https://doi.org/10.1142/9789811276675_0005

Chapter 5

Universal Cognitive Intelligence, from Cognitive Consciousness, and Lambda (Λ)

Selmer Bringsjord, Naveen Sundar Govindarajulu, James Oswald

Rensselaer AI & Reasoning (RAIR) Lab,
RPI, Troy NY 12180, USA

We explain that the concept of *universal cognitive intelligence* ($\mathscr{UCI}$) can be derived in part by generalization from the previously introduced (and axiomatized) *theory of cognitive consciousness*, and the framework, Λ, for measuring the degree of such consciousness in an agent at a given time. $\mathscr{UCI}$ (i) covers intelligence that is artificial or natural (or a hybrid thereof) in nature, and intelligence that is not merely Turing-level or less, but also beyond this level; (ii) reflects a psychometric orientation to AI; (iii) withstands a series of objections (including e.g. the opposing position of David Gamez on tests, intelligence, and consciousness, and the complaint that so-called "emotional intelligence" is beyond the reach of any logic-based framework, including thus $\mathscr{UCI}$); and (iv) connects smoothly and symbiotically with important formal hierarchies (e.g., the Polynomial, Arithmetic, and Analytic Hierarchies), while at the same yielding its own new all-encompassing hierarchy of logic machines: $\mathfrak{LM}$. We end with an admission: $\mathscr{UCI}$ by our lights, for reasons previously published, cannot take account of any form of intelligence that genuinely exploits *phenomenal* consciousness.

Contents

List of Figures

1. Introduction; Plan for the Paper

We show herein that the concept of *universal cognitive intelligence* ($\mathscr{UCI}$) can be easily derived in significant part by generalization from the previously introduced and axiomatized *theory of cognitive consciousness* (TCC), and the framework, Λ, for measuring the degree of such consciousness in an agent at a given time (or over an interval of times). Λ (of the finitary variety), in conjunction with TCC and its axiomatization in the system $\mathcal{CA}$, is presented in (Bringsjord & Govindarajulu 2020). TCC was explained before that publication, more informally and less computationally, in (Bringsjord, Bello & Govindarajulu 2018). Herein, we do not simply reprise this earlier material. Instead, we move from it to $\mathscr{UCI}$, which applies to all cognitive agents, period. In the case of Λ, we explain efficiently how the new, *in*finitary version of it is very easy to erect.

But what, it will be immediately asked, is cognitive intelligence? We are quite confident that you have considerable cognitive intelligence, and to begin straightaway to characterize $\mathscr{UCI}$ we take a few minutes here in the Introduction to justify our confidence. Our confidence arises in no small part from the fact that performance on certain tests can confirm our attitude, so to a test we now turn.

Imagine that Alice, a math teacher in whose class you are in, asks you to consider the proposition that it's not the case that some proposition ϕ implies another proposition ψ. (Alice is currently teaching basic deductive reasoning in the logic known as the *propositional calculus*, or for us $\mathscr{L}_{PC}$.) Alice's two-part Test 1 for you is that, given this negation (i.e. that $\neg(\phi \rightarrow \psi)$), ...

Test 1

> ... does it follow deductively that ϕ holds? And part two: Prove that your answer is correct.

We encourage you to take a minute to reflect.

The answer is "Yes," and a valid (and, for exposition here, verbose) proof is easy enough to come by.[1] We assume that you can handle any test question like this at the level of $\mathscr{L}_{PC}$. That is, we assume that you are an agent (and we further assume in particular that you are a *natural*, not an *artificial*, agent) who can decide, for a set Φ of formulae of $\mathscr{L}_{PC}$ whether an arbitrarily given formula ψ also of this logic can be proved from Φ. Where the customary provability relation $\vdash$ is used, and we subscript it with the inference schemata for the particular logic in question, this is the question

[1] E.g., this will do:

> **Proof**: Our given is that $\neg(\phi \rightarrow \psi)$. Suppose for indirect proof that $\neg\phi$ holds. Suppose in addition, for a sub-proof using *reductio* as well, that $\neg\psi$; and suppose too — to obtain a conditional directly inconsistent with our given — that ϕ. We deduce ψ by *reductio* from the now-available contradiction, and then since our supposition of ϕ has led to ψ, we deduce the conditional $\phi \rightarrow \psi$. But now we have before us the contradiction of this conditional with our given, and infer by *reductio* on (1) that ϕ. ■.

