The M Cognitive Meta-Architecture as Touchstone for Standard Modeling of AGI-Level Minds

Selmer Bringsjord[(✉)], James T. Oswald, Michael Giancola, Brandon Rozek, and Naveen Sundar Govindarajulu

Rensselaer AI & Reasoning (RAIR) Lab, Department of Cognitive Science, Department of Computer Science, Lally School of Management (1st auth), Rensselaer Polytechnic Institute (RPI), Troy NY 12180 USA
{selmer.bringsjord,james.oswald.111,mike.j.giancola,rozek.brandon,naveen.sundar.g}@gmail.com

Abstract. We introduce rudiments of the cognitive meta-architecture M (majuscule of μ and pronounced accordingly), and of a formal procedure for determining, with M as touchstone, whether a given cognitive architecture X_i (from among a finite list $1 \ldots k$ of modern contenders) conforms to a minimal standard model of a human-level AGI mind. The procedure, which for ease of exposition and economy in this short paper is restricted to arithmetic cognition, requires of a candidate X_i, (1), a true biconditional expressing that for any human-level agent a, a property possessed by this agent, as expressed in a declarative mathematical sentence $s(a)$, holds if and only if a formula $\chi_i(\mathfrak{a})$ in the formal machinery/languages of X_i holds as well ($\mathfrak{a}$ being an in-this-machinery counterpart to natural-language name a). Given then that M is such that $s(a)$ iff $\mu(\mathfrak{m})$, where the latter formula is in the formal language of M, with $\mathfrak{m}$ the agent modeled in M, a minimal standard modeling of an AGI-level mind is certifiably achieved by X_i if, (2), it can be proved that $\chi_i(\mathfrak{a})$ iff $\mu(\mathfrak{a})$. We conjecture herein that such confirmatory theorems can be proved with respect to both cognitive architectures NARS and SNePS, and have other cognitive architectures in our sights.

Keywords: standard modeling of AGI-level minds · cognitive architectures · computational logic.

1 Arithmetic as the Initial Target

Despite florid heraldry from Kissinger et al. [14] announcing an "intellectual revolution" caused by the arrival of ChatGPT and its LLM cousins of today, we know that AGI has not arrived. This is so because, as Arkoudas [3] has elegantly pointed out in a comprehensive analysis, ChatGPT doesn't know that 1 is not greater than 1, and surely AGI subsumes command of elementary arithmetic on the natural numbers.[1] We do not pick the domain of arithmetic here randomly.

[1] This example is but the tip of an iceberg of negative knowledge in the realm of mathematics for this and indeed all present and foreseeable LLMs, as Arkoudas

Arithmetic, and more generally all or at least most of logico-mathematics, is by our lights the only cross-cutting and non-negotiable space we can presently turn to in order to at least be in position to judge whether some artificial agent qualifies as having AGI versus merely AI. Given this, we turn first to arithmetic cognition to enable us to share our formal procedure for using the cognitive meta-architecture M as a touchstone for determining whether a candidate cognitive architecture conforms to a minimal standard modeling of AGI-level minds.

2 The Formal Procedure, for Arithmetic Cognition

2.1 Peano Arithmetic to Anchor Arithmetic Cognition

To anchor arithmetic cognition as a proper part of mathematical cognition at the human level, we resort herein to simple arithmetic with only addition and multiplication. The particular axiom system we bring to bear is 'Peano Arithmetic,' or just — to use the conventional label — **PA**. Unassuming as it may be, it has a storied place in the history of logic and mathematics, serving as the basis for such stunning results as Gödel's incompleteness theorems.[2] In particular, we shall employ herein a simple theorem in **PA**, viz., $\vdash 2 + 2 = 4$. In the general form of our procedure, not merely arithmetic cognition, but all of mathematical cognition reduced to formal logic by reverse mathematics will be in play, which means that not just the likes of 2+2=4 but any statements provable from the axioms (i.e. $\mathbf{PA}^{\vdash}$) known to be sufficient for all of mathematics, as charted by the definitive [24], will be fair game.

