poetic ground*, the non-conceptual experience, out of which knowledge grows: see, for example, Ibri (2016). See, also, Houser (2018).

considerations (human-centric values) as illogical constraints on progress. That is a dark thought and maybe a farfetched one. But even if singularity never arrives, and even if AI never manages to surpass humans in overall intelligence, the AI revolution is still upon us and, notwithstanding its many benefits, there is good reason for concern. Most troubling, I believe, is who is in control of AI research and development. I don't doubt that some early AI research, much of which was funded by universities, was motivated by a sincere desire to use advances in computer programming and technology to learn more about mentality and human cognition. But AI research is expensive and now its principal funders are billionaire entrepreneurs, big businesses, and government agencies. Profit, power, and control are the great motivators. Although the stated mission of OpenAI is **"to ensure that artificial general intelligence […] benefits all of humanity**",[17] in September 2024 it was reported that it is transitioning into a traditional for-profit company (Isaac; Metz, 2024). If OpenAI begins to deviate from its stated mission, it will be able to blame Adam Smith's invisible hand.

The largest funders of AI research today are the U.S. and Chinese governments, but other countries are busy trying to catch up. Governmental support for AI research is largely motivated by a desire for power or to maintain an edge over other countries in advanced industrial development and warfare. While these may be legitimate interests, the consequences may be malign: invasive surveillance and loss of privacy, more effective tools for propaganda and psychological control, science fiction-like instruments of war – but not fiction. I cannot help but think of the Manhattan Project and how some project scientists believed that there was at least a remote possibility that the initial test detonation might set off a chain reaction that would destroy life on earth. Yet the challenge for the scientists, and the prospect of wielding such power for the U.S. Government, was too great to forego. This will also be the case for AI weaponry.

Of course, it is possible that AI could open a pathway to advanced civilization. Imagine the benefits of a vast network of robots relieving humans of the work they perform only to keep food on their tables. If societies distributed the benefits of AI generated wealth fairly and wisely, the new unemployed could live comfortably enjoying their lives and, hopefully, spending time in educational and cultural pursuits. But I fear it is more likely that the leisure class will become smaller and smaller, characterized by conspicuous consumption (a term coined by Thorstein Veblen, a one-time student of Peirce at Johns Hopkins) and the increasing number of unemployed will be left to fend for themselves.[18]

The main point of my talk tonight has been that unregulated AI based on a theory of intelligence limited to logical information processing, without the normative guidance of values and constraints naturally acquired in the give and take of evolution, is a threat to civilization. Pragmatic naturalism, and Peirce especially, provides a theory of intelligence that does not presume to be post-Darwinian by dismissing the continuing importance of instinctive and emotional intelligence. But I don't want to leave you with the idea that I believe natural human intelligence is ideal. Bearing in mind the astounding damage humans have wrought across the face of the Earth, it is obvious that it isn't. Surely Hawkins is right in laying some of the blame for the evils wrought by humans on the influence of the Old Brain. But to conclude that the development of natural intelligence has reached such a pinnacle in producing the neocortex that the Old Brain has outlived its evolutionary importance is the wrong way to stray from the pragmatic ground of intelligence. There is a better way. That is to develop the intellectual potential of the higher-level functions in harmony with the reactive and normative intelligence of the Old Brain. The evolution of language, of symbol-based semiosis more generally, put humans on a path that has led far beyond the natural pragmatic ground of intelligence, and I suppose it was in large part the increasing use of language that led to the evolution of a larger and larger neocortex. I also suppose that greater processing capacity for the neocortex would not have been selected, as we made our evolutionary journey from the past, if language use had not harmonized with and generally enhanced the survival priorities of

17 HTTPS//OpenAI Charter | OpenAI/

18 Thorstein Veblen introduced the expression "conspicuous consumption" in Veblen (1899, chapter 4).

the Old Brain. But even though human language may have evolved in support of vital concerns of life, it has expanded the scope of mind far beyond programmed instinctive behaviors and the integrated system of beliefs that constitute natural intelligence and extends human mentality deeply into conceptual space where thought can be focused on complex theoretical and aesthetic concerns and interests. With the aid of language and other complex symbol systems, humans have interwoven their natural intelligence with an extended cultural mind (embodied not only in language but in the manifold of cultural artifacts and practices that convey and promote that culture) and they have assumed theoretical and aesthetic goals far removed from natural evolutionary requisites. Language and abstract thought have facilitated such a sophisticated expansion of human intelligence that thought can be turned critically back on itself to examine the core beliefs that give vitality and sustainability to the supporting culture. Straying from the pragmatic ground of natural intelligence in this way is not to abandon it but to build on it. That, I believe, is the path forward for bona fide intelligence, benefitting as it can from AI but not yielding normative or moral agency to it.