Note that only such a proof will do, if we care about scalability, and given the nature of $\mathscr{UCI}$ we very much do. We can't e.g. use a truth table, because such things don't scale up to first-order logic = $\mathscr{L}_1$ and beyond.

(type[2]) we assume you can handle:

$$(\dagger) \quad \Phi \vdash_{\mathscr{L}_{PC}} \psi?$$

It may appear to the reader that so far we have said precious little that links cognitive consciousness with intelligence. A skeptic may specifically say Test 1 plausibly relates to cognitive intelligence, but appears to have nothing to do with consciousness. Many readers with little serious exposure to philosophy or psychology in our experience take it for granted that there is a deep nexus between consciousness and intelligence; but those skeptical about our project may claim outright that intelligent behavior in the complete absence of conscious content is entirely possible; and this claim, if left standing, threatens the very foundation of the $\mathscr{UCI}$ conception and framework.

Fortunately (for us, anyway), the claim is already undermined by Test 1; you can see this by returning to this test now, for closer study. Test 1, recall, was presented to you by teacher Alice, whose class, you believe, is intended by her to bring it about in you via your attention-and-perception abilities that you know the answers to instantiations of (†). Let the instantiation given above be denoted by ($\dagger^*$). Now, by definition, your level of cognitive consciousness in successfully negotiating Test 1 is present, and indeed quite significant. For, if you are denoted by '$\mathfrak{a}_y$,' and Alice by '$\mathfrak{a}_a$,' and we use for each cognitive verb clearly operative in this simple scenario, we can encapsulate the situation by way of the immediately following formula, whose import is likely quite clear to the reader, even without an accompanying tutorial regarding our highly expressive cognitive calculi (to which we return below):

$$(+) \qquad \mathbf{B}(\mathfrak{a}_y, \mathbf{I}(\mathfrak{a}_a, \mathbf{K}(\mathfrak{a}_y, (\dagger^*)))).$$

In essence, this formula says that the test-taking agent (the role played by you, the reader), believes that Alice intends that you know the instantiation of (†) holds. This belief on your part intuitively constitutes are significant state of cognitive consciousness. Hence, we see that, albeit quickly and at an

[2]The question type here is the "grandfather" of all **NP**-complete problems (though usually the question is put in terms of satisfiability). It thus follows immediately that $\mathscr{UCI}$ includes early on a level of cognitive intelligence in the Polynomial Hierarchy (PH) considered by some to be impressive. For us, as stated later, cognitive intelligence at impressive levels is *beyond* this hierarchy, which accordingly we don't bother to include a view of herein. Readers wishing to initiate study of the Polynomial Hierarchy could turn to (Arora & Barak 2009) for a purely computer-science, and for a truly masterful essay that positions computational complexity/PH within deep issues in mathematics, we heartily recommend (Dean 2019).

intuitive level, that there is in fact an intimate connection between cognitive intelligence, manifested by success on Test 1, and cognitive consciousness.

But what good, you may ask, is the cognitive intelligence we have ascribed to you? Well, if you have this level of cognitive intelligence, you can then use it to find the answer to any problem whatsoever that can be reduced to a query $\mathfrak{q}$ issued against information Φ (where of course the query is a formula in the propositional calculus). A computing machine with this power, that is to say an *artificial* agent with this level of cognitive intelligence, which consists in automated reasoning, is not a small thing. Historically speaking, the first great success stories for logic-based (logicist) AI, from Simon and Newell, were systems, Logic Theorist and General Problem Solver (GPS), with only a significant portion of this cognitive power; see (Newell & Simon 1956).

Another way to contextualize the level of cognitive intelligence we have introduced via Test 1 is to consider pure "logic machines" and logic programming. The state of research and development of logic machines at the dawn of AI is chronicled in the important (but often overlooked) book *Logic Machines and Diagrams* by Gardner (1958). Put in "Prolog-ish" logic-programming terms, the level of cognitive intelligence we have identified above roughly corresponds to what a Proplog (note the spelling) program can do. Proplog is a proper subset of Prolog in which predicates are strings representing propositions, and terms are absent. Of course, to build artificial agents able to succeed in many areas of AI (albeit to a humble degree) it suffices to create a suitable Proplog program; this can be seen by turning to how the propositional calculus, conjoined with an automated reasoner for it, is shown in action in the AI textbooks of today (see e.g. Russell & Norvig 2020).