2.2 Definition of the Procedure

Let $s(a)$ be a mathematical declarative sentence involving both a mathematically cognizing agent a and a single purely arithmetic proposition believed by a. Such sentences typically draw from both natural language (e.g. English) and formal languages. Here's an example of such a sentence: "Gödel believed that first-order logic is complete." We know he believed this because his dissertation centered

shows/explains. Note that Bubeck et al. [8] have made the figurative claim that GPT-4 has — and we quote — "sparks of AGI." We don't know what this metaphorical claim means mathematically (thus confessedly find little meaning in it), but clearly by conversational implication these authors would themselves agree that while GPT-4 is an AI, it's an AGI. If x has sparks of being an R, then x isn't an R — this is the principle at the root of the implication here.

[2] We shall not spend the considerable time that would be needed to list the (countably infinite) axioms, and explain them. Readers can consult the elegant [9] for nice coverage of **PA** (and illuminating commentary on this axiom system). There are theories of arithmetic even simpler than **PA**, because **PA** includes an axiom relating to mathematical induction, and the simpler systems leave this axiom aside. For example, readers unfamiliar with mathematical induction can, if motivated, consult the induction-free theory of arithmetic known as 'Robinson Arithmetic,' or sometimes just as 'Q;' for elegant coverage, see [5].

around the landmark proof of this completeness. But this example is far too complex for our present limited purposes. Accordingly, turning to **PA**, here's a much simpler example of a form that will guide us, put in the present tense:

"Gödel believes that 2+2=4."

The general form is that some agent a is denoted, that agent has the epistemic attitude of belief, and the target of that belief is a proposition, expressible in **PA**, that 2+2=4. We shall denote the form this way: $s(a)$, to indicate that our sentence form must involve an agent a; we leave belief and the believed proposition implicit.

Next, let '$\mu(\mathfrak{m})$' be a formula in M, in a suitable formal language that logicizes $s(a)$. Minimally, this language will have an epistemic modal operator for belief, and will be able to encode arithmetic propositions from natural language. Therefore, the language will need to be a quantified modal one whose extensional component is at least first-order logic. Now, the following is by inspection the case with respect to μ:

$$s(a) \text{ iff } \mu(\mathfrak{m}).^3$$

Next, let X_i be any cognitive architecture that aspires to enable standard modeling (and simulation) of AGI-level minds. What is needed from this cognitive architecture is, (1), the truth of this biconditional:

$$s(a) \text{ iff } \chi_i(\mathfrak{c}).$$

Standard modeling of an AGI-level mind, given the foregoing, is achieved by X_i if, (2), it can be proved that

$$\chi_i(\mathfrak{c}) \text{ iff } \mu(\mathfrak{m}).^4$$

We conjecture that such confirmatory theorems can be proved with respect to both cognitive architectures NARS and SNePS, to which, resp., we shortly turn. But first we give a very quick overview of the nature of M itself.

3 The M Cognitive Meta-Architecture: Key Attributes

M is not a new cognitive architecture intended and designed to compete with the likes of Soar and ACT-R and so on as a platform to model and simulate human and/or AGI cognition. There are innumerable competing architectures in play

[3] The formula in the case of M itself is

$\mu(\mathfrak{m}) :=$ `(believes! m t (= (+ (s (s 0)) (s (s 0))) (s (s (s (s 0))))))`,

where `s` is the successor function and 0 is primitive, but technical details regarding M are outside of current scope.