References

BERNSTEIN, Richard J. *Pragmatic Naturalism:* John Dewey's Living Legacy. New York: Graduate Faculty Philosophy Journal (The New School), 2020.

GATES, Bill. Is this how your brain works?. *GatesNotes: The Blog of Bill Gates,* Nov. 22, 2021.

HAWKINS, Jeffrey C. *A Thousand Brains*: a New Theory of Intelligence. New York: Basic Books, 2022.

HOOKWAY, Christopher. The Pragmatist Maxim and the Proof of Pragmatism. *Cognitio: Revista de Filosofia,* v. 6, n. 1, p. 35-42, 2005.

HOOKWAY, Christopher. The Pragmatist Maxim and the Proof of Pragmatism (2): After 1903. *Cognitio: Revista de Filosofia,* v. 9 n. 1, p. 57-72, 2008.

HOOKWAY, Christopher. The Pragmatist Maxim and the Proof of Pragmatism (3): Habits and interpretants. *Cognitio: Revista de Filosofia,* v. 12 n. 1, p. 89-104, 2011.

HOUSER, Nathan. Pragmatism and the Loss of Innocence. *Cognitio: Revista de Filosofia,* v. 4, n. 2, p. 197-210, 2003.

HOUSER, Nathan. The Church of Pragmatism. *Semiotica,* v. 178, n. 1/4, p. 105-114, 2010a. https://doi.org/10.1515/semi.2010.007

HOUSER, Nathan. The Increasing Attractiveness of Classical Pragmatism. *Cognitio: Revista de Filosofia*, v. 11, n. 2, p. 224-240, 2010b.

HOUSER, Nathan. Ibri's Peirce: Poetry and Play in the Life of the Mind. *In:* GHIZZI, E. B.; DANTAS, L. F. N. S.; MADEIRA, M. GONZALEZ, M. E. Q; AIUB, M. (eds.). *Sementes de Pragmatismo na Contemporaneidade;* Homenagem a Ivo Assad Ibri. São Paulo: FiloCzar, 2018. p. 33-45.

HOUSER, Nathan. Peirce's Ongoing Semiotic Project. *In*: BORGES, P. M.; FRANCO, J. O. R. (eds.). *Tempo da colheita: homenagem à Lucia Santaella* [*Harvest time: Festschrift for Lucia Santaella*. São Paulo: FiloCzar, 2023. p. 27-52.

HOUSER, Nathan. The Vicissitudes of Experience. *In*: DE WAAL, C. (ed.). *The Oxford Handbook of Charles S. Peirce*. Oxford: Oxford University Press, 2024. p. 77-93.

IBRI, Ivo A. Linking the Aesthetic and the Normative in Peirce's Pragmaticism: A Heuristic Sketch: Charles S. Peirce Society 1016 Presidential Address. *Transactions of the Charles S. Peirce Society,* v. 52, n. 4, p. 598-610, 2016. https://doi.org/10.2979/trancharpeirsoc.52.4.04

ISAAC, Mike; MEYZ, Cade. OpenAI Executives Exit as C.E.O. Works to Make the Company For-Profit. *The New York Times,* Sept. 25, 2024.

PEIRCE, Charles S. *The Essential Peirce:* Selected Philosophical Writings, 2 vols: v. 1. N. Houser; C. Kloesel (eds.); v. 2. Peirce Edition Project (ed.). Bloomington: Indiana University Press, 1992-1998. Referred to as EP.

PEIRCE, Charles S. *Logic of the Future.* Writings on Existential Graphs. 3 vols in 4. A.-V. Pietarinen (ed.). Berlin: De Gruyter, 2020-2025.

PIETARINEN, Ahti-Veikko. Comments on Hookway, 'The Pragmatist Maxim and the Proof of Pragmatism (2): After 1903'. *Cognitio: Revista de Filosofia*, v. 9, n. 1, p. 85-91, 2008.

QUINE, W. V. Two Dogmas of Empiricism. *The Philosophical Review,* v. 60, n. 1. p. 20-43, 1951. https://doi.org/10.2307/2181906

QUINE, W. V. The Pragmatists' Place in Empiricism. *In*: MULVANEY, R. J.; ZELTNER, P. M. (eds.). *Pragmatism, its Sources and Prospects*. Columbia, SC: University of South Carolina Press, 1981. p. 21-39.

QUINE, W. V. *From Stimulus to Science.* Cambridge: Harvard University Press, 1995.

VEBLEN, Thorstein. *The Theory of the Leisure Class. New York: MacMillan, 1899.*

OPEN ACCESS

COGNITIO

Revista de Filosofia
Centro de Estudos de Pragmatismo

São Paulo, v. 26, n. 1, p. 1-13, jan.-dez. 2025
e-ISSN: 2316-5278

https://doi.org/10.23925/2316-5278.2025v26i1:e72521