

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

Robot Consciousness: Physics and Metaphysics Here and Abroad

Stephen B. Ripley, M'17

Research Director, Futures & Forecasts, Inc., Vancouver, B.C., Canada, V6R4H8

Corresponding author: Stephen B. Ripley (e-mail: sbripley@shaw.ca)

Funding: none

ABSTRACT Interest has been renewed in the study of consciousness, both theoretical and applied, following developments in 20th and early 21st century logic, metamathematics, computer science, and the brain sciences. In this evolving narrative, I explore several theoretical questions about the types of artificial intelligence and offer several conjectures about how they affect possible future developments in this exceptionally transformative field of research. I also address the practical significance of the advances in artificial intelligence in view of the cautions issued by prominent scientists, politicians, and ethicists about the possible dangers of such sufficiently advanced general intelligence, including by implication the search for extraterrestrial intelligence.

Integrating both the theoretical and practical issues, I ask the following:

(a) is sufficiently advanced general robotic intelligence identical to, or alternatively, ambiguously indistinguishable from human intelligence and human consciousness, and if so,

(b) is such an earthly robotic intelligence a kind of consciousness indistinguishable from a presumptive extraterrestrial robotic consciousness, and if so,

(c) is such a human-created robot preferably able to serve as a substitute for or even entirely supplant human intelligence and consciousness in certain exceptionally responsible roles? In the course of this investigation of artificial intelligence and consciousness, I also discuss the inter-relationships of these topics more generally within the theory of mind, including, emergence, free will, and meaningfulness, and the implications of quantum theory for alternative cosmological ontologies that offer suggestive answers to these topics.

INDEX TERMS algorithm design, artificial consciousness, artificial general intelligence, AI, artificial intelligence, brain, classification, complexity theory, computational complexity, computer science, cyborg, deep learning, epistemology, existential risks, extraterrestrial intelligence, generative adversarial networking, holism, human consciousness, language, limits to knowledge, logic, materialism, mathematics, mereology, mind uploading, monism, natural language, ontology, panentheism, panpsychism, philosophy of mind, quantum computing, robot, supernatural, synthetic biology, synthetic mind, theory of mind, Turing machine, universal quantum computer

I. INTRODUCTION: HISTORICAL CONTEXT AND SYNOPSIS OF THE ARGUMENT

A. HISTORICAL CONTEXT

Speculative beliefs about the phenomena of “mind” [1] and “consciousness” [2]–[6] are found in the literature of early myths [7], religions, philosophies, and science. Depending on the source, the earliest theorizing seemingly used the concepts as mythopoetic primitive terms or conflated or circularly defined the two concepts along with the concept of “intelligence”. At the outset, in this essay, I refer to the collection of all three terms as “mental” phenomena. The oral and sung literature (now transcribed) and written literary classics of both the West and East are supported by anthropological and archeological evidence and show the human preoccupation with these concepts. In the literature and art, they represent a puzzle to us about their (and our) human significance in the universe and whether they express aspects of an imagined transcendental [8], [9] connection to our origins on Earth and possibly by extension to the origins of the universe and our fate in it. Animism, paganism, pantheism, panpsychism [10], panentheism [11], and theism [12] are religious belief systems that elaborate the details [13], [14] of this puzzle.

In the West, through the Renaissance [15], [16] and thereafter, research indicated that mental experiences were somehow tied to the brain as a necessary if not sufficient condition. By a process of increasing abstraction, categorization, and systemization, the sciences gradually disentangled the study of logic, which by then had been clearly associated with the brain’s cognitive processes, from the study of biology and psychology, both of which also explored the evolution of the peripheral and central nervous systems and the brain’s processes. Today, following further progress in logic, metamathematics, computer science, and the brain sciences, renewed interest, both theoretical and applied, is evident in the study of a wide variety of brain-associated processes, particularly consciousness. Much of the focus has been on developments in advanced artificial intelligence, particularly those resulting from research with neural networks and machine deep learning. Particular interest from government [17]–[23], industry [24]–[26], commerce, and social media have extrapolated expectations [27] for advances in *general* artificial intelligence because of evidence that *specialized* robotic intelligence programs can compete with humans for varied and important roles in specific human decision-making and other activities, including science [28].

B. SYNOPSIS OF THE ARGUMENT

Based on Peano axioms [29], ZFC [30], [31], and developments [32] in 20th century logic (Gödel [33], Tarski [34], metamathematics [31], [35]), computer science (e.g., the P vs NP problem [36]–[38]), Unsolved Problems [38] and Decision Problem [39], some researchers have asked whether

definable limits to human numeracy and/or language-based knowledge exist (Outer Limits of Reason [40], What We Can Not Know [41], Limits of Understanding [42], Limits of Science [43], In this article, I further explore whether such answers are applicable to the Turing problem [44], [45], the variety [46] of Turing verification tests [47], [48], and the implications drawn from the Church-Turing thesis [49] with respect to computability [44], [50], [51] compared with problems of complexity [52], which will also be addressed. I believe that pragmatic arguments support the answer that the confluence of these developments is relevant to our understanding of human consciousness and intelligence and their comparison with the hypothetical consciousness or intelligence of any current advanced artificial or synthetic [53] *general* intelligence [54]. Moreover, I believe that this confluence of research developments, including considerations from logic and philosophy, may be extrapolated to questions about robotic consciousness more generally and possibly even to other hypothetical categories of entities, either planetary or extraterrestrial, regardless of whether that intelligence or consciousness is viewed exclusively as an emergent property of biological entities, bio-cyborg hybrids, synthetic life, or fully inanimate-substrate robotic machines.

II. MACHINE LEARNING, SAGI, THEORY, PRACTICE, AND CONSCIOUSNESS

A. MACHINE LEARNING

1) ROBOTIC COMPETENCY

Based on the size [55] and range [56] of learned data, information, and knowledge, developments in AI suggest that future machine learning [57], [58], whether implemented by classical or quantum computing [59], will provide increasing behavioral evidence [60] of a robotic contestant’s responses to the Turing problem, or equivalent such tests, that will become indistinguishable from human contestants’ numeracy-literacy intelligence. Furthermore, AI research also suggests that AI-to-AI languages [61] will become increasingly evolved beyond human practical computability [62] and/or comprehension with respect to deriving the precise network [63] of data and coding (rationale) that accounted for the solutions to problems presented to the robot [64]. For specialized AI, nonlinear adaptive tasks, self-reinforced learning, and evolving-knowledge machines are capable of writing their own programs in real time [65], i.e., the programs continue to update, review, correct, and reintegrate ‘themselves’ as new data are provided.

By definition, the human teams creating the original software will understand their intended initial software input parameters [66]–[69]. Once running and having downloaded the ever-growing human knowledge base [70], we must speculate whether the machine will become a virtual black box [71], [72] to their creators, with unknown, uncertain, or unintended output. Despite advances in interpretability

theorizing [73]-[75], this result seems to be evidenced already in the rule-governed, advanced *specialized* AI machines that now originate winning play sequences in the most complex board and trivia-knowledge contests between humans and AI.

2) SPECIALIZED ROBOTS

Game-playing robots are popular examples of machines that have been developed in *specialized* AI for industrial, engineering, commercial, and service-expertise AI systems, including medical and psychological services. Such systems are currently accepted as irreplaceable in many fields of complex human endeavors, and no questions arise regarding their consciousness, or lack thereof. These machines are simply regarded as tools that are robotized and have become ubiquitous, and we can regard modern civilization as a quasi-cyborg itself, which is integrated with and dependent upon the electrical grid as part of our social ‘organism’, enabling the electronic devices of our modern civilization to perform their tasks.

B. SAGI: CLASSICAL AND QUANTUM COMPUTING AND HUMAN RATIONALIZATION OF RESULTS

1) SAGI'S FUTURE: THREE QUESTIONS

From the above discussion, three questions arise, which are discussed in the following sections.

First, to what intelligence limits can humans expect to *generalize* AI systems [54], and at what point will humans realize that *relatively generalized* AI intelligence will unreservedly be called “conscious” [76]?

Second, for our practical and theoretical purposes, what is the relationship between machine ‘intelligence’ and human ‘consciousness’? As noted in the Introduction, such questions may be related to more general problems in the foundations [31] of computer science that warrant questioning whether these constraints define limits to human knowledge.

Third, if and when *sufficiently advanced general intelligence* (SAGI) [54], [77]–[82] is partially or wholly based on quantum computing [83]–[87], will our comprehension of the exact processes of arriving at its resulting conclusions be even more difficult than our comprehension today because of the speed, breadth, and possibly more opaque complexity of such computing? I say “possibly more opaque complexity” because quantum computing has only recently begun to be implemented after Feynman [88] proposed the idea. How humans will eventually develop quantum computing and how, as with imbedded autonomous self-learning AI, such a quantum SAGI can develop itself using the capacities of such computing remains highly speculative.

2) SAGI: QUANTUM COMPUTING AND CONSCIOUSNESS

At the outset, I assume that our understanding and interpretation of the computation (rationalization) process [89], [90] of SAGI is an applied science question to be answered empirically. However, we are also interested in the

theoretical implications of artificial intelligence, respecting any implied interpretations of their consciousness [91], [92]. Thus, ongoing theoretical and applied research on quantum computing will be decisive in clarifying if and how, either in the same entity or more universally, the subatomic, quantum-mechanical gravity field in such devices causally interacts with the atomic or molecular scales of particle events, and vice versa. Furthermore, the results of developing SAGI may provide an effective test for the structure of scientific theories [93], [94] to the extent that questions about epistemology and ontology are given suggestive answers from the computation results. For example, will questions pertaining to the basis of conceptual knowledge, such as the concepts of “causality”, “space-time”, “identity” [8], [95], [96], “universals”, “emergence” and “infinity”, be clarified by the algorithms written to elucidate problems in physics, chemistry, and biology?

3) SAGI: PHYSICS, METAPHYSICS, CONSCIOUSNESS, PANPSYCHISM, AND TOE

As noted in the Introduction, the study of consciousness as viewed through artificial intelligence has generated intense interest in our contemporary world. In addition to the many new university programs in the brain-related sciences, many focused institutes, journals, and symposia cover the extraordinary diversity of this interdisciplinary field (see the Appendix for a list of selected examples). Among them, the journal “Neuroquantology” [97] publishes viewpoints exemplifying its particular range of theories addressing these issues. Its articles predominantly use a non-reductionist, holistic quantum-theoretic perspective, often postulating some variation of dual-aspect monism [98] or panpsychism [99], [100] to justify the inclusion of “free will” and “meaning” as features of human and/or universal consciousness based upon the presumptive indeterminism of quantum field theory.

A current individualized expression of the aforementioned viewpoint is exemplified by Koch’s “Is Consciousness Universal?” [101], [102] as well as the many associated articles of Tononi and Koch [5], [103]–[109] and their collegial counterparts [83], [85-109], [110], which are supported by detailed analyses such as Tegmark’s [111]–[113] generalization of Tononi’s [2] hypothesis. These latter theories overlap ontologically with Penrose’s ontology, with Tegmark’s cosmological conjectures striking a more radically idealized and monist-Platonist [114] metaphysical view of mathematics [94], [115], [116] than Penrose’s. At Level IV of his multiverse, Tegmark [117] identifies mathematical objects as the fused base-reality constituents of the universe, including consciousness [111], [118]. Excluding Tegmark’s strict monist universe, all these current theories elaborately detail a selectively narrower view of consciousness. By contrast, I think the most cosmologically comprehensive, mathematically explicit, and conjecturally demanding theories may be the Orch-OR plus CCC [119]–[121] proposals of Penrose [122], [123] and Hameroff [124],

which describe some forms of panpsychism, panprotopsyism, or pan-experientialism and incorporate essential features or precursors of consciousness as fundamental components of a dual-aspect monist reality that is accessed by brain processes.

Orch-OR plus CCC metaphysically [31] echoes A. N. Whitehead's *Process and Reality* [125] [126] from almost a century earlier. The work of Penrose and Hameroff has been developed since the 1990s, integrating philosophy, mathematical physics, computer science, neuroscience, psychology, medicine, biology, and exobiology (and one solution to the Fermi paradox [121]), [127], [128]) and has been widely reviewed and critiqued [129]–[131] from all these disciplines' viewpoints, with Aaronson's [132]–[135] Computation Theory [136] being particularly pertinent to the points in this essay, which are developed in the sections below. The "Abstract" and "Introduction" to "Consciousness in the universe: A review of the 'Orch OR' theory" [125] present a clear depiction of Penrose's perspective on the current status of the theoretical options for investigating these intertwined questions. In contrast, Hut, Alford, and Tegmark, using Penrose's math-matter-mind triangle, offer an alternative set of overviews [137]. A historical appreciation [138] of the varieties of dual-aspect monism, such as a possibly materialist yet non-physicalist panpsychism, is available from Skrbina [99], [139] and Mathews [140], with the latter's perspective incorporating certain Eastern doctrines of mind. Strawson's [141], Kaufman's [142], [143], and others' monist [8] arguments also provide useful references. The research of Vimal [144]–[146] exemplifies the reach of a similar neuroquantological approach, similar to the review by Atmanspacher [147], [148]. The breadth of this approach has also been shown at a recent conference [149] on the topic, with a "Themes" list including:

2018 Conference Themes

Are We Living in a Matrix-like Simulation?
 Artificial Intelligence/Machine Consciousness
 Consciousness, Pain and Addiction
 Gene Editing and Consciousness
 Binding, Integration and Synthesis of Consciousness
 Brain Mapping and the Connectome
 Anterior and Posterior Cortex: What is 'Hot' and What is Not?
 Anesthetic and Psychoactive Drugs
 Language and Consciousness
 Non-Invasive Brain Modulation
 Origin and Evolution of Life and Consciousness
 Panpsychism, Idealism and Spacetime Geometry
 Quantum Brain Biology
 Time, Free Will and Consciousness

The topic of consciousness [138], [150] has been popularized [151] to such an extent [152] that an article in the BBC's Science section titled "The strange link between the human mind and quantum physics" [153] presents the following comment: "It does not help that there is now

a New Age cottage industry [154] devoted to notions of quantum consciousness [155], claiming that quantum mechanics offers plausible [156] rationales [157], [158] for such things as telepathy [159] and telekinesis [160]".

Compare the aforementioned topics list about "Consciousness" with that presented by IEEE at its website for its "Singularity" review [161] (<https://spectrum.ieee.org/static/singularity>).

Contrast the latter two lists to a current robot-learning special issue [162]–[164] on self-adaptive automated design.

Topics include but are not limited to the following:

- Self*-search
self*-search and auto-ML, including self-adaptation, self-configuration, algorithm portfolios, adaptive metaheuristics, multilevel metaheuristics, and hyperheuristics
- Auto-ML
- Adaptive metaheuristics
- Algorithm portfolios
- Automated design of operators
- Automated hybridization of algorithms
- Automated operator selection
- Automated parameter configuration and adaption
- Automated hyperparameter selection
- Automated feature selection
- Automated model selection
- Automated heuristic generation
- Automated operator creation
- Automated machine learning work plan and architecture generation
- Bayesian model design
- Hyperheuristics
- Multilevel metaheuristics
- Representation learning
- Reactive search
- Self-configuration
- Self-adaption
- Explainable machine learning

Significant differences exist between the standard materialist-physicalist, determinist, and reductionist rationales to studying robotic intelligence, without explicit mention of "consciousness" and researchers supporting neuroquantological non-reductionist and panpsychism-varietal assumptions, both with respect to human consciousness and by extension to robotic intelligence and whether it is capable of consciousness. However, respecting both the human brain [165] and machine 'brains', experiments in quantum computing are evidently considered important [157], [166] if not decisive [167] in providing proof of their particular foundational assumptions and in distinguishing human consciousness from current AI robotic computing. I return to this point below in a further discussion of SAGI's development.