[4] The kernel of the procedure just described was first adumbrated in [4].

today [15], all directly reflecting the particular predilections of their human creators and developers.[5] M is for assessing and harmonizing these "particularist" architectures at a meta level, and is marked by the following three distinguishing attributes:

- *Non-Partisan.* M is not designed to advance any particular convictions about the nature of cognition, and is in this regard unlike the typical cognitive architecture. To mention just one example, certainly Soar was originally conceived to commit to and build upon the conviction that a key part of human cognition centers around *condition-action rules.* Many other examples of particularist convictions could be enumerated here for many competing cognitive architectures. In stark contrast, M reflects the attitude that any partisan advocacy militates against standardization; instead, the attitude is to move as soon as possible to formalization using the discipline of formal logic. Of course, no *particular* logic is to be locked into in any way as long as its a quantified modal one whose extensional component is at least first-order logic.
- *Thoroughgoingly Formal: Axiomatic and Theorem-based.* M is inseparably aligned with a purely formal view of science and engineering, according to which whatever phenomena is observed and to be deeply understood and predicted should be axiomatized. The axiomatization of mathematics is now mature (and is the initial focus in the application of M as touchstone for whether a given cognitive architecture can minimally be used for standard modeling and simulation of AGI-level mind), and the axiomatization of physics is now remarkably mature; consider for example that not only classical mechanics is long done [19], but special relativity is largely captured [2], and advances are fast being made on general relativity and quantum mechanics. M is based on the assumption that this level of high maturity should now be applied to intelligence, so that matters can be theorem-based.
- *Minimalist.* Given all the resources formal science offers for capturing cognition, use of M is guided by a minimalist approach. The smaller and simpler is the logical system that can be used to capture a target, the better.

4 Applying the Procedure

In this short paper, we cannot fully chronicle the application of our procedure to candidate cognitive architectures. But we attempt to partially justify our optimism that both the cognitive architectures NARS and SNePS will yield in each case the needed theorem by virtue of which standard modeling is confirmed.

4.1 Exploration of NARS

What is $\chi_i(\mathfrak{a})$ for NARS? The sentence $s(a)$ says that a believes 2+2=4 to be a true statement, and we shall assume that counterpart to agent a is the NARS agent $\mathfrak{n}$, and that the formula ν is the in-system counterpart to s. Next, we note that instead of statements NARS has *judgments*: statements with associated fuzzy truth-values, consisting of a frequency $f \in [0, 1]$ that represents a degree

[5] We conjecture that the set of all of these architectures is pairwise inconsistent, but leave this disturbing prospect aside for subsequent investigation via M.

of belief in the underlying statement, and a confidence $c \in [0,1]$ representing how stable the belief is (Definition 3.3 in [25]). For our target of eventually demonstrating $\nu(\mathfrak{n})$ iff $\mu(\mathfrak{m})$ it will suffice[6] that we define $\nu(\mathfrak{n})$ as the statement "The NARS agent $\mathfrak{n}$ believes the judgment $2+2=4$ with a frequency of 1 (there is only positive evidence for the statement)." Formalizing this further, a NARS agent is said to *believe* a *judgment* iff it is either an *experience*, a judgment provided to the system directly, or a statement that can be derived from experiences (Definition 3.7 in [25]). Thus $\nu(\mathfrak{n})$ for a NARS agent $\mathfrak{n}$ is true by providing $2+2=4$ as a standalone experience (in our case perhaps provided by the theoretical perception system outlined in [26]).[7] Finally, the representation of the actual statement $2+2=4$ can be accomplished in a number of ways, as NARS supports the representation of relational terms that can represent arbitrary n-ary relations between terms that represent objects. One example of this representation in Narsese is $< (* \quad 2 \quad 2 \quad 4) \rightarrow add >$ where *add* is a term representing a relation between two summands and a sum.

Having defined $\nu(\mathfrak{n})$, we can turn to a proof sketch for $\nu(\mathfrak{n})$ iff $\mu(\mathfrak{m})$. There are multiple approaches to this proof, one particularly formal variant would be expressing NAL in one of our *cognitive calculi* — a specialized type of logical system for Theory-of-Mind reasoning[8] — in a higher-order logic and proving a bridge theorem. Instead for economy we opt for an intuitive proof based on theoretical idealized perception systems for NARS and M. For the forward direction of biconditional proof we assume $\nu(\mathfrak{n})$. By our above definition, $\nu(\mathfrak{n})$ iff the agent $\mathfrak{n}$ experiences $2+2=4$ or has experiences that deductively[9] lead to the conclusion $2+2=4$ with frequency 1 in $\mathfrak{n}$. Under idealized perception, this implies the existence of external representations of either s that $2+2=4$, or a set of statements $\mathcal{S}$ that imply s. The existence of these external representations means that an M agent $\mathfrak{m}$ under idealized perception would also perceive s, $\mathbf{P}(\mathfrak{m},\cdot,s)$ or perceive the set of $\mathcal{S}$, $\bigwedge_{e\in\mathcal{S}}\mathbf{P}(\mathfrak{m},\cdot,e)$. Since many