Of course, as you and all those working in (or even just seriously interested in) AI will know, the level of cognitive intelligence identified thus far is quite humble. $\mathscr{UCI}$, as is soon seen, scales progressively higher and higher in the levels of cognitive intelligence available, thereby covering agents of increasing cognitive intelligence, *ad infinitum.*

The remainder of the present essay realizes a straightforward plan: We first recapitulate, in brief, TCC, its axiomatization in $\mathcal{CA}$, and (the finitary version of) Λ (§2). We then explain in the context of prior work why a move to universal *cognitive* intelligence is to our minds wise (§3); quickly root $\mathscr{UCI}$ in the psychometric approach to AI introduced and defended by the first author rather long ago (§4); briefly explain how $\mathscr{UCI}$ is naturally associated with Λ of the infinitary form, which is based upon transfinite numbers and infinite matrices (§5); explain that $\mathscr{UCI}$ is connected to the Arithmetic and Analytic Hierarchies, and to the first author's new hierarchy $\mathfrak{LM}$ that subsumes these and indeed all standard, established hierarchies (§6); briefly discuss some work by others that relates to $\mathscr{UCI}$ (§7); and present and respond in Sec. 8 to a series of objections to $\mathscr{UCI}$ (including the position of Gamez that consciousness and a test-based approach to intelligence are not compatible, let alone synergistic (§8.3); and the claim that $\mathscr{UCI}$ can't cover forms of intelligence (e.g., so-called "emotional intelligence"; §8.4). We wrap up by conceding that non-occurrent mental states are untouched by our framework, and that another form of consciousness, so-called *phenomenal* consciousness, is beyond the reach of $\mathscr{UCI}$ and its ingredients, such as Λ.

2. Cognitive Consciousness, Its Axiomatization ($\mathcal{CA}$), and Λ: Recapitulation

In the present section we briefly recapitulate first TCC, and then Λ.

2.1. *The Theory of Cognitive Consciousness*

As at least philosophers working on consciousness well know, Block (1995) introduced a fundamental dichotomy of types of consciousness: *phenomenal* consciousness (p-consciousness) on the one hand, and *access* consciousness (a-consciousness, abbreviated) on the other. P-consciousness, as most of our readers will know, is "what it's like" consciousness. There is something it feels like to taste, analyze, and enjoy a great bottle of Nerello Mascalese. There is also something it feels like to skydive for the first time. And so on. When you enter these states, you are p-conscious. At least according to the first author, p-consciousness is not only impossible to capture in computation (Bringsjord 1999) — it's not even possible to rationally take a first genuine engineering step toward building an artificial agent that is p-conscious (Bringsjord 2007). We return to p-consciousness at the very end of the present paper.

What about a-consciousness? What is it? We would very much like to provide you with a formal definition, but the concept is, as we've said, due to Block [certainly with antecedents he (1995) discusses], and by any relevant standard he leaves this concept informal.[3] For confirmation to the reader, we convey that Block's (1995) informal definition (p. 231) is as follows: A state of some agent is a-conscious if and only if it is poised (a) to be used as a premise in reasoning, (b) for rational control of action, and (c) for rational control of speech.[4]

Now, a third kind of consciousness is the one near and dear to our hearts: *cognitive consciousness*, or just *c-consciousness*. This brand of consciousness is present only when the agent that bears it has a robust ensemble of cognitive attitudes, which correspond directly to a relevant set of verbs that signal parts of cognition long investigated in cognitive psychology and cognitive science (e.g. see the cognitive verbs that anchor a number of the chapters in the authoritative Ashcraft & Radvansky 2013). The set of these verbs includes: *believing*, *knowing*, *perceiving*, *communicating* (in a natural language, and perhaps also a formal language that might be used in, say, mathematics), *hoping*, *fearing*, *intending*, and so on *ad indefinitum*.[5] For the most part, c-conscious states can be denoted by use of gerundive nominals and specifically the schema '$\mathfrak{a}$'s V-ing that ϕ,' where, respectively, these variables take agents, cognitive verbs, and declarative propositions. For us, not only must ϕ be a formula in the formal language $\mathcal{L}$ of some formal logic $\mathscr{L}$, but the schema itself must correspond to some formula in some formal language $\mathscr{L}'$ of some *intensional* logic $\mathscr{L}$. Intensional logics allow for the representation and reasoning over propositions whose meanings are not

[3] This is confirmed by the fact that we know now what is meant by 'reasoning' here. For that matter, what is "rational" control of action? In fact, what is rationality? Because of the obscurity of Block's definition, the first author long ago issued a recommendation to discard the term 'a-consciousness' in favor of using instead terms that refer to the kinds of things this umbrella term is supposed to cover (Bringsjord 1997).