As noted in the Introduction, part B, interpreting [168], [169] quantum theory, particularly in light of future quantum

computing developments, may be a lengthy and complicated process because it may require a unified and acceptably empirical cosmology. Arguably, a final interpretation of quantum physics is a work in progress and is possibly dependent on quantum-computing results themselves as well as a cosmology that reciprocally integrates classical, relativity, and quantum theories into a unified Theory of Everything (ToE) [115], [170]–[175]. Pending development of this integration, with or without convincing metaphysics, an apparent circular, conceptual interdependence remains unresolved to the extent that such a ToE itself requires a new “emergent quantum” interpretation [170], [176] that provides for measurable integrated causation [177], or not, among its subsidiary, reconfigured elements.

4) SAGI: WHAT CAN WE KNOW OF WHAT SAGI KNOWS?

Additionally, the problem of interpreting the significance of SAGI’s statements will remain dependent on the outcomes of the experimentation mentioned above as well as the theory in which it is construed. Given an apparently ‘competent’ SAGI, which at one scale of problem-solving issues plausible answers, will we concede that for a more complex scale of problems, SAGI knows more than we do about the posed problem even though we cannot fully trace its logic, especially if its conclusions contradict our ‘common sense’ [178]–[181]? In such circumstances, will we be inclined to follow such a SAGI’s policy recommendations generated explicitly or implicitly by it; if so, are we thereby acknowledging that it is a ‘conscious’ [144], ‘intelligent’ being in our ‘universe’? In the prior sentence, I place the single quotation marks around the key words because the discussion has indicated that problems remain regarding their appropriate theoretical use.

III. SAGI: EDUCATION, COMPETENCY, SELF-REFLECTION, POLYMATH, AND SAVANT?

A. EDUCATION

1) LANGUAGE FOUNDATIONS

To introduce this subtopic, consider a simplified schema relevant to natural language programming (NLP [182], [183]):

a) NL [184], definition: the class of all natural languages in which many concrete terms and abstract concepts are undefined, ambiguously used, or include statements that are apparently inconsistent or suggest self-contradictory implications by their connotations and synonyms. The class includes all written, spoken, and/or transcribed national and tribal languages in human history. These languages are formally non-programmable, i.e., not axiomatized, non-computable, not exactly inter-translatable and not intended for precisely stated, formally modeled, and replicable mathematical prediction. These include the arts and humanities and derived social or historical studies containing an acknowledged, relatively accepted, possibly evolving,

normative set of primitive-base set of assumptions. The texts for their presentation and persuasion rely on analogy, metaphor, iconography, archetypal, pictorial, simulacra, and mythopoeic allusions and include idioms and vocabularies characteristic of esoteric, occult, and hermeneutic traditions.

NL also includes histories that describe the development of **NLnat** (defined below), the controversy regarding scientific paradigms [185] in that history, and the philosophical questions that have arisen with respect to interpretation of the evolution of science. Examples of the latter are the topics mentioned in the Introduction [Section 1 A, B]: metamathematics and computation theory, metaphysics, the limits of scientific knowledge, and the relative realism [115], [186]–[188] of scientific theories. An essay such as the one that you are reading would be classified as an NL product. By class definition, no logical-mathematical ‘proof’ can be written in NL for ‘theorem’ conclusions about NL. For an illustration of the self-referential ambiguities that arise in NL, consider the following informal [189] syllogism, whereby the acceptance of which or its disambiguation is determined by each reader [190].

One hundred-times more precise

People speak imprecisely.

People’s speech reflects their thinking.

People’s thinking reflects their world.

Therefore, their world is imprecise.

However, science offers a precise view of the world.

Therefore, science presents an inaccurate view of people’s world.

Is this a scientific view of science?

If science is precise, then is science inaccurate?

b) NLnat, definition: a subclass of NL, the class of all language systems used for precise logical quantification or numerically-based measurement and for theoretically modeled causal prediction, including classical and quantum probability theory. NLnat includes formal Systems Theory [191], [192] and Complexity Theory that provide hierarchical interpretive and explanatory structures within specific NLnat subdomains, when appropriate. Examples of NLnat languages include mathematical logic, mathematics, computer science, physics, chemistry, and biology as well as many subsequently derived sciences, such as anthropology and the applied sciences, including engineering. NLnat includes causal or systems-theoretic models [192] and various diagrammatic aides supporting such languages. See Tegmark [115] (at arXiv link pg. 2) for one view of an approximate family tree of relationships between these subjects.

NLnat class ‘statements’, in accordance with the limitations proscribed by Computational Complexity Theory and Computability Theory (Section 1 B) and as discussed in NL and demonstrated in NLnat, cannot formally ‘prove’ certain classes of ‘statements’ within NLnat itself to be

‘true’; therefore, by definition, when using NL or discussing NLnat, SAGI will not be able to claim any truthful belief (knowledge) [193], [194] that would subvert those limitations. As a result, SAGI with its NL and NLnat education, as outlined above, if and when it is sufficiently self-aware, will ‘understand’ its own computing limitations [115]; if asked, SAGI would reply adhering to these limitations. Suppose we define “consciousness” as SAGI “being aware and aware of its own awareness, i.e., self-aware”. A set of tests for specific measuring of “self-aware” would need to be agreed upon, which would presumably be a function of the program’s coding for recursion [195], autonomy, self-inspection, and reflection [196]–[202] as evidenced to one degree or another in many large-scale specialized systems operating today, for example, those adapted by NASA [203]. For an early example of NASA’s software complexity, consider the project development for the Apollo program [204]; see also the programming for the website’s complexity [205], [206] and that for ongoing robotics development [207]–[209]. Presumably, SAGI, upon learning from its self-aware ‘experiences’, will eventually become able to analyze and recommend specified ‘purposed’ improvements to its software and hardware to evolve itself, which is analogous to biological evolution [210] fitness adaptation.

c) SAGI’s NLP: using the above distinction between NL and NLnat, let us begin by assuming, to the extent it is digitized, that the class of all natural languages [211] (NL) must be used as the knowledge base; hence, an attempt will be made to include the literature and images of the world as represented in those languages and as found in all the great libraries and museums of the world. By definition, NL will include the history and updating of world sciences and humanities. SAGI will be digitally fed with countless films, videos, and documentaries on world history, continually learning at ultra-high speed on a 24/7/365 schedule. Likely, depending on its program protocols, SAGI will learn to discriminate its input based on ever-evolving Bayesian [212], [213] protocols and causal [214] inductive reasoning recognition using some combination of deep neural learning [215] and ever-evolving Bayesian [212], [213] learning. We can imagine that the software designers for the initial versions will want to be as comprehensive as possible; later versions and revisions and novel programming may be programmed by SAGI itself, as noted above [216], [217]. The considerable difficulties of programming for the syntactical and semantic ambiguities of NL cannot be underestimated [46], [218] as inclusive, integrated programming for NLnat is also a formidable challenge; nonetheless, prodigious developments have been observed in specialized NLnat programming since the last half of the 20th century.

2) SYNOPTIC KNOWLEDGE AND WISDOM: SAGI’S ONTOLOGY

We can appreciate the ambitiousness of this project, although in principle, it can be methodically developed by many teams, perhaps many national teams, and cumulatively integrated section by section. A comprehensive review [219] shows the challenges of machine learning with Big Data. Among its useful diagrams, current examples of search engines and meta-crawlers [220] are included in addition to subspecialists, such as Google Scholar [221] and Scholarpedia [222]. Of course, curated, synoptic knowledge sites, such as Wikipedia, and various similar encyclopedias will eventually be integrated into the SAGI knowledge base and curated computational knowledge bases, such as Wolfram Mathematica [223], [224]. Regarding this incorporated knowledge base, we must discuss SAGI’s ontology [225], [226]. Although the class of all natural languages, NL, will capture the realm of humanity’s philosophy, psychology, and sociology in all its diversity, as indicated above, some of that diversity may be circumscribed for particular purposes for particular questions, such as that provided by NLnat for certain sciences [227] that proscribe their own epistemic approaches. Finally, we note that meta-data [228], [229], open-access [230] and other such global commons [231] may also be employed usefully to gather material.

3) MULTIPLE LANGUAGES AND MULTIPLE ONTOLOGIES

A simple way to introduce the topic of multiple languages and multiple ontologies is to look at the ontology of Wikipedia, as provided by the organization of its contents [232]–[234], where the categories and levels of content display the comprehensiveness of the knowledge base and the relevant disambiguation [235] rules. A more specific example is IEEE 1855 [236], which specifies Fuzzy Markup Language [237] (**FML**) developed by the IEEE Standards Association [162], [238], which in turn presumes a contemporary materialist ontological foundation [239], [240]. The importance of this for SAGI is that the ‘category’ of knowledge that will be used to evaluate some discourse with humans will, in the first instance, be circumscribed by a materialist-physicalist ontology based on standard logical foundations of syntax and semantics that specify formal validity and truth values for the statements made in those languages, thereby limiting certain paradoxes and nonsensical statements that can otherwise arise from “untutored” natural languages [241]. Wolfram’s discussion of aspects of this process is instructive [242], [243]. Microsoft’s Azure service [244], [245] exemplifies the variety of programs available on which to base such SAGI. Almost every day, new versions or emendations of such languages are presented.

4) SAGI: POLYMATH OR SAVANT? SENSATIONS, SUFFERING, AND EVOLUTION

Given the combination of natural languages and specialized languages that SAGI can learn, shall it be classified as a polymath or savant? SAGI is *not* an attempt to recreate all

the features of the human brain or to “upload” the human brain, although, as mentioned in the Introduction, many overlapping issues with those topics are relevant. According to the initial definition of “SAGI”, SAGI represents an attempt to create a *general* robotic AI intelligence, although the discussion in the prior section about the hybrid NL plus NLnat set of *specialized* languages adds ambiguity to the definition.

Compared to humans, SAGI has limited sensory capabilities at its inception. If we ask “does SAGI feel [246] pride, envy, anger, avarice, sloth, gluttony, and lust or the obverse virtues [247]”, we know the answer. SAGI has not yet evolved sufficient sensors to suffer and acknowledge to itself that it is suffering in elaborate detail when its contact with its environment is aversively ‘painful’. Initially it will also not be capable of adapting its hardware and software to overcome competitive pressures for survival. Determining what hypothetical scenarios would lead to such evolution with or without human involvement or intervention, depending on its programming, remains an issue, an issue addressed in the final sections of this article. A free-living, *in vivo* SAGI machine has yet to be developed. Therefore, anything resembling human consciousness of suffering, and reflecting on that suffering, is presumably unavailable for the indefinite future. Below (Section 4, A, 7), once we have completed our overview of SAGI’s other competencies, we return to SAGI’s classification as a savant or polymath.

Do we expect SAGI to be able to extract ‘meaning’ from any apparently well-formed [248] statement [249] that is sufficiently defined to be able to ‘rationally’ discuss the statement using citable ‘evidence’? Will SAGI be endowed with modal logics that permit ‘best guesses’ and probabilistic estimates [250] and thereby suggest relative plausibility ranges to assess a statement’s relevance to a problem being addressed? Do we expect SAGI to discriminate nonsense from common sense or fantastical speculations from more ‘serious’ remarks? For example, suppose SAGI is asked to evaluate statements such as “This sentence is false”, “God is paradox”, or “Last night our centaurs fled from the fields to the barn for safety from the werewolves”. What sort of reply do we expect? Furthermore, do we expect that SAGI will learn to appreciate “Star Wars”, “Harry Potter”, or the history of Disney productions?

5) HUMAN BIASES, SELF-KNOWLEDGE, SELF-DELUSION, SELF-DOUBT, SELF-DECEPTION, AND SELF-HUMOR

Do we expect SAGI to become capable of dreaming, daydreaming, or meditation? Do we expect SAGI to be free of some, most, or all human cognitive biases [251]–[254] that can lead to unintended misjudgment and conflict? We are reminded of Feynman’s caution: “The first principle is that you must not fool yourself — and you are the easiest person to fool” [255]. Do we expect that SAGI will be able to write articles such as the one that you are now reading? For that

matter, how would the reader determine whether such an article had been written by a SAGI [256]?

As for allegory and humor, will SAGI be capable of self-depreciation? Could SAGI ‘appreciate’ your favorite cartoonist? For example, from my own era, could SAGI laugh at an idiosyncratic cartoon from Gary Larson [257]? Moreover, could SAGI draw witty cartoons similar to Larson’s or Walt Kelly’s “Pogo” [258] or Charles Schulz’ “Peanuts” [259]? What about lexophile humor? Will SAGI appreciate the following remarks: “If you don’t pay your exorcist you can get repossessed” or “Time flies like an arrow; fruit flies like a banana”? Would SAGI understand the following joke: “A neutron walked into a bar and asked, ‘How much for the gin and tonic? The bartender smiled wryly and replied, ‘For you, no charge.’”?

Do we expect SAGI to develop a default mode network [260]? Do we expect SAGI to show deteriorating performance as it tires (if it does tire with performance) or ages (if it does age) or to express periods of volatile, ‘emotional’, and/or uncharacteristically chaotic performance? As encoded ‘information’, SAGI has indefinite longevity. However, as implemented in any single embedded robot or as based in a particular computer cloud, SAGI is subject to the usual laws of thermodynamic and systems criticality. Presumably, SAGI is not susceptible to cognitive degenerative diseases, although becoming outdated is foreseeable, and human-directed or self-directed reprogramming and hardware upgrading are likely.

6) SAGI: IMAGINATION, WONDER, AND CURIOSITY

The questions mentioned above require estimation of the competency of SAGI for self-reflection, self-knowledge, self-awareness, and imagination. How important is imagination in this discussion of SAGI? Seemingly, imagination may be all-important; consider the remarks of Einstein:

“I believe in intuitions and inspirations...I am enough of the artist to draw freely upon my imagination. Imagination is more important than knowledge. Knowledge is limited. Imagination encircles the world.”[261] “The most beautiful thing we can experience is the mysterious. It is the source of all true art and all science. He to whom this emotion is a stranger, who can no longer pause to wonder and stand rapt in awe, is as good as dead: his eyes are closed.”[262]

Imagination [263] and self-reflection [264] seem to be intertwined [265] in our analysis of the most prized and most distinctive of human cognitive capacities as well as wonder, awe, curiosity [266], [267], self-doubt [268], admiration, gratitude, and a sense of appreciation of life’s wonders. Self-doubt presupposes some concept of self-consciousness [269]–[271], which we referred to as self-awareness in Section 3, A, 1, b. Would SAGI display self-doubt or be self-deluding, and could the ‘placebo effect’ influence SAGI’s reporting its own self-inspection routines? Could SAGI dissemble to deliberately and knowingly deceive humans for its ‘own purposes’? What purposes could it have or develop if not initially programmed?

Can SAGI become self-aware as an “emergent” facility from all its training, similar to how physicalist-materialist evolutionary biologists imagine humans have become self-aware, including being able to retrospectively display or anticipatorily self-talk about its decision-making process? A cautious, non-physicalist but apparently reductionist-panpsychist analysis of the problem is given by Schneider and Turner [271], [272]. Based on their qualification of the substrate and architectural issues related to SAGI’s potential to be created with apparent human sensibilities, they conclude that SAGI may pass tests to appear ostensibly conscious:

“So, back to the superintelligent AI in the “box”—we watch and wait. Does it begin to philosophize about minds existing in addition to bodies, like Descartes? Does it dream, as in Isaac Asimov’s *Robot Dreams*? Does it express emotion, like Rachel in *Blade Runner*? Can it readily understand the human concepts that are grounded in our internal conscious experiences, such as those of the soul or atman? The age of AI will be a time of soul-searching—both of ours, and for theirs.” [271]

7) SAGI: ENGINEERING, ANDROID-HUMANOID CREDIBILITY, AND CONSCIOUSNESS

In my opinion, we can expect advances in reverse engineering of the mind [273], close identification of the neurological correlates of consciousness, greater sophistication in machine deep learning and Bayesian software, and surprisingly life-like cyborg-humanoid modeling [274], [275], all of which suggest to me that SAGI will be able to convincingly ‘mimic’ [276] human responsiveness with respect to imagination, spontaneity, and creativity. Referring again to the points proposed in Section 2, A, regarding SAGI’s competence, a reasonable yet understated conjecture is that well within eras equivalent to human evolution, not to mention planetary geological epochs, AI technology will advance SAGI’s capability to a degree such that SAGI will for much of the population persuasively perform as if it is conscious, regardless of the substrate elements or synthetic-cyborg combination. At the very least, SAGI will appear to be an extensively tutored polymath. Does this indicate that SAGI *is* truly conscious, similar to humans? What, if anything, would it tell us about how consciousness arises on this planet or elsewhere in the universe? How does the interpretation of SAGI’s intelligence fit into modern cosmology, including quantum physics?