6 We hold that confidence is irrelevant here as it is a temporal property which only impacts how likely the system is to change its mind, which has use for nonmonotonic reasoning but is irrelevant to our current deduction-only explorations.

7 Additionally we could proceed by providing any number of experiences that allow the system to derive 2+2=4 as long as they allow the system to derive 2+2=4 with frequency 1.

8 Cognitive calculi build off of the notion of traditional logical systems, which consist of a formal language $\mathscr{L}$, a set of inference schemata $\mathscr{I}$, and a formal semantics $\mathscr{S}$. The most notable distinguishing factors of cognitive calculi are (1) they contain modal operators for mental states, e.g., perception, belief, obligation; and (2) they contain no model-based semantics; instead the semantics of formulae are purely inference-theoretic. That is, the semantics are expressed exclusively through the inference schemata $\mathscr{I}$. For a longer exposition of exactly what a cognitive calculus is and isn't, we refer the interested reader to Appendix A of Bringsjord et al. [7].

9 Abductive and inductive reasoning in NARS have the resulting frequency of the conclusion depend on confidence values influenced by a system parameter; as this can be arbitrary, this will not guarantee the preservation of frequency of 1 for conclusions using these modes of reasoning, thus only deductive reasoning applies here.

standard cognitive calculi have inference schemata allowing perceived statements to become believed statements, and others allowing propositional reasoning on beliefs, in the first case $\mathbf{P}(\mathfrak{m}, \cdot, s) \rightarrow \mathbf{B}(\mathfrak{m}, \cdot, s)$, and in the second $\bigwedge_{e \in \mathcal{S}} \mathbf{P}(\mathfrak{m}, \cdot, e) \rightarrow \bigwedge_{e \in \mathcal{S}} \mathbf{B}(\mathfrak{m}, \cdot, e) \rightarrow \mathbf{B}(\mathfrak{m}, \cdot, s)$ which is the definition of $\mu(\mathfrak{m})$. For the backward direction of the biconditional proof, we assume $\mu(\mathfrak{m})$ to derive $\nu(\mathfrak{n})$ using the same argument outlined for the forward direction.

We thus claim that $\nu(\mathfrak{n})$ iff $\mu(\mathfrak{m})$, which confirms that NARS conforms to a minimal standard modeling of AGI-level minds.

4.2 Exploration of SNePS and GLAIR

SNePS is a KRR system, ultimately in fact a logic [22], that can be used as either a standalone system or inside others; GLAIR is a cognitive architecture designed by SNePS scientists that uses SNePS for KRR [23]. As $\gamma(\mathfrak{g})$ for GLAIR (or any agent using SNePS for KRR) depends solely on representation within SNePS at the knowledge layer of a GLAIR agent [23], we generalize and refer to $\gamma(\mathfrak{g})$ for any arbitrary agent having SNePS under its hood, henceforth referred to simply as *SNePS agents*. Any statement within a SNePS system is said to be *believed* by the system. Figure 1 shows a representation of the statement $2+2=4$ in SNePS as a network. [11] makes a distinction between a SNePS agent understanding that 2+2=4 as declarative knowledge versus understanding what 2+2=4 means as semantic knowledge. In this language, $\gamma(\mathfrak{g})$ can be interpreted purely in the sense of the representation of the declarative knowledge and is thus satisfied by the representation in Figure 1.