[4] Oddly, Block admits (p. 231) that condition (c) isn't necessary, since — as he sees matters — non-linguistic creatures can be a-conscious by virtue of their states satisfying only (a) and (b). This admission seems to us to be indicative of just how murky a-consciousness is — so murky for us that we refrain from addressing such questions as: Is c-conscious content coextensive with Block's admittedly ill-posed definition of a-conscious content? Does a-consciousness play a large role e.g. in certain conceptions of intentional action, keeping in mind that Λ takes account of any intensional operator for 'intends'? We believe that if a-consciousness was defined using the tools and techniques of logic-based AI or CogSci (e.g. see Bringsjord, Giancola & Govindarajulu (forthcoming), Bringsjord 2008), it would be possible to venture formal definitions of c-consciousness. But such an investigation is out of scope for us in the present essay.

[5] As far as we can tell, any agent or system that is cognitively conscious (= i.e. that enters into a series of c-conscious states through an interval of time) is necessarily a-conscious during this stretch. In general, we see no harm in viewing cognitive consciousness to be the most important type of a-consciousness identified by human scientists and engineers thus far.

determined compositionally. For instance, if an agent $\mathfrak{a}$ believes that Selmer is short, and let this be represented by the formula $\mathbf{B}(\mathfrak{a}, S(s))$ (where $\mathbf{B}$ is an intensional operator), whether or not Selmer is in fact short will have no bearing on the truth of this formula. In extensional logics, which include all the elementary classical logics students of logic and mathematics learn when they first start out, for instance the logics $\mathscr{L}_{PC}$ and $\mathscr{L}_1$, if one knows the semantic value of constituents of formulae one can calculate the value of the overarching formula. For instance, if $S(s)$, as an atomic formula in $\mathscr{L}_1$ in which S is a unary relation and s a constant, is true, then we know immediately that the disjunction $S(s) \vee \psi$ is true as well, for any instantiation to ψ.[6] We have introduced an infinite family of logics, *cognitive calculi*, that include intensional operators for all significant cognitive attitudes. For a characterization of cognitive calculi, readers are directed to Appendix A in (Bringsjord, Govindarajulu, Licato & Giancola 2020). For presentation and use of a cognitive calculus that we have made considerable use of in our AI work, the *Deontic Cognitive Event Calculus* ($\mathcal{DCEC}^*$), see (Govindarajulu & Bringsjord 2017, Bringsjord, Govindarajulu & Giancola 2021).

At this point, we recommend that the reader see Fig. 1, which depicts some key aspects of the discussion thus far. The logics along all three "rays" are needed by $\mathscr{UCI}$, and by the new hierarchy $\mathfrak{LM}$ for it. (For those readers wishing to look ahead figure-wise, this new hierarchy is referred to in Figs. 5 and 7, and is generated by the conception of logic programming summarized in Fig. 6.)

2.1.1. *Regarding the Axiom System ($\mathcal{CA}$) for Cognitive Consciousness*

We refrain from covering in the present paper all the axioms of TCC, since our main purpose is to introduce $\mathscr{UCI}$. For full coverage, the reader interested in more is directed first to the introduction of the axioms (and cognitive consciousness in general) provided in (Bringsjord et al. 2018), and then, for a more detailed (and more technical) presentation of the axioms geared to those in AI, to (Bringsjord & Govindarajulu 2020). It will suffice here to show the reader but two of the simplest axioms of the axiom system $\mathcal{CA}$, and here's the first:

[6]For a fuller discussion, but still an economical one, we direct the reader to (Fitting 2015).