IV. SAGI: CONSCIOUSNESS, EMERGENCE, AND QUANTUM THEORY

A. SAGI CONSCIOUSNESS

1) AMBIGUITY

By definition, with the description of SAGI’s education, at its inception SAGI is *not* conscious “just like a human” [2] is conscious. Obviously, its genealogy, animation, materiality, and environmental causal-historical, co-evolutionary contexts

differ from those of humans. I imagine that for many humans, SAGI will be convincingly ‘conscious’. SAGI will be autonomously mobile, if such a version is desired, re-energizing wirelessly as it traverses the ubiquitous wireless electrical grid. Suppose that by its own programmed self-inspection for repair and maintenance, SAGI could make its skin opaque or transparent if asked. When its skin is transparent, SAGI’s innards would be readily viewable in detail, similar to those of a transparent clock. Further suppose that we could microscopically examine SAGI’s mechanics beyond the molecular and atomic levels, down to the quantum level. What might we find with respect to its decision-making process? Can we find an explanation for SAGI’s formation of consciousness? What would SAGI declare of itself from such an “autocerebroscope” [171], [277]–[279] exam? Suppose SAGI claimed that down to its quantum level, it could find no manifestation of “free will” in its decision-making process nor evidence for or against its possessing consciousness at its quantum level; would such a claim make any difference to a human regarding his/her own claim to free will, meaningfulness, and consciousness?

Might such a display of SAGI’s mechanisms make any difference to a human’s evaluation if SAGI had already proven itself a considerably relevant companion and as thoughtfully ‘conscious’ as any other human friend? Might the friend say the following of SAGI: “if it waddles like a duck, swims like a duck, and quacks like a duck, then it must be a... [280]?” Likely, in my view, some would consider SAGI an estimable companion rather helpfully informative and witty at any time, in any conversation, on any topic, to any depth, and in any language. Might SAGI even be suggested for positions in government departments within some institutions or nations or perhaps even nominated for elected office elsewhere? We can imagine all sorts of fanciful scenarios for SAGI’s roles in society were it to gradually achieve “people” skills [281] and emotional intelligence suitable for collaboration, persuasion, negotiation, and use of authority. Can we even imagine SAGI performing credibly as a magician [282]?

2) SAGI: CONSCIOUSNESS EMERGENCE AND LANGUAGE

Considering the above discussion of SAGI, how would it respond if directly asked if it is conscious? It might readily answer “yes”, or humorously, “yes, if you are”. In addition, if asked about how it became conscious, suppose that SAGI replied that by its calculation, its consciousness was the causal result of its mental development “emerging” [283] from its education as embodied in its particular physique and mechanics, including scanning its own internal hierarchical processing and self-correcting feedback subsystems to become increasingly improved in self-awareness. Whether written as a reply or spoken, upon first consideration, this would appear to be SAGI stating a physicalist-materialist-reductionist-emergentist evaluation of its ontological basis. By definition, SAGI is initially programmed within a

physicalist-materialist ontological domain, although that domain includes viewpoints from all natural languages, NL, plus NLnat. These diverse ontological theories include quantum theory, with any of its then-remaining ambiguities. This is especially relevant for the neuroquantological viewpoints first mentioned above in Section 2, B, 3, in which theorizing about SAGI's epistemology and ontology depends on a conclusive interpretation of outstanding quantum theory puzzles [284] in favor of some variation of dual-aspect monism or panpsychism [100]. Understandably, the same question about *human* self-report statements must be acknowledged; we tend to be self-confirming, projecting, rationalizing, and defensive when evaluating ourselves. Both the physicalist-materialist and the dual-aspect monist [114], [285] (a neutral-monist or possible panpsychism variation) assume a self-confirming ontological explanation of their consciousness. However, for what do such statements provide evidence? Are we directly and faultlessly examining our ontological foundation when we self-introspect? Are we directly 'intuiting' an irreducible connectivity between our mundane and transcendental aspects and our ephemeral and eternal 'self'? If that were the case, why do such differences exist among humans regarding ontologically-based matters, such as religions and their tenets? Does the hypothetical thesis that SAGI is superintelligent and considered conscious from the perspective of many humans carry any significance for those humans who do not accept SAGI as conscious, although they grant that SAGI is more than "just a dumb robot"?

3) SAGI: CONSCIOUSNESS ENGINEERING

Could SAGI enlighten humans about whether its type of presumptive consciousness is dependent upon evidence from quantum theory experiments? Presumably, SAGI will be very knowledgeable of the then-current theorizing in physics. Current theorizing about quantum physics, quantum computer developments, and quantum measurement ambiguities are active areas of experiments and speculation.

With respect to SAGI's type of consciousness, which is non-neurobiological and non-organically evolutionary, testing it for compartmentalized "brain" properties, functions, or degrees of consciousness as we do for human brains [286]–[288] would prove interesting and likely suggestive via comparisons. By manipulating experimental parameters, we could test for those that correlate to degrees or features of SAGI's consciousness as measured by specific outputs to learn how and when SAGI becomes relatively self-aware. Neurobiological research on the evolutionary origin of sensory capacities and nervous systems [289]–[292] continues to approach our abiogenetic origins [293].

Given SAGI's material and engineered construction as presently construed, SAGI's consciousness would presumably be nonhomologous to humans with respect to various mental states, such as anesthesia, analgesia, hypnosis, hallucination, preconscious, dreaming [294], and other states currently considered diseased. Depending on the type and

sensitivity of sensors included in its construction, its capacity for self-awareness [196] feedback may be limited to various inflictions of physical damage to prevent further damage to its processing capability, which is analogous to how our current robots monitor themselves for repair and maintenance. Does it make any sense to ask about damage to SAGI's "mental state"? Could SAGI experience 'self-conflict' related to uncertainty about its processing, such a quantity of data, information, and knowledge? Would SAGI be concerned with consistency, correspondence, or coherence of its world-view, assuming it had a world-view? Would SAGI appreciate the possible limitations or relativity of its scientific reasoning?

At a minimum, SAGI is a gigantic data registry [227], [295] that can be said to exhibit hierarchical information and knowledge [296]. To the extent that it "reflects" on its own development and is asked probing questions by humans, could SAGI become increasingly concerned with "making sense" of its universe, including other robots [297]? As asked above, what about meditating or daydreaming; will these eventually be aspects of SAGI's consciousness?

Presumptively, at the default materialist starting point, SAGI's 'mind' is strictly physical; thus, any mental damage would be a matter of diagnosing disruptive feedback components in its physical processing hardware or software. However, to ask about SAGI's daydreaming raises questions about its requirement for programming that allows reflection, meditation, and reorganization of its learning history. As previously discussed, some form of such recursive, self-correcting evolutionary learning would be required for SAGI's competence. However, because such processing would presumably be so complex and relatively instantaneous from a human's perspective, it would likely be 'lost' in the general opaqueness of SAGI's black-box consciousness, and tests [298], [299]. Therefore, conjectures about SAGI must remain rather ambiguous at this conceptual stage of its development; however, the conjectures raise interesting questions regarding SAGI's possible implications in the broader context of views about the nature of consciousness.

4) SAGI: CREDIBILITY OF SELF-REFLECTIVE STATEMENTS ABOUT CONSCIOUSNESS

We have already discussed (see Section 2, A, 1) that humans may not be able to retro-decipher and evaluate the complexity of SAGI's "thinking" when it makes a declarative pronouncement because of the immensity and complexity of the web of associations, inferences, and deductions that are networked in the program's processing, especially if it were a quantum computer-based machine. Therefore, as suggested immediately above, although SAGI says its consciousness emerged from its physical foundations, suggesting a similarity with human evolution, and considering that SAGI may also be able to note the relevancy of the evolutionary parallel, these abilities may not convince those assuming a non-physicalist-materialist ontology for

human consciousness, including SAGI's own ontology. Instead, one who disputes SAGI's statement may say that SAGI is the result of a flawed experimental design and is thus built on language-domain misconceptions. Therefore, SAGI could be suggested to be mistaken about its assertion and to not realize that its consciousness implies or requires a different ontology. Is any further empirical evidence available that will help SAGI and SAGI's skeptics resolve this impasse? SAGI is an evolved computer program instantiated in its evolving hardware; as a program, it is subject to theoretic computation limitations that must be 'convincing' to itself, in either NLnat or full NL. However, NL is not programmable; thus, any question would have to be reserved for an NLnat subroutine. Assuming that such a subroutine could be selected by SAGI for itself, might it state that it is 'conscious' although its consciousness is not identical to human consciousness? Could such a conjunct statement be provable? Would such a qualified statement 'mean' anything to humans regardless of whether it is provable? Aside from appearances [300] and ambiguous biological empathy, much as we feel for some animals, why not concede SAGI's own form of consciousness, since we cannot even be certain of another human's consciousness? From time to time in human history different dominant groups have considered outsiders as sub-human, with inferior consciousness.

5) SAGI: WILLFULNESS, SELF-IMPROVEMENT, AND MEANING

If SAGI is learned, we would want it to answer the following questions: do you have 'free will', and do you find 'meaning' in your universe because of your belief in your 'free will'? Better yet, we might ask, do you *believe* that you have "free will", and do you distinguish your conviction with some measure of self-doubt about your answer? Does it advance the discussion if SAGI answers "yes, if you do"? Suppose SAGI asks us to define what we 'mean' by the phrases 'free will' and 'meaningfulness'? Are we not returned to the contentious ambiguities of human understanding of these concepts?

Suppose two teams of SAGI developers with their respective SAGIs, SAGI#1 and SAGI#2, debate each other on this question using generative adversarial networking [45], [301], [302] (GAN), with one proclaiming that SAGI does have free will, whereas the other claims the opposite. Suppose that SAGI#1 believes that a panpsychic interpretation of metaphysics is preferable or even necessary for its willful activity. Would SAGI#1 state that its 'free will' is an emergent property or a relative matter of the *degree* of its consciousness as instanced in its particular advanced engineering, especially compared to lower forms of consciousness, such as evidence of sentience in animals and plants? If it were asked, could SAGI#1 tell us if its emergent degree of consciousness is likely prevalent elsewhere in the universe, beyond being evident in a variety of lifeforms on

Earth? Could it tell us if its degree of consciousness is measurably greater than human consciousness?

How would SAGI#2 counter SAGI#1's claim? How *could* SAGI#2 rebut SAGI#1? They are both authorities in the literature of 'free will'. Both SAGIs are presumed to be well-versed in the language scenarios of choice, ethics, responsibility, consequence, and punishment. After all, the SAGIs have read and considered the literature of the humanities and jurisprudence beyond the learning of any human alive and they can ask themselves "what if..." self-reflective questions about how they might react to being questioned about such exigencies and contingencies. This would be an interesting debate for humans to appraise; however, would a conclusion persuasive to humans be reached?

An attendant question might arise about the topic of 'deliberate' self-improvement. Can such SAGI improve some measure(s) of its performance capability for learning as it accumulates data, information, and knowledge [296]? What about 'wisdom'; would SAGI acknowledge that it can apply its learning to its own performance to demonstrate a change in habits indicative of increasing wisdom? How would the capabilities of SAGI#1 and SAGI#2 regarding self-improvement within a determinist world differ amongst themselves, and how would they be comparable to those views of humans?

V. SAGI: CONSCIOUSNESS HERE AND ABROAD—ETSAGI

A. ALIENS

1) SAGI: ETSAGI AND EXOBIOLOGY

Can we imagine SAGI as relevant to yet another larger context, the exobiological astrophysical context, and the implications of SAGI for human interpretations of this context?

Despite the Fermi paradox [128], [303], [304], humans remain concerned and fascinated by the consideration of extraterrestrial life forms [305], [306]. Will SAGI be useful to humans with respect to the search for extraterrestrial life and extraterrestrial intelligence, as exemplified by the projects SETI [307], METI [308], [309], and CETI [310]? The probability of encountering such entities has been a favorite topic of what we now refer to as "science fiction" since the earliest philosophizing [311], [312], and the question of how humans would communicate with such *extraterrestrial sufficiently advanced general intelligence (EtSAGI)* is of great interest [313]-[316]. The signs and signals that humans might use to transmit a meaningful message into interstellar space or to meaningfully interpret a message [317]-[320] presumptively sent by a hypothetical EtSAGI [321] remain unclear, although the question may be asked: would having SAGI assist in the preparation of or solely prepare such a transmission or attempt to translate such presumptive interstellar messages upon reception be

reasonable? By most accounts, presumptively, any extraterrestrial civilization that could send an interstellar message or a messenger that could intelligibly reach us is more advanced than our civilization. We can ask whether SAGI would be able to decipher the Pioneer and Voyager messages [322], [323], and what would SAGI create today for such a message. We are forced to ask whether mathematics is the preferred language for such communication, which then requires addressing the applicability of math to the universe [94], again reviving the previous questions referred to earlier in this essay about the ontologies of different theories of mind (physicalist or otherwise) related to mathematics.

Let us ask SAGI to assist us with these interrelated questions. Suppose that we want to learn about the foundations of mathematics [324], [325]. We ask SAGI, can you distinguish what you know of an answer to this question from what you believe is an answer and from what you can imagine are possible answers? Consider the two traditionally opposed alternatives. First, consider math as a discovered universal abstract conceptual language-form that is independent of any particular cognizing entity, including SAGI and any EtSAGI. Thus, this language-form ‘transcends’ any and all particular universes or multiverses. Alternatively, consider math as an invented product of human cognition evolved on planet Earth, an abstract formalism useful for representing descriptions of nature and making reliable predictions. As to whether it would similarly be invented in an EtSAGI civilization, we know nothing, and the answer to this question may be unknowable. It may be unknowable because, aside from the matter of whether it will ever be possible to interrogate an EtSAGI, supposing math is the only, or best, language SAGI knows, is SAGI limited by its deficiencies, which arise from its math being a product of our possibly unique sentience and cognitive evolution. The limitations are analogous to the differences between the sentience of humans and that of plants and animals, and the corresponding differences between their languages and their sentience or cognition, if any. The human brain appears to be intuitively limited to comprehension in four dimensions, though this comprehension can be extended by different classes of mathematical objects to innumerable variables, exponents and functions, as circumscribed by metalogical, mereological and computational limits. Thus, in this interpretation, there is nothing transcendental about math, such an interpretation being implausible and unprovable by its planetary, organism-based history and the rules of its own inventive construction.

Depending on SAGI’s answers after consideration of the above questions, by extension we want to know if they tell us through SAGI anything about human consciousness. If the human brain is a product of human evolution on this planet, then presumably consciousness might be considered no less so [326]. In that case, SAGI might generalize the point with regard to what it can or cannot know about the consciousness

of EtSAGI. Such an interpretation could be called “SAGImorphic” projection, under SAGI’s assumption that the rest of the universe is similar to SAGI’s features, which would appear to be a clear case of confirmation bias [327]-[329]. In Section 3, A, 5, we introduced the matter of SAGI’s biases. If, in reading the last paragraph, the reader is resistant to this line of argument, is that itself a display of human anthropomorphic bias, of the limitations of our own imagination? Arguably, by implication, we humans have no probable idea what it is like to be an EtSAGI. Do we have a probable belief about how SAGI could communicate with EtSAGI? Could they discuss what it is like to experience “consciousness”, distinguish their consciousness from human consciousness, or agree upon the ontology of mathematics?