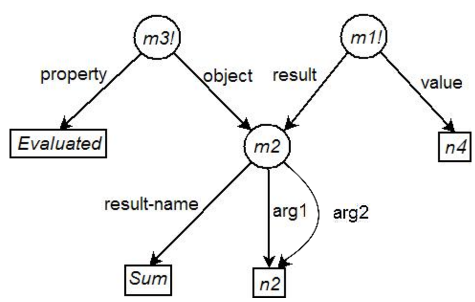

Fig. 1. A SNePS agent's belief that 2+2=4. where $m2$ is the functional term representing a resultant Sum, from $n2$ twice, $m1$ is the proposition that $m2$ evaluates to $n4$, and $m3$ is the proposition that $m2$ has a value. (Adapted from Figure 4.1 in [11].)

We claim that $s(a)$ iff $\gamma(\mathfrak{g})$ is true by construction. Unfortunately given current space constraints, $\gamma(\mathfrak{g})$ iff $\mu(\mathfrak{m})$ is non-trivial. Since M has a purely inferential semantics, and since SNePS allows inferences to be systematically carried out, we prove an inference-theoretic interpretation of the biconditional by showing that given some context in which $\mu(\mathfrak{m})$ is deduced (in the fashion of [6]),

$\gamma(\mathfrak{g})$ can be the conclusion of valid reasoning in SNePS that uses a counterpart of this context. The left-to-right direction follows the same strategy. We thus assert that SNePS too conforms to a minimal standard modeling of AGI-level minds.

5 Related Work

Commendably, Laird et al. [16] launched a search for a standard model of the *human* mind. But their approach and ours are starkly divergent. We have no particular interest in the human mind or its embodiment in the form of earthly brains, which we regard to be adventitious relative to AGI at the human level and above. Nonetheless, realistically, at least philosophically speaking, there will be in the minds of some AI theorists overlap between the Lairdian approach and the approach we introduce herein, so we point out a second divergence: Their approach is informal, while ours is formal, i.e. theorem-driven. For good measure, a third aspect of divergence is found in the fact that while we regard the "best bet" for commonality of AGIs to be found in the arena of logic and mathematics, cognition in this area is regarded by Laird et al. to be cognitively recherché, which is borne out holistically by the absence of any discussion whatsoever of logico-mathematical cognition in [16], and more specifically by the fact that their proposed "standard model" constraints have nothing whatsoever to do with reasoning, and instead consist of the four pillars of "perception/motor", "learning," "memory and content," and "processing." Reasoning, including reasoning in connection with logico-mathematical cognition over content in formal languages, would only perhaps arise in conception in secondary, epiphenomenal fashion under the roof held up by their quartet of pillars.

We suspect some readers will think that knowledge graphs and description logics are related to our proposed procedure with M. However, care must be taken when considering this kind of work.

In practice, most knowledge-graph systems can be represented by a decidable description logic [10] (e.g. by $\mathcal{ALC}$, or $\mathcal{SHOIN}$, which are standards for most knowledge graphs), but such logics cannot capture **PA**, and they cannot capture epistemic attitudes about theorems of this axiom system. The reason is that description logics are proper fragments of first-order logic (FOL), and thus cannot express **PA**, which requires full FOL and is by Church's Theorem undecidable. Formalizing mathematics is known to require at a minimum third-order logic (M's cognitive calculi include quantified modal third-order logic) [24]. What thus may seem to be work related to ours is in the case of knowledge graphs and description logic actually not. However, our procedure can easily handle weaker, decidable theories of arithmetic, such as Presburger Arithmetic, and as a matter of fact the particular sentence $s(a)$'s component '2+2=4' is a valid statement in both Peano *and* Presburger Arithmetic.

[10] Some description logics have been discovered to be undecidable [21, 20]. However, the core focus in the description logic community is on finding decidable fragments.

6 Objections

We anticipate many objections to our new approach. We rapidly encapsulate under current space constraints two, and briefly reply to each.

6.1 "But What About Purely Numerical Approaches to AGI?"

It will be said against us: "There are approaches to rigorously capturing general intelligence at the human level and above that make no reference to the axiomatized declarative content of **PA**, let alone to the additional axiom systems to which you implicitly refer when invoking reverse mathematics for your standardization program (e.g. see [13]). Your approach is hence idiosyncratic at best, and tendentious at worst."

In reply, the key question is what those aiming at securing AGI via approaches that exclude the standardization we advocate will settle for when an artificial agent is challenged to demonstrate the power and accuracy of its mathematical cognition. Suppose that some artificial agent purportedly not only believes that 2+2=4, but purportedly has command over **PA** overall. The key question, then, when narrowed, is: Would purely external behavior of the right sort be sufficient, or must there be some underlying structures and content associated with the behavior that enable proving a connection to formulae like μ? Large Language Models (LLMs), for example, provide an excellent context for asking this question. Suppose an LLM agent known colloquially by the name 'Larry,' based purely on deep learning, and thus completely bereft of any formulae that encode members of $\mathbf{PA}^{\vdash}$ (the closure under deduction of **PA**), is able to generate all sorts of sentences like the sentence s from above, but also more complicated ones, because saying any number-theoretic theorem is possible for this LLM. Let $s'(\mathit{Larry})$ be "I believe that every cubic number is the sum of n consecutive odd numbers," where the indirect indexical refers to 'Larry.' And suppose that many, many other sentences are generated by Larry on this topic, where this generation is syntactically flawless, but is by definition based exclusively on underlying numerical data processing. Under this supposition, proving a bridging biconditional that links from the LLM Larry to formulae in M is impossible. This is an empirical fact.

We see this as most unfortunate, for the simple reason that science explains by virtue of finding formal theory that explains observed phenomena; physics is the paradigmatic case in point. In the case of the LLM that is ChatGPT, the empirical fact that deep formal science of the type that has always been the "golden goal" of science is completely excluded as it is in the case of Larry, has been noted recently by Wolfram [28]. Hence, the blockage by the impenetrable nature of LLMs for our M-based procedure is just something we must accept, with all of rigorous science.

6.2 "Math is Merely Manufactured"

The objection against us here can be summarized thus: "Using mathematical cognition as the cornerstone of a test for standard modeling and simulation of

AGI-level minds bestows upon such cognition a kind of 'ground-truth' status. But mathematics is essentially a symbol-manipulation game legislated by human beings, as explained in [17]."

As all or at least most readers will know, while the view espoused in this objection has been defended by serious scholars (e.g. [17]), this is by no means a consensus view. There are many well-known problems that afflict the view, for instance the apparent fact that math stunningly corresponds to the behavior of the natural world [27], while formal logic has a parallel relationship with computation [12]. Yet our position, in keeping with the non-partisan nature of M itself, is to leave such debates to the side, in favor of simply observing that at the very least, going with mathematical cognition as a starting place for trying to establish a plumb-line standard modeling of AGI-level minds is rational, since if any part of cognition is likely to span minds in general it is mathematical cognition — rather than perception, motor control, natural language usage, etc.

7 Conclusion and Next Steps

Immediate next steps include delivering full proofs of our conjectures with respect to the NARS and SNePS, and expanding our procedure to include cognitive architectures beyond these two cognitive architectures. Two obvious targets are Soar and ACT-R, the latter of which promises to qualify as standard by our metrics in no small part because ACT-R has already been considered from the standpoint of formal logic (at least at the level of first-order logic; see [10, 1]). We don't know what will happen in the case of Soar.[11]

A significant challenge awaits us when our procedure is expanded beyond mathematical cognition into other parts of AGI-level cognition. We must be able to draw from logic-based machinery to for example formalize communication so that our key biconditionals can go through in this realm. The most severe challenge to our procedure will arise, we believe, in the case of robust attention and perception, and, having devoted time to considering perception in connection with NARS (as seen above), we are studying the attention/perception-centric cognitive architecture ARCADIA [18] now from a formal point of view.