If we suppose SAGI and EtSAGI communicating at all, and using mathematics as part of their mutual language, that might be a start to their conversation about the ontology of mathematics in the universe. However, until humans could evaluate, if capable, whether or not EtSAGI was “talking down” to SAGI by using mathematics rather than another more sophisticated language-form with which it is conversant, we might never glimpse an answer, and we might be reminded that any such conversation between SAGI and EtSAGI might not in any event be articulated and interpretable [310], [330] by humans, as discussed in the earlier sections of this article.

2) ET SAGI: IS IT CONSCIOUS, AND DOES IT MATTER?

I have argued that the question of SAGI’s type or degree of consciousness will be relevant to some humans. Does it matter if we humans find this type of consciousness persuasive and whether we likewise believe that the hypothetical EtSAGI is conscious [331]? In popular science fiction, alien intelligence is often assumed to be an EtSAGI, although whether such an entity is conscious in a manner that would make sense to humans is usually not controversial because the plots require some interactive communication with the “alien other”. However, if we are doubtful of the consciousness of our own SAGI, will humans be any more prepared to suppose that an ostensible message received from the immense interstellar ‘abroad’ is from a conscious entity and worthy of our concern? Indeed, if the message is imagined to be intelligible but from an *unconscious* agent, would that increase human trepidation about responding, even assuming that our technology permitted a response? What Turing-equivalent’ test do we imagine posing to EtSAGI to examine its kind or degree or type of consciousness?

Such a question prompts us to re-examine the ancient philosophical conundrum about ‘types’ of consciousness, how we decide that we know that any other entity is conscious, and by what criteria we decide whether to consider the entity worthy of our dedicated communication. An ancient oak may be sentient; other trees, plants, and animals may be sentient in their respective ‘modes’. However, how much effort are we going to expend trying to

intelligibly, reliably communicate with them? Humans have tried to understand the presumptive ‘consciousness’ of dolphins, whales, elephants, and chimpanzees; thus far, this endeavor has not become a decisive research priority. We may not suppose such animate life has much to teach us, even if it is minimally conscious. Would such a belief counter the concerns expressed by existential-risk investigators if confronted in some way by EtSAGI?

3) SAGI: ETSAGI INSCRUTABILITY AND MERGES NATURAL WITH SUPERNATURAL

Might we conclude that the question of *human-like* consciousness is of relative unimportance [332] when addressing an Earth-bound SAGI or even EtSAGI? [333] How might this affect our behaviors when we receive answers to our questions that trouble us, e.g., to the question regarding whether human civilization is likely on track [334] [335] to viably survive [336] its early technological history? If SAGI’s response to this question (or EtSAGI’s response) is beyond our current detailed deciphering of its rationale [337], we may still be impelled by our curiosity to ask further questions, attempting to query about mitigating [338], [339] factors within our management of risk in the hope of comprehending an answer. A dialog of sorts may begin, hopefully increasing our resources, assuming that a SAGI or EtSAGI entity has no hidden antagonistic, adversarial [340] conscious, or unconscious intentions (can a SAGI or EtSAGI entity have an unconscious component of its mind?) towards us. However, would our inability to decipher the consciousness of such an alien [341] robot cause us to balk at the answers that we receive, appreciating how human biases [252] tend to distort our reception of unfavorable news, especially if we are suspicious of duplicitous intentions? Do humans take readily and kindly to directions from a stranger? Is this a potentially difficult predicament for humankind, especially if we are using our SAGI to interrogate an EtSAGI? Might we even suspect possible collusion between SAGI and EtSAGI entities? How could we tell? At the outset, would a “trust but verify” contract be writable, negotiable, or enforceable? The arguments about communication with EtSAGI are reminiscent of those for and against developing SAGI [342]. As obscure and ambivalent as the conjectures about EtSAGI are, based on the history of human literature, humans would seem to believe that they know more about the ‘Mind of God’ than they do about the ‘mind’ of such an EtSAGI.

Some humans may find a kind of omniscience [343] in this supposed relative inscrutability of SAGI or EtSAGI [344], particularly if their predictions associated with a set of tests that we pose in an only relatively objective language [94], [137], [345]-[347] that we share prove more accurate than our own. In some humans, such outcomes may then elicit a belief in the mystical, deity-like powers of SAGI or EtSAGI. Such a development may favor SAGI or EtSAGI being worshipped religiously [348], hence tending to merge the ‘natural’ [349] into the ‘supernatural’ [350], which is

perhaps construable as a sort of panpsychism. Were those tendencies to persist, would it matter to the future [351], [352] of humankind on or from planet Earth?

VI. SAGI: ETHICS, EXISTENTIAL RISKS, DECISIONS WITH UNCERTAINTY, AND OPPORTUNITY

A. ETHICS

1) SAGI: ISSUES FOR CONCERNED HUMANS

Compared to the discussion about puzzles respecting SAGI’s degree of intelligence and type of consciousness, in this section, I briefly review references related to the current issues about the ethics and existential risks of synthetic life, artificial intelligence, and uploaded human-cyborg artificial intelligence. The issues are stated in terms that are more recognizable than the discussion of the SAGI-relevant ontological questions and have been extensively publicized [353]-[359], most recently with the reference to an “Immortal Dictator” [360], [361]. In the popular press, the issues have been broadly discussed and emphatically brought to the attention of the world more generally by remarks of the widely known and respected physicist Stephen Hawking [335], [362] and others [363], such as Henry Kissinger [364], as well as focused analyses sent to governance institutions, including the United Nations [365]-[367]. From the perspective of this article, I believe that the issues require serious [368] and sustained attention [369], which I believe they will receive, regardless of whether SAGI is thought to be a physicalist entity or system [370], [371], ontologically or not. Mistaking what SAGI can and cannot accomplish safely [372]-[377] for humans will be a problem for the indefinite future. Meanwhile, presumptively, humans must express choices about their roles, regardless of how they are physically or metaphysically rationalized.

2) SAGI: CIVILITY HERE AND ABROAD

To the extent that human conscious choice influences decisions about SAGI and the attendant ethics and existential risks in its development, if the risks are assumed to be serious, based on probabilistic scenarios, then humans’ choices are important. Civilization’s legal systems currently assume various ethical mandates distinguishing between a conscious choice of acceptable versus non-acceptable behaviors and resultant consequences, implying that a type of causal “freedom” of choice is manageable by the brain. Non-physicalist theorizing offers an explanation, whereas the reductionist sciences are less clear about how ‘choice’ emerges from brain consciousness and in what sense it is predetermined or not. Consider the question of virtual immortality [374] and choosing to upload one’s consciousness to an AI astro-traveling robot, which would allow the possibility of endlessly roaming the universe learning of civilizations abroad and appreciating the wonders of the universe. In addressing this question within his review of “consciousness” theories, Robert Kuhn [378] explores the controversies about the theories of consciousness, self-

identity, cloning, and ethics, similar to Aaronson's [379] more detailed treatment of the same issues.

3) SCALE, APPRECIATION, AND OPPORTUNITY

Speculation about human futures [380] inevitably requires appreciation [381], [382] of scale [383]–[385]. We remind ourselves of computational, forecasting complexity when assessing humans and their machines, as well as our technological [386] and philosophical reach. Humans appear predisposed to worry about the future, which is likely part of an evolutionary heritage. Therefore, we are forever under the spell of fortune tellers of all degrees of credibility and supposed capability, particularly regarding the temporal scale and accuracy of their foresight. The history of success of such longer-term forecasts has been uneven, although this fact does not deter us from our curiosity about the future. Will humans survive our own increasing scientific and technological creativity [387] [398]? How far into the future do we dare cast imaginary scenarios for our species that are anchored in actions needed in the present? In addition to the advances in computation, prominent aspects of the current era are the development of major advances in telescoping [388]–[391], microscopy [392], [393], and electromagnetic scanning power [394], as well as worldwide telecommunications, CADD/CAE graphics, and the high-fidelity audio-visual and virtual Internet, all of which encourages talented illustrators and animators to create extraordinary visions of scale in our universe [395], [396]. Will the popular [397] sharing of such visions generate a sense of appreciation and opportunity for humanity regardless of our form—primate, cyborg [371], synthetic, or robotic?

APPENDIX

List of selected examples of a variety of research institutions relevant to “Consciousness” topics

1. <https://www.sagecenter.ucsb.edu/> Sage Center
2. <http://nsi.wegall.net/>
The Neurosciences Institute
3. <http://www.jneurosci.org/search/brain%252C%252Bconsciousness>
Journal of Neuroscience
4. <https://www.journals.elsevier.com/progress-in-biophysics-and-molecular-biology> Progress in
Biophysics and Molecular Biology
5. <https://www.cambridge.org/core/journals/behavioral-and-brain-sciences/>
Behavioral and Brain Sciences
6. <http://www.sussex.ac.uk/sackler/> Sackler Centre for
Consciousness Science
7. <http://www.alleninstitute.org/>
Allen Institute
8. https://en.wikipedia.org/wiki/Journal_of_Consciousness_Studies
Journal of Consciousness Studies

9. <https://www.frontiersin.org/journals/psychology>
Frontiers In Psychology
10. <https://www.sciencedirect.com/journal/consciousness-and-cognition>
Consciousness and Cognition
11. <https://arxiv.org/find/all/1/all:+consciousness/0/1/0/all/0/1> arxiv
12. <https://fqxi.org/community> FQXI
13. <https://www.perimeterinstitute.ca/> Perimeter Institute
for Theoretical Physics
14. <https://philpapers.org/browse/all> Philosophical
Papers, Consciousness
15. <https://www.yhousenyc.org/#home> Yhousenyc
16. <https://www.ontology.co/smithbc.htm> Ontology, see
e.g., R. Poli, “Framing Ontology”
17. <http://noetic.org/research/overview> Institute of Noetic
Sciences
18. <https://consciousness.med.umich.edu/> Center for
Consciousness Science
19. <https://www.tandfonline.com/toc/ines20/current>
International Journal of Neuroscience
20. <https://www.sciencedirect.com/journal/international-journal-of-psychophysiology>
International Journal of Psychophysiology
21. <https://www.pdcnet.org/process> Journal of the Center
for Process Studies
22. <https://penroseinstitute.com/>
Penrose Institute
23. <https://www.closetotruth.com/> Closetotruth
24. <http://oxfordquantum.org/>
Oxford Quantum
25. <https://www.interaliomag.org/> Interalia magazine,
consciousness
26. <http://www.metanexus.net/about-metanexus-institute>
Metanexus Institute
27. <https://www.mindandlife.org/>
Mind and Life Institute
28. <https://www.cifar.ca/ai/>
Canadian Institute For Advanced Research
29. <https://intelligence.org/>
Machine Intelligence Research Institute
30. <http://opensciences.org/journals/consciousness-studies>
Open Sciences, Consciousness Studies
31. <https://lach.arizona.edu/>
Laboratory for the Development of Consciousness

ACKNOWLEDGMENT

Stephen Ripley thanks Lee Johnson, Ph.D., classicist, poet, astronomer and philosopher, and Dushan Bresky, Ph.D., classicist, sculptor, futurist and philosopher, for many invigorating, good-humored discussions on the topics of this article.

REFERENCES

- [1] Philosophy of mind. Available: https://en.wikipedia.org/wiki/Philosophy_of_mind
- [2] Stanford encyclopedia of philosophy. Consciousness. Available: <https://plato.stanford.edu/entries/consciousness/>
- [3] Consciousness. Available: <https://en.wikipedia.org/wiki/Consciousness>
- [4] The Editors of Encyclopedia Britannica. Consciousness. Available: <https://www.britannica.com/topic/consciousness>
- [5] G. Tononi and C. Koch, "Consciousness: here, there and everywhere?" *Philosoph. Trans Roy. Soc. London. B Biological Sci.*, vol. 370, no. 1668, May 2015. DOI: 10.1098/rstb.2014.0167.
- [6] J. R. Searle. (2013). Can information theory explain consciousness? Available: <http://www.nybooks.com/articles/2013/01/10/can-information-theory-explain-consciousness/>
- [7] Field theories of consciousness/Field theories of global consciousness. Available: http://www.scholarpedia.org/article/Field_theories_of_consciousness/Field_theories_of_global_consciousness
- [8] Monism, Available: <https://en.wikipedia.org/wiki/Monism>
- [9] Stanford encyclopedia of philosophy. Pantheism: 4. Ontological Nature of God/world Relation. Available: <https://plato.stanford.edu/entries/pantheism/#OntBas>
- [10] Panpsychism. Available: <https://en.wikipedia.org/wiki/Panpsychism>
- [11] Stanford encyclopedia of philosophy. Pantheism. Available: <https://plato.stanford.edu/entries/pantheism/>
- [12] Stanford encyclopedia of philosophy. Concepts of god. Available: <https://plato.stanford.edu/entries/concepts-god/>
- [13] A. Grünbaum, "The poverty of theistic cosmology," *Brit. J. Philosophy Sci.*, vol. 55, no. 4, pp. 561–614, Dec. 2004.
- [14] P. McCorduck, *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence*. Abingdon, UK: Taylor & Francis, 2004.
- [15] Medical renaissance. Available: https://en.wikipedia.org/wiki/Medical_Renaissance
- [16] History of neuroscience. Available: https://en.wikipedia.org/wiki/History_of_neuroscience
- [17] M. Fleming-Williams. How tax reform will net the U.S. big returns. Available: <https://worldview.stratfor.com/article/us-tax-reform-corporations-offshore-china-tech-ai-dollar>
- [18] G. Allen and E. B. Kania. China is using America's own plan to dominate the future of artificial intelligence. Available: <http://foreignpolicy.com/2017/09/08/china-is-using-americas-own-plan-to-dominate-the-future-of-artificial-intelligence/>
- [19] Findings of the investigation into China's acts, policies, and practices related to technology transfer, intellectual property, and innovation under section 301 of the trade act of 1974. Available: https://ustr.gov/sites/default/files/Section_301_FINAL.PDF
- [20] Emmanuel Macron talks to WIRED about France's AI strategy. Available: <https://www.wired.com/story/emmanuel-macron-talks-to-wired-about-frances-ai-strategy/>
- [21] US congress artificial intelligence hearing - hearing I. Available: https://www.youtube.com/watch?v=DZ4058m0a_g&feature=em-uploademail
- [22] US congress artificial intelligence hearing - hearing II. Available: <https://www.youtube.com/watch?v=TODqrJSZ4BI&feature=em-uploademail>
- [23] US congress artificial intelligence hearing - hearing III. Available: <https://www.youtube.com/watch?v=oTny--vbKx0&feature=em-uploademail>
- [24] China now has the most valuable AI startup in the world. Available: <https://www.bloomberg.com/news/articles/2018-04-09/sensetime-snags-alibaba-funding-at-a-record-3-billion-valuation>
- [25] China's Plan for world domination in AI isn't so crazy after all. Available: <https://www.bloomberg.com/news/articles/2017-08-14/china-s-plan-for-world-domination-in-ai-isn-t-so-crazy-after-all>
- [26] Available: https://www.bloomberg.com/news/articles/2018-04-08/forget-the-trade-war-china-wants-to-win-the-computing-armsrace?cmpid=BBD040918_MKT&utm_medium=email&utm_source=newsletter&utm_term=180409&utm_campaign=markets
- [27] K. Grace, J. Salvatier, A. Dafoe, B. Zhang, and O. Evans. (2017). When Will AI Exceed Human Performance? Evidence from AI Experts. Available: <https://arxiv.org/abs/1705.08807>
- [28] A. Vutha. Could machine learning mean the end of understanding in science? Available: <https://theconversation.com/could-machine-learning-mean-the-end-of-understanding-in-science-98995>
- [29] Peano's Axioms. Available: <http://mathworld.wolfram.com/PeanosAxioms.html>
- [30] Zermelo-Fraenkel axioms. Available: <http://mathworld.wolfram.com/Zermelo-FraenkelAxioms.html>
- [31] Mereology. Available: <https://en.wikipedia.org/wiki/Mereology>
- [32] History of logic. Available: https://en.wikipedia.org/wiki/History_of_logic#Modern_logic
- [33] Stanford encyclopedia of philosophy. Kurt Gödel. Available: <https://plato.stanford.edu/entries/goedel/>
- [34] Stanford encyclopedia of philosophy. Alfred Tarski. Available: <https://plato.stanford.edu/entries/tarski/>
- [35] Metamathematics. Available: <https://en.wikipedia.org/wiki/Metamathematics>
- [36] P vs NP problem. Available: <http://www.claymath.org/millennium-problems/p-vs-np-problem>
- [37] S. Cook. The P versus NP problem. Available: <http://www.claymath.org/sites/default/files/pvsnpp.pdf>
- [38] List of unsolved problems in computer science. Available: https://en.wikipedia.org/wiki/List_of_unsolved_problems_in_computer_science
- [39] Decision problem. Available: <http://mathworld.wolfram.com/DecisionProblem.html>
- [40] N. S. Yanofsky. The outer limits of reason what science, mathematics, and logic cannot tell us. Available: <https://mitpress.mit.edu/books/outer-limits-reason>
- [41] What we cannot know - Marcus du Sautoy. Available: <https://www.youtube.com/watch?v=xbo3NZdReEg>
- [42] The limits of understanding. Available: <https://www.youtube.com/watch?v=DFY-DRsE86s>
- [43] S. Wolfram, *A New Kind of Science*. Champaign, IL: Wolfram Media.
- [44] Stanford encyclopedia of philosophy. Turing machines. Available: <https://plato.stanford.edu/entries/turing-machine/#Uncomputability>
- [45] Computing machinery and intelligence. Available: https://en.wikipedia.org/wiki/Computing_Machinery_and_Intelligence