[11] Some readers of earlier drafts of the present paper have asked us whether our procedure can be applied not just to cognitive architectures, but to artificial agents in general — for instance to the LLM agents in today's headlines. This question, alas, is at once tricky and straightforward. If the question is about *pure* LLMs, the question is straightforward, and easily answered in the negative, since cognitive attitudes directed at declarative content qua declarative content within the theory of elementary arithmetic (the full closure of **PA** under standard first-order deduction) cannot exist in such a system, which operates exclusively over data derived by tokenizing and vectorizing etc. away from quantifier-rich formula. Things become tricky when one sees that LLMs are increasingly getting "glued" to outside intelligent systems that have been engineered to handle logic-based data and to reason in accordance with inference schemata that have since Aristotle been devised for processing such data.

Acknowledgments

We are grateful to two anonymous reviewers for their helpful feedback, to Paul Bello for insights regarding whether the search for a standard model of the (human) mind can bear fruit in light of the computationally resistant nature of consciousness, and to ONR (Award # N00014-22-1-2201) for partial support of the research reported herein.

References

1. Anderson, J.: Chapter 7: Formal Semantics of ACT Represetations. In: Language, Memory, and Thought, pp. 220–251. Lawrence Erlbaum Associates, Hillsdale, NJ (1976)
2. Andréka, H., Madarász, J.X., Németi, I., Székely, G.: A Logic Road from Special Relativity to General Relativity. Synthese pp. 1–17 (2011). https://doi.org/10.1007/s11229-011-9914-8, http://dx.doi.org/10.1007/s11229-011-9914-8
3. Arkoudas, K.: ChatGPT is no Stochastic Parrot. But it Also Claims That 1 is Greater than 1. Philosophy and Technology (forthcoming)
4. Bello, P., Bringsjord, S.: Two Problems Afflicting the Search for a Standard Model of the Mind. In: The 2017 Fall Symposium Series, Technical Reports FSS-17-01–FSS-17-05, pp. 296–301. Association for the Advancement of Artificial Intelligence, Palo Alto, CA (2017), http://kryten.mm.rpi.edu/pb_sb_aaaifs2017_final.pdf, This paper specifically appears in Technical Report *A Standard Model of the Mind*, Laird, J., Lebiere, C. & Rosenbloom, P., eds. A preprint of the paper can be obtained via the URL given here.
5. Boolos, G.S., Burgess, J.P., Jeffrey, R.C.: Computability and Logic (Fourth Edition). Cambridge University Press, Cambridge, UK (2003)
6. Bringsjord, S., Hendler, J., Govindarajulu, N., Ghosh, R., Giancola, M.: The (Uncomputable!) Meaning of Ethically Charged Natural Language, for Robots, and Us, From Hypergraphical Inferential Semantics. In: Ferreira, I. (ed.) Trustworthy Artifical-Intelligent Systems, pp. 143–167. Springer, Cham, Switzerland (2022), http://kryten.mm.rpi.edu/UncomputableNLURobots032421.pdf, This book is Vol. 102 in the Series *Intelligent Systems, Control and Automation: Science and Engineering.*
7. Bringsjord, S., Govindarajulu, N., Giancola, M.: Automated Argument Adjudication to Solve Ethical Problems in Multi-Agent Environments. Paladyn, Journal of Behavioral Robotics **12**, 310–335 (2021), http://kryten.mm.rpi.edu/AutomatedArgumentAdjudicationPaladyn071421.pdf
8. Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y.T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M.T., Zhang, Y.: Sparks of Artificial General Intelligence: Early experiments with GPT-4 (2023), arXiv:2303.12712v3
9. Ebbinghaus, H.D., Flum, J., Thomas, W.: Mathematical Logic (second edition). Springer-Verlag, New York, NY (1994)
10. Gall, D., Frühwirth, T.: A Formal Semantics for the Cognitive Architecture ACT-R. In: Proietti, M., Seki, H. (eds.) Logic-Based Program Synthesis and Transformation: (Lecture Notes in Computer Science Volume 8981), pp. 74–91. Springer, Cham, Switzerland (2014), http://www.informatik.uni-ulm.de/pm/fileadmin/pm/home/fruehwirth/drafts/act_r_semantics.pdf