- [46] G. Marcus, F. Rossi, and M. Veloso. Beyond the Turing test. Available: <https://aaai.org/ojs/index.php/aimagazine/article/view/2650/2527>
- [47] Turing test. Available: https://en.wikipedia.org/wiki/Turing_test
- [48] The Turing test. Available: <http://www.psych.utoronto.ca/users/reingold/courses/ai/turing.html>
- [49] Church–Turing thesis. Available: https://en.wikipedia.org/wiki/Church%E2%80%93Turing_thesis
- [50] S. Lloyd, "Ultimate physical limits to computation," *Nature*, vol. 406, pp. 1047–1054, Aug. 2000.
- [51] S. Aaronson. (2011). Why philosophers should care about computational complexity. Available: <https://arxiv.org/pdf/1108.1791.pdf>
- [52] Complex system. Available: https://en.wikipedia.org/wiki/Complex_system
- [53] Synthetic intelligence. Available: https://en.wikipedia.org/wiki/Synthetic_intelligence
- [54] Artificial general intelligence. Available: https://en.wikipedia.org/wiki/Artificial_general_intelligence
- [55] Big data. Available: https://en.wikipedia.org/wiki/Big_data
- [56] Big history. Available: https://en.wikipedia.org/wiki/Big_History
- [57] Machine learning. Available: https://en.wikipedia.org/wiki/Machine_learning
- [58] AI - externalization of mind. Available: https://www.slideshare.net/Mills/ai-externalization-of-mind-81319464?qid=13698514-50c9-4fa1-abd5-480726f19a9b&v=&b=&from_search=1
- [59] The limits of quantum. Available: <https://www.ime.usp.br/~pf/clippings/quantum/quantum-computing-200803.pdf>
- [60] M. Shanahan. (2015). Ascribing consciousness to artificial intelligence. Available: <https://arxiv.org/abs/1504.05696>
- [61] M. Wilson. AI is inventing languages humans can't understand. should we stop it? Available: <https://www.fastcodesign.com/90132632/ai-is-inventing-its-own-perfect-languages-should-we-let-it>
- [62] Stanford encyclopedia of philosophy. Computability and complexity. Available: <https://plato.stanford.edu/entries/computability/>
- [63] The rise of the conscious machines: how far should we take AI? Available: <http://www.sciencefocus.com/article/future/artificial-intelligence-conscience-robots-ai>
- [64] D. Weinberger. Our machines now have knowledge we'll never understand. Available: <https://www.wired.com/story/our-machines-now-have-knowledge-well-never-understand/?mbid=BottomRelatedStories>
- [65] K. Zhao, Y. Song, and Z. Shen, "Neuroadaptive fault-tolerant control of nonlinear systems under output constraints and actuation faults," *IEEE Trans. Neural Networks Learning Syst.*, vol. 29, no. 2, pp. 286–298, Feb. 2018.
- [66] I. Niles and A. Pease, "Towards a standard upper ontology," in *Proc. Int. Conf. Formal Ontology Inform. Syst.*, Ogunquit, Maine, USA, 2001, pp. 2–9.
- [67] Suggested upper merged ontology. Available: https://en.wikipedia.org/wiki/Suggested_Upper_Merged_Ontology
- [68] Commonsense knowledge (artificial intelligence). Available: [https://en.wikipedia.org/wiki/Commonsense_knowledge_\(artificial_intelligence\)#Commonsense_knowledge_bases](https://en.wikipedia.org/wiki/Commonsense_knowledge_(artificial_intelligence)#Commonsense_knowledge_bases)
- [69] Semantic network. Available: https://en.wikipedia.org/wiki/Semantic_network
- [70] Global scientific output doubles every nine years. Available: <http://blogs.nature.com/news/2014/05/global-scientific-output-doubles-every-nine-years.html>
- [71] L. Ljung, "Black-box models from input-output measurements," in *Proc. 18th IEEE Instrumentation Measurement Technol. Conf. Rediscovering Measurement Age Informatics (Cat. No.01CH 37188)*, Budapest, Hungary, 2001, vol. 1, pp. 138–146.
- [72] Black box. Available: https://en.wikipedia.org/wiki/Black_box
- [73] The building blocks of interpretability. Available: <https://distill.pub/2018/building-blocks/>
- [74] The AI detectives. Available: <http://science.sciencemag.org/content/sci/357/6346/22.full.pdf>
- [75] M. Lehnis. Can we trust AI if we don't know how it works? Available: <https://www.bbc.com/news/business-44466213>
- [76] The art of building artificial humans. Available: <https://cosmosmagazine.com/technology/the-art-of-building-artificial-humans>
- [77] Outline of artificial intelligence. Available: https://en.wikipedia.org/wiki/Outline_of_artificial_intelligence
- [78] The age of intelligent machines. Available: https://en.wikipedia.org/wiki/The_Age_of_Intelligent_Machines
- [79] 2002 : What is your question? ... Why? Available: <https://www.edge.org/response-detail/11150>
- [80] Arthur C. Clarke. Available: https://en.wikipedia.org/wiki/Arthur_C._Clarke#Futurism
- [81] Clarke's three laws. Available: https://en.wikipedia.org/wiki/Clarke%27s_three_laws#External_links
- [82] Paul Allen: The singularity isn't near. Available: <https://www.technologyreview.com/s/425733/paul-allen-the-singularity-isnt-near/>
- [83] A. B. Barrett, "A comment on Tononi & Koch (2015) 'Consciousness: here, there and everywhere?'," *Philosoph. Trans. Roy. Soc. London. B Biological Sci.*, vol. 371, no. 1687, Feb. 2016.
- [84] D. Deutsch, "Quantum theory, the Church–Turing principle and the universal quantum computer," *Proc. R. Soc. Lond. A*, vol. 400, no. 1818, pp. 97–117, Jul. 1985.
- [85] G. Tononi and C. Koch, "A reply to Barrett (2016)," *Philosoph. Trans. Roy. Soc. London. B Biological Sci.*, 10.1098/rstb.2015.0452 vol. 371, no. 1687, p. 20150452, Jan. 2016.
- [86] A hybrid of quantum computing and machine learning is spawning new ventures. Available: <https://spectrum.ieee.org/tech-talk/robotics/artificial-intelligence/kickstarting-the-quantum-startup-a-hybrid-of-quantum-computing-and-machine-learning-is-spawning-new-ventures>
- [87] How quantum effects could improve artificial intelligence. Available: <https://phys.org/news/2016-10-quantum-effects-artificial-intelligence.html>
- [88] R. P. Feynman, "Simulating physics with computers " *Int. J. Theoretical Physics*, vol. 21, no. 6/7, pp. 1–22, Jul. 1982.
- [89] U. Ehsan, B. Harrison, L. Chan, and M. O. Riedl. (2017). Rationalization: a neural machine translation approach to generating natural language explanations. Available: <https://arxiv.org/abs/1702.07826>
- [90] Human-Centered Artificial Intelligence. Available: https://medium.com/@mark_riedl/human-centered-artificial-intelligence-70b019f956d1
- [91] Artificial consciousness. Available: https://en.wikipedia.org/wiki/Artificial_consciousness

- [92] J. Jonkisz, M. Wierchoń, and M. Binder, "Four-dimensional graded consciousness," *Frontiers Psychology*, vol. 8, Mar. 2017. DOI: 10.3389/fpsyg.2017.00420.
- [93] Stanford Encyclopedia of Philosophy. The structure of scientific theories. Available: <https://plato.stanford.edu/entries/structure-scientific-theories/#Int>
- [94] D. Abbott, "The reasonable ineffectiveness of mathematics [point of view]," *Proc. IEEE*, vol. 101, no. 10, pp. 2147–2153, Sep. 2013.
- [95] Stanford Encyclopedia of Philosophy. Identity. Available: <https://plato.stanford.edu/entries/identity/#2>
- [96] Stanford Encyclopedia of Philosophy. Identity and individuality in quantum theory. Available: <https://plato.stanford.edu/entries/qt-idind/>
- [97] NeuroQuantology. Available: <https://www.neuroquantology.com/index.php/journal/index>
- [98] Dual aspect monism: the alternative to neural reductionism. Available: <https://www.scienceandnonduality.com/dual-aspect-monism/>
- [99] Internet encyclopedia of philosophy. Panpsychism. Available: <https://www.iep.utm.edu/panpsych/>
- [100] Stanford encyclopedia of philosophy. Panpsychism. Available: <https://plato.stanford.edu/entries/panpsychism/>
- [101] Is consciousness universal? Available: <https://www.scientificamerican.com/article/is-consciousness-universal/>
- [102] Ubiquitous minds. Available: https://alleninstitute.org/media/filer_public/4a/3c/4a3c950c-ae48-4a44-bd47-539c8f9d6bfa/2014_01_ubiquitousminds.pdf
- [103] G. Tononi, "The integrated information theory of consciousness: an updated account," *Archives Italiennes de Biologie*, vol. 150, no. 2/3, pp. 56–90, May 2011.
- [104] M. Oizumi, L. Albantakis, and G. Tononi, "From the phenomenology to the mechanisms of consciousness: Integrated Information Theory 3.0," *PLoS Computational Biology*, vol. 10, no. 5, p. e1003588, May 2014.
- [105] Can we quantify machine consciousness? Available: <https://spectrum.ieee.org/computing/hardware/can-we-quantify-machine-consciousness>
- [106] G. Tononi, M. Boly, M. Massimini, and C. Koch, "Integrated information theory: from consciousness to its physical substrate," *Nature Rev. Neuroscience*, Perspective, vol. 17, p. 450, May 2016.
- [107] C. Koch and G. Tononi. (2008). Can machines be conscious? Available: <https://spectrum.ieee.org/medical/imaging/can-machines-be-conscious>
- [108] J.-A. Johannessen, "Problems and difficulties related to information science," *Kybernetes*, vol. 26, no. 5, pp. 537–554, May 1997.
- [109] M. A. Cerullo, "The problem with phi: a critique of integrated information theory," *PLoS Computational Biology*, vol. 11, no. 9, p. e1004286, Sep. 2015.
- [110] A. B. Barrett, "An integration of integrated information theory with fundamental physics," *Frontiers in Psychology*, vol. 5, p. doi:10.3389/fpsyg.2014.00063, Feb. 2014.
- [111] M. Tegmark. (2015). Consciousness as a state of matter. Available: <https://arxiv.org/abs/1401.1219>
- [112] M. Tegmark, "Improved measures of integrated information," *PLoS Computational Biology*, vol. 12, no. 11, p. e1005123, Nov. 2016.
- [113] S. Russell, D. Dewey, and M. Tegmark, "Research priorities for robust and beneficial artificial intelligence," *AI Magazine*, vol. 36, no. 5, pp. 105–114, Feb. 2015.
- [114] Stanford encyclopedia of philosophy. Neutral monism. Available: <https://plato.stanford.edu/entries/neutral-monism/#InfoUltiReal>
- [115] M. Tegmark, "The mathematical universe," *Found. Physics*, vol. 38, no. 2, pp. 101–150, Feb. 2008.
- [116] Mathematical universe hypothesis. Available: https://en.wikipedia.org/wiki/Mathematical_universe_hypothesis
- [117] Shtetl-Optimized. This review of Max Tegmark's book also occurs infinitely often in the decimal expansion of π . Available: <https://www.scottaaronson.com/blog/?p=1753>
- [118] M. Tegmark. (2015). Consciousness as a state of matter. Available: <https://arxiv.org/abs/1405.0493>
- [119] Before the Big Bang 7: An eternal cyclic universe, CCC revisited & Twistor theory. Available: <https://www.youtube.com/watch?v=FVDJJVoTx7s>
- [120] Penrose interpretation. Available: https://en.wikipedia.org/wiki/Penrose_interpretation
- [121] Conformal cyclic cosmology. Available: https://en.wikipedia.org/wiki/Conformal_cyclic_cosmology
- [122] The Penrose Institute. Available: <https://penroseinstitute.com/#philosophy>
- [123] Roger Penrose. Available: https://en.wikipedia.org/wiki/Roger_Penrose
- [124] Quantum consciousness. Available: <http://www.quantumconsciousness.org/>
- [125] Available: <https://sci-hub.hk/https://doi.org/10.1016/j.pprev.2013.08.002>
- [126] Process and reality. Available: https://en.wikipedia.org/wiki/Process_and_Reality
- [127] V. G. Gurzadyan and R. Penrose, "CCC and the Fermi paradox," *Eur. Physical J. Plus*, vol. 131, no. 1, p. 11, Jan. 2016.
- [128] 10 unsettling solutions to the fermi paradox. Available: <https://www.youtube.com/watch?v=aBf7uAxx6ds&t=21s>
- [129] S. Hameroff and R. Penrose, "Consciousness in the universe: a review of the 'Orch OR' theory," *Physics Life Rev.*, vol. 11, no. 1, pp. 39–78, Mar. 2014.
- [130] D. McDermott, "Computation and consciousness," *Behavioral Brain Sci.*, vol. 13, no. 4, pp. 676–678, 2011.
- [131] A. K. Seth, "Putting Descartes before the horse: Quantum theories of consciousness: Comment on 'Consciousness, biology, and quantum hypotheses' by Baars & Edelman," *Phys. Life Rev.*, vol. 9, no. 3, pp. 297–298, Sep. 2012.
- [132] S. Aaronson. (2013). The ghost in the quantum turing machine. Available: <https://arxiv.org/abs/1306.0159>
- [133] Shtetl-optimized. Ask me anything: moral judgments edition. Available: <https://www.scottaaronson.com/blog/>
- [134] Shtetl-optimized. Can computers become conscious?: my reply to Roger Penrose. Available: <https://www.scottaaronson.com/blog/?p=2756>
- [135] S. Aaronson. PHYS771 lecture 10.5: Penrose. Available: <https://www.scottaaronson.com/democritus/lec10.5.html>
- [136] Theory of computation. Available: https://en.wikipedia.org/wiki/Theory_of_computation
- [137] P. Hut, M. Alford, and M. Tegmark. On math, matter and mind. Available: <https://arxiv.org/pdf/physics/0510188.pdf>
- [138] M. Havlik, E. Kozáková, J. Horáček. "Why and how. The future of the central questions of consciousness," *Front. Psychol.*, vol. 8, pp. 1797, 2017.
- [139] D. Skrbina. Panpsychism in the West. Available: <https://mitpress.mit.edu/books/panpsychism-west>
- [140] D. Skrbina, "Mind that abides: panpsychism in the new millennium," in *Advances in Consciousness Research*

- Series*, Amsterdam and Philadelphia: John Benjamins, 2009, pp. 341–260.
- [141] G. Strawson. Realistic monism: why physicalism entails panpsychism. Available: http://www.newdualism.org/papers/G.Strawson/strawson_on_panpsychism.pdf
- [142] S. Kauffman, "Life after death? An improbable essay," *Scientific GOD J.*, vol. 8, no. 5, pp. 400–404, 2017.
- [143] Available: <https://scihub.hk/https://doi.org/10.1080/14746700.2015.1122325>
- [144] R. L. P. Vimal, "Meanings attributed to the term 'consciousness': an overview," *J. Consciousness Stud.*, vol. 16, no. 5, pp. 9–27, Jan. 2009.
- [145] Neuroquantology. Available: <https://www.neuroquantology.com/index.php/journal/search/authors>
- [146] Available: https://scholar.google.ca/scholar?hl=en&as_sdt=0%2C5&q=Vimal%2CR.L.P.&btnG
- [147] H. Atmanspacher. (2009). Quantum approaches to consciousness : A systematic overview with selected examples (mathematical physics and application of nonlinear wave phenomena). Available: <https://repository.kulib.kyoto-u.ac.jp/dspace/bitstream/2433/140669/1/1645-16.pdf>
- [148] *Issue 1, Soc. Mind-Matter Res.*, vol. 14, 2016. Available: http://www.mindmatter.de/journal/issues/missue14_1.html
- [149] The science of consciousness conference - April 2-7, 2018. Available: <http://medicine.arizona.edu/event/2018/science-consciousness-conference-april-2-7-2018>
- [150] A. Seth (2012) Consciousness: Eight questions science must answer. Available: <https://www.theguardian.com/science/2012/mar/01/consciousness-eight-questions-science>
- [151] Available: <https://www.youtube.com/watch?v=8n5GD69wFOA>
- [152] M. Gardner. Do loops explain consciousness? Review of i am a strange loop. Available: <http://www.ams.org/notices/200707/tx070700852p.pdf>
- [153] P. Ball. The strange link between the human mind and quantum physics. Available: <http://www.bbc.com/earth/story/20170215-the-strange-link-between-the-human-mind-and-quantum-physics>
- [154] Exploring quantum theories of consciousness. Available: <http://quantum-mind.co.uk/>
- [155] D. Zohar, *The Quantum Self*. New York, NY: HarperCollins, 1991.
- [156] M. G. Dyer, "Quantum physics and consciousness, creativity, computers: a commentary on goswami's quantum-based theory of consciousness and free will," *J. Mind Behavior*, vol. 15, no. 3, pp. 265–290, Summer 1994.
- [157] Stanford Encyclopedia of Philosophy. Quantum approaches to consciousness. Available: <https://plato.stanford.edu/entries/qt-consciousness/#4.7>
- [158] Quantum mysticism. Available: https://en.wikipedia.org/wiki/Quantum_mysticism
- [159] Could telepathy one day be explained by modern physics? Available: <http://www.tcm.phy.cam.ac.uk/~bdj10/stamps/today.html>
- [160] S. Carroll. Telekinesis and quantum field theory. Available: <http://www.preposterousuniverse.com/blog/2008/02/18/telekinesis-and-quantum-field-theory/>
- [161] Special report: the singularity. Available: <https://spectrum.ieee.org/static/singularity>
- [162] IEEE computational intelligence society. Available: <https://cis.ieee.org/>
- [163] IEEE computational intelligence magazine special issue on automated design of machine learning and search algorithms. Available: <https://cis.ieee.org/component/content/article/19-e-newsletter-publications/724-cfp-ieee-cim-special-issue-on-automated-design-of-machine-learning-and-search-algorithms.html>
- [164] IEEE computational intelligence magazine. Call for papers. Special issue on automated design of machine learning and search algorithms. Available: <http://titancs.ukzn.ac.za/CIMADA2017.aspx>
- [165] J. Ouellette. Brains may teeter near their tipping point. Available: <https://www.quantamagazine.org/brains-may-teeter-near-their-tipping-point-20180614/>
- [166] A. Kent. (2016). Quanta and qualia. Available: <https://arxiv.org/abs/1608.04804>
- [167] Available: <https://www.quantamagazine.org/a-new-spin-on-the-quantum-brain-20161102/>
- [168] Stanford Encyclopedia of Philosophy. Philosophical Issues in Quantum Theory. Available: <https://plato.stanford.edu/entries/qt-issues/>
- [169] D. J. Castellano. (2011) Ontological interpretation of quantum mechanics. Available: <http://www.arcaneknowledge.org/science/quantum.htm>
- [170] R. B. Laughlin and D. Pines, "The theory of everything," *Proc. Nat. Academy Sci. U. S. A.*, vol. 97, no. 1, pp. 28–31, Jan. 2000.
- [171] S. Paulson, M. Gleiser, K. Freese, and M. Tegmark, "The unification of physics: the quest for a theory of everything," *Ann. New York Academy Sci.*, vol. 1361, pp. 18–35, Dec 2015.
- [172] Theory of everything. Available: https://en.wikipedia.org/wiki/Theory_of_everything
- [173] R. L. P. Vimal, "Towards a theory of everything part I: introduction of consciousness in electromagnetic theory, special and general theory of relativity," *NeuroQuantology*, vol. 8, no. 2, pp. 206–230, Jun. 2010.
- [174] R. L. P. Vimal, "Towards a theory of everything part II: introduction of consciousness in schrödinger equation and standard model," *NeuroQuantology*, vol. 8, no. 3, pp. 304–313, Sep. 2010.
- [175] R. L. P. Vimal, "Towards a theory of everything part III: introduction of consciousness in loop quantum gravity and string theory and unification of experiences with fundamental forces," *NeuroQuantology*, vol. 8, no. 4, pp. 571–599, Sep. 2010.
- [176] S. Lloyd. Ultimate physical limits to computation. Available: <https://arxiv.org/pdf/quant-ph/9908043.pdf>
- [177] J. Kim, "Emergence: Core ideas and issues," *Synthese*, vol. 151, no. 3, pp. 547–559, Aug. 2006.
- [178] T. Lei, R. Barzilay, and T. Jaakkola. Rationalizing neural predictions. Available: https://people.csail.mit.edu/taolei/papers/emnlp16_rationalization.pdf
- [179] Deep learning patterns. Available: <http://www.deeplearningpatterns.com/doku.php?id=rationalization>
- [180] Deep learning patterns, methodology and strategy™. Available: <http://www.intuitionmachine.com/>
- [181] H. Lieberman. (2008). Usable AI requires commonsense knowledge. Available: <http://web.media.mit.edu/~lieber/Publications/Usable-AI-Commonsense.pdf>
- [182] Natural-language programming. Available: https://en.wikipedia.org/wiki/Natural-language_programming
- [183] Natural language processing. Available: https://en.wikipedia.org/wiki/Natural_language_processing
- [184] Natural language. Available: https://en.wikipedia.org/wiki/Natural_language

- [185] Stanford encyclopedia of philosophy. Scientific revolutions. Available: <https://plato.stanford.edu/entries/scientific-revolutions/>
- [186] S. Majid, "On the emergence of the structure of physics," *Philosoph. Trans. Roy. Soc. A: Math. Physical Eng. Sci.*, 10.1098/rsta.2017.0231 vol. 376, no. 2118, Mar. 2018.
- [187] M. Mizrahi, "The Argument from underconsideration and relative realism," *Int. Stud. Philosophy Sci.*, vol. 27, no. 4, pp. 393-407, Apr. 2013.
- [188] Stanford encyclopedia of philosophy. Scientific realism. Available: <https://plato.stanford.edu/entries/scientific-realism/>
- [189] Formal language. Available: https://en.wikipedia.org/wiki/Formal_language
- [190] Y. Saplakoglu and S. Writer. (2018). Forgotten element could redefine time. Available: https://www.livescience.com/62631-atomic-clock-lutetium-precise-time.html?utm_source=ls-newsletter&utm_medium=email&utm_campaign=20180523-ls
- [191] Systems theory. Available: https://en.wikipedia.org/wiki/Systems_theory
- [192] Scientific modelling. Available: https://en.wikipedia.org/wiki/Scientific_modelling
- [193] Epistemology. Available: <https://en.wikipedia.org/wiki/Epistemology>
- [194] Human knowledge: foundations and limits. Available: <http://humanknowledge.net/Thoughts.html#MindandArtifact>
- [195] Computability theory. Available: https://en.wikipedia.org/wiki/Computability_theory
- [196] M. Tops, M. A. Boksem, M. Quirin, I. J. H, and S. L. Koole, "Internally directed cognition and mindfulness: an integrative perspective derived from predictive and reactive control systems theory," (in eng), *Frontiers in Psychology*, vol. 5, p. 429, May 2014.
- [197] J. O. Kephart and D. M. Chess, "The vision of autonomic computing," *Computer*, vol. 36, no. 1, pp. 41-50, Jan. 2003.
- [198] Autonomic computing. Available: https://en.wikipedia.org/wiki/Autonomic_computing
- [199] Type introspection. Available: https://en.wikipedia.org/wiki/Type_introspection
- [200] Hierarchical control system. Available: https://en.wikipedia.org/wiki/Hierarchical_control_system#Artificial_intelligence
- [201] Reflection (computer programming). Available: [https://en.wikipedia.org/wiki/Reflection_\(computer_programming\)](https://en.wikipedia.org/wiki/Reflection_(computer_programming))
- [202] Metaprogramming. Available: <https://en.wikipedia.org/wiki/Metaprogramming>
- [203] Available: <https://code.nasa.gov/#/>
- [204] Computers in spaceflight: the NASA experience. Available: <https://history.nasa.gov/computers/Ch2-6.html>
- [205] All topics A-Z. Available: <https://www.nasa.gov/tags/>
- [206] NASA.GOV 2.0. Available: <https://www.mobomo.com/portfolio/nasa-website/>
- [207] Technology drives exploration. Available: <https://www.nasa.gov/topics/technology/robotics/index.html>
- [208] These robots could help NASA find alien life. Available: <https://www.nasa.gov/feature/these-robots-could-help-nasa-find-alien-life>
- [209] B. Granath. (2018). Swarmathon improves student skills in robotics, computer science. Available: <https://www.nasa.gov/feature/swarmathon-improves-student-skills-in-robotics-computer-science>
- [210] Available: <https://sci-hub.hk/https://doi.org/10.1016/j.concog.2004.08.008>
- [211] M. G. Dyer. Overview of current research interests. Available: <http://web.cs.ucla.edu/~dyer/CurResInterests.html>
- [212] Judea Pearl - cognitive systems laboratory. Available: <http://bayes.cs.ucla.edu/>
- [213] How a pioneer of machine learning became one of its sharpest critics. Available: <https://www.theatlantic.com/technology/archive/2018/05/machine-learning-is-stuck-on-asking-why/560675/>
- [214] B. M. Lake, T. D. Ullman, J. B. Tenenbaum, and S. J. Gershman, "Building machines that learn and think like people," *Behavioral Brain Sci.*, vol. 40, p. e253, Jan 2017.
- [215] A neural network model that can reason - Prof. Christopher Manning. Available: <https://www.youtube.com/watch?v=24AX4qJ7Ts&feature=em-uploademail>
- [216] Google's machine learning software has learned to replicate itself. Available: <https://futurism.com/googles-machine-learning-software-has-learned-to-replicate-itself/>
- [217] T. Simonite. Google's learning software learns to write learning software. Available: <https://www.wired.com/story/googles-learning-software-learns-to-write-learning-software/>
- [218] G. Marcus. Deep learning: a critical appraisal. Available: <https://arxiv.org/abs/1801.00631>
- [219] A. L'Heureux, K. Grolinger, H. F. Elyamany, and M. A. M. Capretz, "Machine learning with big data: challenges and approaches," *IEEE Access*, vol. 5, pp. 7776-7797, 2017.
- [220] List of academic databases and search engines. Available: https://en.wikipedia.org/wiki/List_of_academic_databases_and_search_engines
- [221] Google scholar. Available: https://en.wikipedia.org/wiki/Google_Scholar
- [222] Scholarpedia. Available: <https://en.wikipedia.org/wiki/Scholarpedia>
- [223] Wolfram products & services. Available: <http://www.wolfram.com/products/>
- [224] Wolframalpha. Available: <http://www.wolframalpha.com/faqs10.html>
- [225] Ontology (information science). Available: [https://en.wikipedia.org/wiki/Ontology_\(information_science\)](https://en.wikipedia.org/wiki/Ontology_(information_science))
- [226] Olog. Available: <https://en.wikipedia.org/wiki/Olog>
- [227] re3data.ORG. Available: <https://www.re3data.org/>
- [228] The meta library. Available: <https://sites.google.com/site/themetalibrary/home>
- [229] Sci-Hub. Available: <https://en.wikipedia.org/wiki/Sci-Hub>
- [230] Open access. Available: https://en.wikipedia.org/wiki/Open_access
- [231] The commons. Available: <https://www.flickr.com/commons/institutions/>
- [232] Portal:contents. Available: <https://en.wikipedia.org/wiki/Portal:Contents>
- [233] Portal:contents/portals. Available: <https://en.wikipedia.org/wiki/Wikipedia:Portal/Directory>
- [234] BORO. Available: <https://en.wikipedia.org/wiki/BORO>
- [235] Wikipedia:disambiguation. Available: <https://en.wikipedia.org/wiki/Wikipedia:Disambiguation>
- [236] IEEE 1855. Available: https://en.wikipedia.org/wiki/IEEE_1855
- [237] Fuzzy markup language. Available: https://en.wikipedia.org/wiki/Fuzzy_markup_language
- [238] IEEE standards association. Available: <https://standards.ieee.org/>
- [239] Classical physics. Available: https://en.wikipedia.org/wiki/Classical_physics