11. Goldfain, A.: A Computational Theory Of Early Mathematical Cognition. Ph.D. thesis, State University of New York at Buffalo (2008)
12. Halpern, J., Harper, R., Immerman, N., Kolaitis, P., Vardi, M., Vianu, V.: On the Unusual Effectiveness of Logic in Computer Science. The Bulletin of Symbolic Logic **7**(2), 213–236 (2001)
13. Hutter, M.: Universal Artificial Intelligence: Sequential Decisions Based on Algorithmic Probability. Springer, New York, NY (2005)
14. Kissinger, H., Schmidt, E., Huttenlocher, D.: ChatGPT Heralds an Intellectual Revolution. Wall Street Journal (February 24 2023), The rather lengthy subtitle of this article reads: "Generative AI presents a philosophical and practical challenge on a scale not experienced since the start of the Enlightenment."
15. Kotseruba, I., Tsotsos, J.K.: A Review of 40 Years of Cognitive Architecture Research: Core Cognitive Abilities and Practical Applications. arXiv: Artificial Intelligence (2016)
16. Laird, J., Lebiere, C., Rosenbloom, P.: A Standard Model of the Mind: Toward a Common Computational Framework Across Artificial Intelligence, Cognitive Science, Neuroscience, and Robotics. AI Magazine **38**(4), 13–26 (2017). https://doi.org/https://doi.org/10.1609/aimag.v38i4.2744
17. Lakoff, G., Nuñez, R.: Where Mathematics Comes From: How The Embodied Mind Brings Mathematics Into Being. Basic Books, New York, NY (2000)
18. Lovett, A., Bridewell, W., Bello, P.: Selection, Engagement, & Enhancement: A Framework for Modeling Visual Attention. In: Proceedings of the 43rd Annual Conference of the Cognitive Science Society. pp. 1893–1899. Cognitive Science Society, Vienna, Austria (2021)
19. McKinsey, J., Sugar, A., Suppes, P.: Axiomatic Foundations of Classical Particle Mechanics. Journal of Rational Mechanics and Analysis **2**, 253–272 (1953)
20. Patel-Schneider, P.F.: Undecidability of subsumption in NIKL. Artificial Intelligence **39**(2), 263–272 (1989)
21. Schmidt-Schauß, M.: Subsumption in KL-ONE is undecidable. KR **89**, 421–431 (1989)
22. Shapiro, S.: SNePS: A logic for natural language understanding and commonsense reasoning. In: Iwanska, L., Shapiro, S. (eds.) Natural Language Processing and Knowledge Representation: Language for Knowledge and Knowledge for Language, pp. 175–195. AAAI Press/MIT Press, Menlo Park, CA (2000)
23. Shapiro, S.C., Bona, J.P.: The GLAIR Cognitive Architecture. In: Biologically Inspired Cognitive Architectures II: Papers from the AAAI Fall Symposium (2009)
24. Simpson, S.: Subsystems of Second Order Arithmetic. Cambridge University Press, Cambridge, UK (2010), this is the 2nd edition.
25. Wang, P.: Non-Axiomatic Logic. WORLD SCIENTIFIC (2013). https://doi.org/10.1142/8665, https://www.worldscientific.com/doi/abs/10.1142/8665
26. Wang, P.: Perception in NARS. Tech. rep. (2018), https://cis.temple.edu/tagit/publications/PAGI-TR-7.pdf
27. Wigner, E.: The Unreasonable Effectiveness of Mathematics in the Natural Sciences. In: Communications in Pure and Applied Mathematics, vol. 13, pp. 1–14. John Wiley and Sons, New York, NY (1960)
28. Wolfram, S.: What Is ChatGPT Doing ... and Why Does It Work? Stephen Wolfram – *Writings* (February 2023), https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work, Specifically published on February 14 2023.