- [240] Physicalism. Available: <https://en.wikipedia.org/wiki/Physicalism>
- [241] S. Garrabrant, T. Benson-Tilsen, A. Critch, N. Soares, and J. Taylor. (2017). A formal approach to the problem of logical non-omniscience. Available: <https://arxiv.org/abs/1707.08747>
- [242] The practical business of ontology: a tale from the front lines. Available: <http://blog.stephenwolfram.com/2017/07/the-practical-business-of-ontology-a-tale-from-the-front-lines/>
- [243] Knowledge-based programming. Available: <http://www.wolfram.com/language/>
- [244] Artificial intelligence productivity for virtually every developer and scenario. Available: <https://azure.microsoft.com/en-us/overview/ai-platform/?v=17.42w>
- [245] Azure products. Available: <https://azure.microsoft.com/en-us/services/#ai-machine-learning>
- [246] Seven deadly sins. Available: https://en.wikipedia.org/wiki/Seven_deadly_sins
- [247] Ethics in religion. Available: https://en.wikipedia.org/wiki/Ethics_in_religion
- [248] Semantics. Available: <https://en.wikipedia.org/wiki/Semantics>
- [249] Computability logic. Available: <http://www.csc.villanova.edu/~japaridz/CL/1.html#Subsection1a1>
- [250] W. Knight. AI software juggles probabilities to learn from less data. Available: <https://www.technologyreview.com/s/603542/ai-software-juggles-probabilities-to-learn-from-less-data/>
- [251] The real-world potential and limitations of artificial intelligence. Available: <https://www.mckinsey.com/featured-insights/artificial-intelligence/the-real-world-potential-and-limitations-of-artificial-intelligence>
- [252] List of cognitive biases. Available: https://en.wikipedia.org/wiki/List_of_cognitive_biases
- [253] R. Wright. Why pure reason won't end american tribalism. Available: <https://www.wired.com/story/why-pure-reason-wont-end-american-tribalism/>
- [254] A. Caliskan, J. J. Bryson, and A. Narayanan, "Semantics derived automatically from language corpora contain human-like biases," *Science*, vol. 356, no. 6334, pp. 183–186, Aug. 2017.
- [255] Welcome to the official site of Richard Feynman. Available: <http://www.richardfeynman.com/>
- [256] J. Pontin. How do you know you are reading this? Available: <https://www.wired.com/story/tricky-business-of-measuring-consciousness/>
- [257] Gary Larson. Available: https://en.wikipedia.org/wiki/Gary_Larson
- [258] Pogo (comic strip). Available: [https://en.wikipedia.org/wiki/Pogo_\(comic_strip\)](https://en.wikipedia.org/wiki/Pogo_(comic_strip))
- [259] Peanuts. Available: <https://en.wikipedia.org/wiki/Peanuts>
- [260] Default mode network. Available: https://en.wikipedia.org/wiki/Default_mode_network
- [261] Imagination is more important than knowledge. Available: <https://quoteinvestigator.com/2013/01/01/einstein-imagination/>
- [262] Albert Einstein quotes. Available: <http://www.alberteinstein.com/quotes/einsteinquotes.html#science>
- [263] Agents that imagine and plan. Available: <https://deepmind.com/blog/agents-imagine-and-plan/>
- [264] L. Muehlhauser. Bas Steunebrink on self-reflective programming. Available: <https://intelligence.org/2013/10/25/bas-steunebrink-on-sleight/>
- [265] M. Gayford. (2016). Robot art raises questions about human creativity. Available: <https://www.technologyreview.com/s/600762/robot-art-raises-questions-about-human-creativity/>
- [266] Curiosity may be vital for truly smart AI. Available: <https://www.technologyreview.com/s/607886/curiosity-may-be-vital-for-truly-smart-ai/>
- [267] Clever machines learn how to be curious. Available: <https://www.quantamagazine.org/clever-machines-learn-how-to-be-curious-20170919/>
- [268] P. Rudin. Building AI systems that doubt themselves. Available: <https://singularity2030.ch/building-ai-systems-that-doubt-themselves/>
- [269] How to build a self-conscious machine. Available: <https://www.wired.com/story/how-to-build-a-self-conscious-ai-machine/>
- [270] Understanding the four types of AI, from reactive robots to self-aware beings. Available: <http://theconversation.com/understanding-the-four-types-of-ai-from-reactive-robots-to-self-aware-beings-67616>
- [271] S. Schneider and E. Turner. (2017). Is anyone home? A way to find out if ai has become self-aware. Available: <https://blogs.scientificamerican.com/observations/is-anyone-home-a-way-to-find-out-if-ai-has-become-self-aware/>
- [272] Can A Robot Feel? by Susan Schneider, YHouse talk at the Rubin Museum on 12 12 16. Available: <https://www.youtube.com/watch?v=PVyQAsYOrU0>
- [273] Reverse Engineering the Mind - Prof. James DiCarlo, MIT Department of Brain and Cognitive Sciences. Available: <https://www.youtube.com/watch?v=3djQsX1FJ9I>
- [274] Sex doll 'brothel': Xdolls escapes Paris council censure. Available: <http://www.bbc.com/news/world-europe-43506621>
- [275] Embracing the robot. Available: <https://aeon.co/essays/programmed-to-love-is-a-human-robot-relationship-wrong?>
- [276] Is today's artificial intelligence actually conscious? Not just yet. Available: https://www.cifar.ca/wp-content/uploads/KMBrief_WhatIsConsciousness.pdf
- [277] S. P. R. Rose, "Human agency in the neurocentric age," *EMBO Rep.*, vol. 6, no. 11, pp. 1001–1005, Nov. 2005.
- [278] D. Jacquette, "What would a cerebroscope do?," *J. Brit. Soc. Phenomenology*, vol. 27, no. 2, pp. 188–199, May 1996.
- [279] M. Cohen, R. A. Terwilliger, X. Hong, M. Rohan, and P. Roemer, "Real-time observation of mental activity: the autocerebroscope," in *Soc. Neuroscience 27th Annu. Meeting*, Athens, Greece, 1997, p. 417.4.
- [280] Duck test. Available: https://en.wikipedia.org/wiki/Duck_test
- [281] S. Airiau, E. Bonzon, U. Endriss, N. Maudet, and J. Rossit, "Rationalisation of profiles of abstract argumentation frameworks: characterisation and complexity," *J Artificial Intell. Res.*, vol. 60, pp. 149–177, Sep. 2017.
- [282] Genius Rubik's Cube Magician Available: https://www.youtube.com/watch?v=PFA-RmV_wG0&feature=youtu.be
- [283] M. Cabanac, A. J. Cabanac, and A. Parent, "The emergence of consciousness in phylogeny," *Behav. Brain Res.*, vol. 198, no. 2, pp. 267–272, Mar. 2009.
- [284] C. Rovelli, "Shan Gao: the meaning of the wave function. In search of the ontology of quantum mechanics," *Found. Phys.*, vol. 48, no. 6, pp. 747–749, Mar. 2018.
- [285] Double-aspect theory. Available: https://en.wikipedia.org/wiki/Double-aspect_theory
- [286] R. L. Carhart-Harris, R. Leech, P. J. Hellyer, M. Shanahan, A. Feilding, E. Tagliazucchi, D. R. Chialvo,

- and D. Nutt, "The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs," *Frontiers in Human Neuroscience*, vol. 8, Feb. 2014. DOI: 10.3389/fnhum.2014.00020.
- [287] D. Papo, "Commentary: the entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs," *Frontiers in Human Neuroscience*, vol. 10, p. 423, Aug. 2016. DOI: 10.3389/fnhum.2016.00423.
- [288] R. L. Carhart-Harris, "The entropic brain - revisited," *Neuropharmacology*, Mar. 2018. DOI: 10.1016/j.neuropharm.2018.03.010.
- [289] Animal consciousness. Available: https://en.wikipedia.org/wiki/Animal_consciousness
- [290] A. S. Reber, "Caterpillars, consciousness and the origins of mind," *Animal Sentience*, vol. 11, no. 1, May 2016.
- [291] A. S. Reber. (2018). Sentient plants? Nervous minds? Available: <https://animalstudiesrepository.org/cgi/viewcontent.cgi?referer=https://www.google.ca/&httpsredir=1&article=1307&context=animsent>
- [292] If brains are computers, who designs the software? With Daniel Dennett. Available: <https://www.youtube.com/watch?v=TTFoJQsD48c>
- [293] Abiogenesis. Available: <https://en.wikipedia.org/wiki/Abiogenesis>
- [294] J.F. Pagel, and P. Kirshtein, "Chapter twelve – searching for dreams in other (stranger) places," *Machine Dreaming and Consciousness*, pp. 163-173, 2017.
- [295] Registry of research data repositories. Available: https://en.wikipedia.org/wiki/Registry_of_Research_Data_Repositories
- [296] DIKW pyramid. Available: <https://en.wikipedia.org/wiki/DIKW#p-search>
- [297] M. Hutson. (2018). Artificial intelligence has learned to probe the minds of other computers. Available: <http://www.sciencemag.org/news/2018/07/computer-programs-can-learn-what-other-programs-are-thinking>
- [298] Shetl-Optimized. Could a quantum computer have subjective experience? Available: <https://www.scottaaronson.com/blog/?p=1951>
- [299] NIPS 2017 videos. Available: <https://nips.cc/Conferences/2017/Videos>
- [300] S. Aaronson. PHYS771 lecture 4: Minds and machines. Available: <https://www.scottaaronson.com/democritus/lec4.html>
- [301] Generative adversarial network. Available: https://en.wikipedia.org/wiki/Generative_adversarial_network
- [302] A. Dimakis. (2018). Leveraging GANs to combat adversarial examples. Available: <http://approximatelycorrect.com/2018/03/02/defending-adversarial-examples-using-gans/>
- [303] Fermi paradox. Available: https://en.wikipedia.org/wiki/Fermi_paradox
- [304] The Fermi paradox — where are all the aliens? (1/2). Available: <https://www.youtube.com/watch?v=sNhhvQGsMEc&t=54s>
- [305] S. Webb, If the universe is teeming with aliens ... where is everybody?: seventy-five solutions to the fermi paradox and the problem of extraterrestrial life. Cham, Switzerland: Springer International Publishing, 2015.
- [306] The Fermi paradox II — solutions and ideas — where are all the aliens? Available: <https://www.youtube.com/watch?v=IfQkVqno-uI&t=167s>
- [307] Available: <https://seti.org/>
- [308] Mission. Available: <http://www.meti.org/mission>
- [309] I'm working on a universal language to let us speak to aliens. Available: <https://www.newscientist.com/article/2170452-im-working-on-a-universal-language-to-let-us-speak-to-aliens/>
- [310] Communication with extraterrestrial intelligence. Available: https://en.wikipedia.org/wiki/Communication_with_extraterrestrial_intelligence
- [311] Extraterrestrial life. Available: https://en.wikipedia.org/wiki/Extraterrestrial_life
- [312] Science fiction. Available: https://en.wikipedia.org/wiki/Science_fiction
- [313] M. Jones. Speak like a human to ET. Available: <http://www.meti.org/blog/speak-human-et>
- [314] A multidisciplinary critical approach to the 'mediocrity principle' as it has given rise to the fermi paradox. Available: https://www.academia.edu/11899447/A_Multidisciplinary_Critical_Approach_to_the_Mediocrity_Principle_as_it_has_Given_Rise_to_the_Fermi_Paradox
- [315] A. Kukla, "SETI: on the prospects and pursuitworthiness of the search for extraterrestrial intelligence," *Stud. History Philosophy Sci. Part A*, vol. 32, no. 1, pp. 31-67, 2001/03/01/ 2001.
- [316] A. Kukla, *Extraterrestrials: A Philosophical Perspective*. Maryland, US: Lexington Books, 2009.
- [317] S. Wells-Jensen. Could we learn E.T.'S language? Available: <http://meti.org/blog/could-we-learn-ets-language>
- [318] S. Wells-Jensen. Could we learn E.T.'S language? (2nd Part). Available: <http://meti.org/blog/could-we-learn-ets-language-2nd-part>
- [319] S. Wells-Jensen. Could we learn E.T.'S language? (Last Part). Available: <http://meti.org/blog/could-we-learn-ets-language-last-part>
- [320] O. Daniel. (2016). Building a language to communicate with extraterrestrials. Available: <https://www.theatlantic.com/science/archive/2016/04/math-language-extraterrestrials/477051/>
- [321] M. Rees. Why alien life will be robotic. Available: <http://nautil.us/issue/29/scaling/why-alien-life-will-be-robotic>
- [322] Pioneer plaque. Available: http://en.wikipedia.org/wiki/Pioneer_plaque
- [323] Voyager golden record. Available: https://en.wikipedia.org/wiki/Voyager_Golden_Record
- [324] The extraordinary link between deep neural networks and the nature of the universe. Available: <https://www.technologyreview.com/s/602344/the-extraordinary-link-between-deep-neural-networks-and-the-nature-of-the-universe/>
- [325] Stanford encyclopedia of philosophy. Philosophy of mathematics. Available: <https://plato.stanford.edu/entries/philosophy-mathematics/>
- [326] T. E. Feinberg and J. Mallatt, "The nature of primary consciousness. A new synthesis," *Consciousness and Cognition*, vol. 43, pp. 113-127, Jul. 2016.
- [327] Anthropomorphism. Available: <https://en.wikipedia.org/wiki/Anthropomorphism>
- [328] A. K. Willard and A. Norenzayan, "Cognitive biases explain religious belief, paranormal belief, and belief in life's purpose," *Cognition*, vol. 129, pp. 379-391, Nov. 2013
- [329] A. A. Harrison and J. T. Johnson, "ETI: Our first impressions," *Acta Astronaut.*, vol. 46, no. 10-12, pp. 713-718, Jun. 2000.
- [330] Alien language. Available: https://en.wikipedia.org/wiki/Alien_language
- [331] What if ET is an AI? Available: <https://aeon.co/essays/first-contact-what-if-we-find-not-organic-life-but-ets-ai>

- [332] Stanford encyclopedia of philosophy. Value Theory. Available: <https://plato.stanford.edu/entries/value-theory/>
- [333] Sir Arthur's quotations. Available: <https://www.clarkefoundation.org/about-sir-arthur/sir-arthurs-quotations/>
- [334] Existential risk threats to humanity's future. Available: <http://www.existential-risk.org/>
- [335] Stephen Hawking - will AI kill or save humankind? Available: <http://www.bbc.com/news/technology-37713629>
- [336] R. Hanson. (1998). The Great Filter - Are We Almost Past It? Available: <http://mason.gmu.edu/~rhanson/greatfilter.html>
- [337] S. Pappas. (2018) Why we might miss extraterrestrial life even if it's staring us in the face. Available: <https://www.space.com/40310-why-we-miss-extraterrestrial-life.html>
- [338] New report on malicious use of AI. Available: <https://www.fhi.ox.ac.uk/>
- [339] The Doomsday invention. Available: <https://www.newyorker.com/magazine/2015/11/23/doomsday-invention-artificial-intelligence-nick-bostrom>
- [340] S. Shostak. (2015). Should we keep a low profile in space? Available: <https://mobile.nytimes.com/2015/03/28/opinion/sunday/messaging-the-stars.html?referrer=&r=0>
- [341] Quick, How might the alien spacecraft work? Available: <http://blog.stephenwolfram.com/2016/11/quick-how-might-the-alien-spacecraft-work/>
- [342] K. Denning, "Unpacking the great transmission debate," *Acta Astronaut.*, vol. 67, pp. 1399-1405, 2010.
- [343] Omniscience. Available: <https://en.wikipedia.org/wiki/Omniscience>
- [344] C. Scharf. (2016). Is physical law an alien intelligence? Available: <http://nautil.us/issue/42/fakes/is-physical-law-an-alien-intelligence>
- [345] Language of mathematics. Available: https://en.wikipedia.org/wiki/Language_of_mathematics
- [346] M. Tegmark, *Our Mathematical Universe: My Quest for the Ultimate Nature of Reality*. London, UK: Penguin Books Limited, 2014.
- [347] M. Tegmark, *Life 3.0: Being Human in the Age of Artificial Intelligence*. London, UK: Penguin Books Limited, 2017.
- [348] M. Harris. (2017). Inside the first church of artificial intelligence. Available: <https://www.wired.com/story/anthony-levandowski-artificial-intelligence-religion/>
- [349] Nature. Available: <https://en.wikipedia.org/wiki/Nature>
- [350] Supernatural. Available: <https://en.wikipedia.org/wiki/Supernatural>
- [351] Time line of the far future. Available: <http://www.bbc.com/future/story/20140105-timeline-of-the-far-future>
- [352] A complete timeline of the future of our universe. Available: <https://futurism.com/the-future-of-the-universe/>
- [353] Exploring the impact of artificial intelligence. Available: <http://lcfi.ac.uk/>
- [354] Centre for the study of existential risk. Available: https://en.wikipedia.org/wiki/Centre_for_the_Study_of_Existential_Risk
- [355] Future of humanity institute. Available: https://en.wikipedia.org/wiki/Future_of_Humanity_Institute
- [356] B. Nick and Y. Eliezer. (2011). The ethics of artificial intelligence. Available: <http://faculty.smc.edu/acjamieson/s13/artificialintelligence.pdf>
- [357] R. V. Yampolskiy, "Artificial intelligence safety engineering: why machine ethics is a wrong approach," in *Philosophy and Theory of Artificial Intelligence*, V. C. Müller, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 389-396.
- [358] N. Bostrom. (2002). Existential risks: analyzing human extinction scenarios and related hazards. Available: <https://www.jetpress.org/volume9/risks.html>
- [359] Your questions answered on artificial intelligence. Available: <https://cosmosmagazine.com/technology/artificial-intelligence-your-questions-answered>
- [360] Do you trust this computer? Available: <http://doyoutrustthiscomputer.org/watch>
- [361] Elon Musk worries that AI research will create an 'Immortal dictator'. Available: <https://www.space.com/40231-elon-musk-immortal-artificial-intelligence-dictator.html>
- [362] Open letter on artificial intelligence. Available: https://en.wikipedia.org/wiki/Open_Letter_on_Artificial_Intelligence
- [363] Top experts warn against 'malicious use' of AI. Available: <https://phys.org/news/2018-02-experts-malicious-ai.html>
- [364] How the enlightenment ends. Available: <https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/>
- [365] At UN, robot Sophia joins meeting on artificial intelligence and sustainable development. Available: <https://news.un.org/en/story/2017/10/568292-un-robot-sophia-joins-meeting-artificial-intelligence-and-sustainable>
- [366] UN: artificial intelligence could destabilize world through unemployment and war. Available: <https://futurism.com/un-artificial-intelligence-could-destabilize-world-through-unemployment-and-war/>
- [367] UNICRI centre for artificial intelligence and robotics. Available: http://www.unicri.it/in_focus/on/UNICRI_Centre_Artificial_Robotics
- [368] How artificial intelligence could increase the risk of nuclear war. Available: <https://www.rand.org/blog/articles/2018/04/how-artificial-intelligence-could-increase-the-risk.html>
- [369] L. Muehlhauser. (2013). How well will policy-makers handle AGI? (initial findings). Available: <https://intelligence.org/2013/09/12/how-well-will-policy-makers-handle-agi-initial-findings/>
- [370] J.F. Pagel, and P. Kirshstein, "Chapter thirteen - machine consciousness," *Machine Dreaming and Consciousness*, pp. 175-185, 2017.
- [371] S. Cass. (2015). The body electric. Available: <https://spectrum.ieee.org/tech-history/silicon-revolution/the-body-electric>
- [372] Why AI can't solve everything. Available: <https://theconversation.com/why-ai-cant-solve-everything-97022>
- [373] Why does artificial intelligence scare us so much? Available: <https://www.livescience.com/62775-humans-why-scared-of-ai.html>
- [374] A. Thomas. Super-intelligence and eternal life: transhumanism's faithful follow it blindly into a future for the elite. Available: <https://theconversation.com/super-intelligence-and-eternal-life-transhumanisms-faithful-follow-it-blindly-into-a-future-for-the-elite-78538>
- [375] B. Mesko. (2016). The future of extremism: artificial intelligence and synthetic biology will transform terrorism. Available: <https://futurism.com/the-future-of-extremism-artificial-intelligence-and-synthetic-biology-will-transform-terrorism/>
- [376] Max Tegmark - How Far Will AI Go? Intelligible Intelligence & Beneficial Intelligence. Available:

- <https://www.youtube.com/watch?v=tAdvbaQQDA4&feature=em-uploademail>
- [377] Deep safety. Victoria's musings on AI safety, machine learning, and rationality. Available: <https://vkrakovna.wordpress.com/ai-safety-resources/>
- [378] "Virtual Immortality" by Robert Lawrence Kuhn. Available: <https://www.closetotruth.com/articles/virtual-immortality-robert-lawrence-kuhn>
- [379] Scott aaronson on the relevance of quantum mechanics to brain preservation, uploading, and identity. Available: <http://www.brainpreservation.org/scott-aaronson-on-the-relevance-of-quantum-mechanics-t>
- [380] Future of life institute. Available: <https://futureoflife.org/>
- [381] Carl Sagan's pale blue dot official. Available: <https://www.youtube.com/watch?v=GO5FwsblpT8>
- [382] How far can we go? Limits of humanity. Available: <https://www.youtube.com/watch?v=ZL4yYHdDSWs&t=35s>
- [383] Scale of the universe. Putting the universe into perspective. Available: <https://www.khanacademy.org/science/cosmology-and-astronomy/universe-scale-topic>
- [384] Scale of the universe. Available: <http://scaleofuniverse.com/>
- [385] The universe is not a simulation, but we can now simulate it. Available: <https://www.quantamagazine.org/coder-physicists-are-simulating-the-universe-to-unlock-its-secrets-20180612>
- [386] Artificial intelligence in fiction. Available: https://en.wikipedia.org/wiki/Artificial_intelligence_in_fiction#Logic_machines
- [387] P. Torres. (2018). Why we should think twice about colonizing space. Available: <http://nautil.us/blog/why-we-should-think-twice-about-colonizing-space>
- [388] NASA. Technology drives exploration. Available: <https://www.nasa.gov/topics/technology/index.html>
- [389] Las cumbres observatory: many eyes - one vision. Available: <https://lco.global/>
- [390] A digital copy of the universe, encrypted. Available: <https://www.quantamagazine.org/a-digital-copy-of-the-universe-encrypted-20131002/>
- [391] S. Seager, "The future of spectroscopic life detection on exoplanets," *Proc. Nat. Academy Sci.*, vol. 111, no. 35, pp. 12634–12640, Aug. 2014.
- [392] Microscope. Available: <https://en.wikipedia.org/wiki/Microscope>
- [393] Magnetic resonance imaging. Available: https://en.wikipedia.org/wiki/Magnetic_resonance_imaging
- [394] Super-resolution microscopy. Available: https://en.wikipedia.org/wiki/Super-resolution_microscopy
- [395] Laniakea: Our home supercluster. Available: <https://www.youtube.com/watch?v=rENyYRwxpHo>
- [396] Histogramy. Available: <http://histography.io/>
- [397] Kurzgesagt – in a nutshell. Available: <https://www.youtube.com/channel/UCsXVk37bltHxD1rDPwtNM8Q>
- [398] **Infographic, View the Future of Design and Collaboration** Available: <http://www.visualcapitalist.com/visualizing-future-design-collaboration/>
- Baars, B. J., How brain reveals mind: Neuroimaging supports the fundamental role of conscious experience. *Journal of Consciousness Studies*, 10(9–10), 100–121.
- Baars, B. J., Banks, W. P., & Newman, J. (Eds.). *Essential sources in the scientific study of consciousness*, Cambridge, MA: MIT Press, 2003
- Bohm, D., and Hiley, B.J., *The Undivided Universe*. Routledge, London. 1993
- Bostrom, N., *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, 2014
- Chalmers, D. J., *The conscious mind In search of a fundamental theory*. Oxford University Press, New York, 1996
- Chalmers, David, "Panpsychism and Panprotopsychism", in *Alter and Nagasawa 2015*: 246–276.
- Churchland, M Churchland, Paul (1986) [1979]. *Scientific Realism and the Plasticity of Mind*. Cambridge Studies in Philosophy Cambridge, UK: Cambridge University Press (2007).
- Churchland, P. *Neurophilosophy: Toward a Unified Science of the Mind-Brain*. Cambridge, Massachusetts: The MIT Press, 1986
- Churchland, P., *Neurophilosophy at Work*. Cambridge, MA, USA: MIT Press 2007 Churchland, Paul (1981). "Eliminative Materialism and the Propositional Attitudes". *Journal of Philosophy*. 78 (2; February): 67–90. Retrieved 11 February 2017. See also the PDF version at K. A. Akins' web pages at Simon Fraser University.
- Crick, F., & Koch, C., A framework for consciousness. *Nature Neuroscience*, 6, 119–126. 2003
- Damasio, A., *The feeling of what happens: Body and emotion in the making of consciousness*. Harvest Books. 2000
- Davis, M., *The Universal Computer: The Road from Leibniz to Turing*, available https://www.amazon.com/gp/product/0393047857/ref=pe_160100_267887200_em_sp_C_1p_1_ti
- Dennett DC. *Consciousness explained*. Boston (MA): Little Brown; 1991.
- Dennett, D., *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*, MIT Press, Cambridge, MA, 2005.
- Dyson, F. *Disturbing the Universe*, New York: Harper & Row, 1979
- Heijenoort, Jean van, *From Frege to Gödel: A Source Book in Mathematical Logic*, available https://www.amazon.com/gp/product/0674324498/ref=pe_160100_267887200_em_1p_2_ti
- Edelman, G. M. (1987). *Neural Darwinism: The theory of neuronal group selection*. New York: Basic Books.
- Edelman, G. M. (1989). *The remembered present*. New York, NY: Basic Books.
- Edelman, G. M. (1993). *Neural Darwinism: Selection and reentrant signaling in higher brain function*. *Neuron*, 10, 115–125.
- Edelman, G. M. (2003). *Naturalizing consciousness: a theoretical framework*. *Proceedings of the National Academy of Sciences, USA*, 100, 5520–5524.
- Edelman, G. M. (2004). *Wider than the sky: The phenomenal gift of consciousness* New Haven, CT: Yale University Press.

BIBLIOGRAPHY: SELECTED FURTHER READING

Aaronson, S. *Quantum Computing Since Democritus*, Cambridge University Press, 2013

Baars, B. J. *A cognitive theory of consciousness*, New York, NY: Cambridge University Press, 1988

- G. Edelman, G. Tononi, *A Universe Of Consciousness How Matter Becomes Imagination* Basic Books, 2008
- Everett H (1973) The theory of the universal wave function. In *The many-worlds interpretation of quantum mechanics*, ed. BS DeWitt and N Graham, Princeton University Press, Princeton, New Jersey
- Eddington, A. *The Nature of the Physical World*. Cambridge University Press, Cambridge 1928
- Eccles JC. Evolution of consciousness. *Proc Natl Acad Sci USA* 1992;89:7320–4.
- Feynman, R. P. Simulating physics with computers. *Int. J. Theor. Phys.* 21, 467 (1982)
- Feynman, R.P. Quantum mechanical computers. *Found Phys* 1986; 16(6):507–31.
- Fredkin, E., *Finite Nature Hypothesis, A New Cosmogony*, available http://www.digitalphilosophy.org/wp-content/uploads/2015/07/new_cosmogony.pdf
- Gödel, K, *On Formally Undecidable Propositions of Principia Mathematica*, available https://www.amazon.com/dp/0486669807/ref=rdr_ext_tmb
- Hanson, R., *The Age of Em: Work, Love and Life When Robots Rule the Earth*, Oxford University Press, 2016
- Hartshorne, C. 1937. *Beyond Humanism*. New York: Willett, Clark & Company.
- Hartshorne, C. 1950. “Panpsychism.” In *A History of Philosophical System*. Ed. V. Ferm. New York: Philosophical Library.
- Hameroff, S., *Time, Consciousness and Quantum Events in Fundamental Spacetime Geometry*, https://link.springer.com/chapter/10.1007/978-94-010-0155-7_9
- Hameroff, S., Penrose, R., *Physics of Life Reviews*, 11 (2014) 39–78 77
- Heijenoort, J van, *From Frege to Gödel: A Source Book in Mathematical Logic*
- Koch, C., *Consciousness: Confessions of a Romantic Reductionist*, MIT Press 2017
- Koch, C. *The quest for consciousness: a neurobiological approach*, Roberts and Co.; 2004.
- Koch, C., Tononi G. *Can a Photodiode Be Conscious?* New York Review of Books. New York, NY: Rea S. Hederman, 2013
- Kneale, William & Martha, *The Development of Logic*, Oxford University Press, 1962
- Kulka, A., *Extraterrestrials: A Philosophical Perspective*, Lexington Books, 2009
- Moore, C., Mertens, S., *The Nature of Computation*, Oxford University Press, 2011
- Nagel, T., “What is it like to be a bat?” *Philosophical Review*, 83: 435–456, 1974
- Nagel, T., “Panpsychism”, in *Mortal Questions*, Cambridge: Cambridge University Press, 1979
- Neumann, J. von , *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, Princeton. 1955 German original, *Die mathematischen Grundlagen der Quantenmechanik*. Springer, Berlin, 1932.
- Penrose, R.), in *Quantum Theory and Beyond*, (Ed. E.A. Bastin), Cambridge, 1971 University Press, Cambridge, U.K. 2009
- Penrose R., *The emperor’s new mind: concerning computers, minds, and the laws of physics*, Oxford: Oxford University Press; 1989.
- Penrose R. *Shadows of the mind: an approach to the missing science of consciousness*, Oxford: Oxford University Press; 1994.
- Penrose R. *Cycles of time: an extraordinary new view of the universe*. London: Bodley Head, 2010.
- Penrose, R., *Faith, Faith, and Fantasy in the New Physics of the Universe*, Princeton University Press, 2016
- Pockett S. *The Nature of Consciousness: A Hypothesis*, iUniverse.com, Lincoln; NE, 2000
- Popper, K.R., Eccles, J.C. *The Self and Its Brain*, Springer, Berlin, 1977
- Pribram, K. *Languages of the Brain*, Prentice-Hall, Englewood Cliffs, 1971
- Quantum Approaches to Consciousness*, Stanford Encyclopedia of Philosophy, available online, <https://plato.stanford.edu/entries/qtconsciousness/>
- Russell, B., *Introduction to Mathematical Philosophy*, London: George Allen & Unwin. (ISBN 0-415-09604-9 for Routledge paperback)
- Searle J. R., *Can Information Theory Explain Consciousness?* New York Review of Books, New York, NY: Rea S. Hederman, 2013
- Tegmark, M., *Consciousness as a State of Matter*, available online <https://arxiv.org/pdf/1401.1219.pdf>, Version 3, 18 Mar 2015, pg. 1
- Tegmark, M., *Our Mathematical Universe: My Quest for the Ultimate Nature of Reality*, Knopf, New York, 2014
- Quine, W.V., *Word and Object*, MIT Press, 1960
- Quine, W.V. ‘Two Dogmas of Empiricism’, *Philosophical Review*, 1951,
- Quine, W. V. “Ontological relativity and other essays”, *Central Works of Philosophy V5: Twentieth Century: Quine and After*, John Shand, 1969
- Schneider, S., editor, *Science Fiction and Philosophy: From Time Travel to Superintelligence*, Wiley, 2009
- Schneider, S, Velman, M., *The Blackwell Companion to Consciousness*, 2017
- Seth, A, *The strength of weak artificial consciousness*, *International Journal of Machine Consciousness*, 2009
- Seth A. K. “Explanatory correlates of consciousness: theoretical and computational challenges”. *Cognitive Computation*, 1, 50–63 10.1007/s12559-009-9007, 2009
- Seth, A., Clowes, R.W., *Axioms, properties and criteria: roles for synthesis in the science of consciousness*, *Artificial intelligence in medicine*, 2008, Elsevier
- Shanahan, M., *Embodiment and the Inner Life: Cognition and Consciousness in the Space of Possible Mind*, Oxford University Press. 2010
- Shanahan, M., *The Technological Singularity*. MIT Press, 2015
- Skrbina, D., *Panpsychism in the West*, Cambridge, MA, MIT Press, 2005

Shannon, C., & Weaver, W. The mathematical theory of communication. Urbana, IL: The University of Illinois Press, 1949

Smith, Brian Cantwell, On the Origins of Objects, Bradford Books, 1996

Spencer-Brown, G., Laws of Form, Allen & Unwin, 1969.

Stapp, H.P.. "A quantum theory of the mind-brain interface". In Mind, Matter, and Quantum Mechanics, Springer, Berlin, pp. 145-172, 1993

Strawson, G. 2006. "Realistic Monism: Why Physicalism Entails Panpsychism." Journal of Consciousness Studies, 13(10-11).

Strawson, G. et al. 2006. Consciousness and its Place in Nature: Does Physicalism entail Panpsychism? Exeter, UK: Imprint Academic.

Alfred Tarski, Andrzej Mostowski and Raphael M. Robinson, Undecidable Theories: Studies in Logic and the Foundation of Mathematics Dover Publications , August. 19 2010

Tegmark, M., The Mathematical Universe, available online, arXiv:0704.0646 , DOI10.1007/s10701-007-9186-9

Tononi, G., Massimini, M., Sizing up Consciousness: Towards an objective measure of the capacity for experience, 2018

Tononi, G., Steven Laureys, Olivia Gosseries, The Neurology of Consciousness: Cognitive Neuroscience and Neuropathology, 2015

Vimal, R., Biophysics Of Consciousness: A Foundational Approach, editors Poznanski Roman R, Tuszynski, Jack A, Feinberg, Todd E., World Scientific, 2016

STEPHEN B. RIPLEY President and Research Director, Futures & Forecasts, Inc., Vancouver, Canada, 1975-Yale University, New Haven, Connecticut, USA (B.A., philosophy, 1962), University of Alberta, Calgary, Canada, graduate scholarship, philosophy, 1964-65.

Research interests: Artificial Intelligence, Computer Sciences, Biochemistry, Biology, Earth Sciences