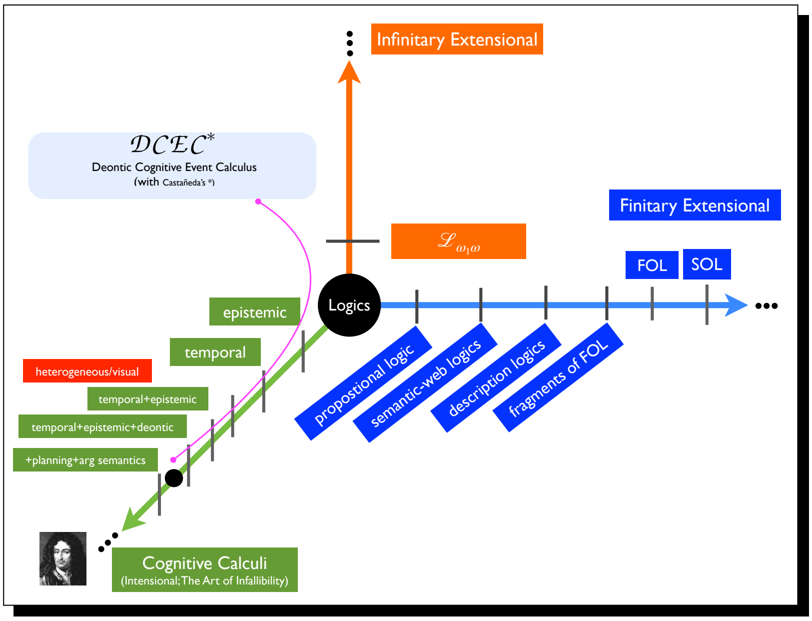

Fig. 1. The Three "Rays" of Logics. *The blue ray (x axis) emanates out from the start/core, passing through ever-increasingly expressive finitary extensional logics. The orange ray (y-axis) goes upwards, passing through increasingly expressive* ***infinitary*** *extensional logics. The green ray is a bit complicated, with details out of scope here, but the basic idea, which has led Bringsjord to proclaim his discovery of Leibniz's "The Art of Infallibility" (in the French used by Leibniz: art d'infaillibilité) is that content from the other rays are combined with full coverage of intensional attitudes in a certain family of intentional logics (called* ***cognitive calculi****, with diagrammatic/visual logics included as well).*

Perception to Belief

P2B Human persons perceive internally[a] and externally,[b] and in both cases the percepts in question are believed (at varying degrees of strength, with external perception at the strength of *evident* (which here can be understood as "overwhelmingly likely"), but never *certain*) by these agents, whereas most of what is internally perceived is indeed certain.

[a]E.g., we perceive that we are in pain when we are.
[b]E.g., we perceive creatures whose behavior indicates to us that they are in pain.

P2B is easy to comprehend (even without our presenting here the likelihood calculus and inductive intensional logics that use this calculus).[7] When we perceive such things as that seven is a prime number (via purely mental activity) or that we seem to be sad, we believe these propositions, and they are *certain* for us. But when we perceive in a garden a pink rose, *ceteris paribus* we believe that there is a pink rose before us, but it could be an illusion. (We may have forgotten that we are wearing rose-tinted sunglasses, and we are in fact looking at a white rose.) In c-consciousness as we rigorize it, belief is stratified by its strength (or confidence), in that a belief is accompanied by a *strength factor* σ. The strength of the cognitive attitude of belief is based directly on the underlying *likelihood* of the proposition. So for example Jones, if having ingested a powerful pain reliever in a hospital, and knowing that such drugs can have serious side-effects, may believe only at the level of *more probable than not* that there is a walrus before him. With strength stratification in place, belief becomes graded on the positive side from *certain* to *certainly false*, and so will knowledge.[8] (The negative part of the spectrum is symmetrical with the positive. E.g., "more likely than not" = σ of 1 has its inverse "less likely than so.") This means that our framework for $\mathcal{CA}$, in contrast with axiom systems based on standard logics (e.g. Peano Arithmetic and axioms for set theory, such as the Zermelo-Fraenkel axioms; these two axiom systems are expressed solely in $\mathscr{L}_1$, which is firmly bivalent), which have binary values TRUE and FALSE, or sometimes those two plus INDETERMINATE, has 13 possible values. In large measure due to the research and engineering of the second author, and significant contributions from Mike Giancola, we have some fairly robust implementations of artificial agents that embody axiom **P2B**, and bring this framework to concrete life; see the simulation, and CPU times, reported in (Bringsjord et al. 2021), and see as well footnote 7, wherein we explain

[7]This is a fitting place to explicitly inform the reader that in the present paper we almost exclusively refer to and use *deductive* logics/calculi. We do this to keep things manageable in the span of a single, reasonably sized essay. Were we to enlarge all that we say, all of our figures, and so on, so as to take account of formal inductive logics, the space of a small book would be needed. The best place for diligent readers to turn should they wish to explore how $\mathscr{UCI}$ can be expanded so as to include coverage of the inductive-logic case is (Bringsjord et al. 2021), in which the inductive cognitive calculus $\mathcal{IDCEC}$ is specified and used in a robust simulation with an automated reasoner (ShadowAdjudicator) for it and other such inductive calculi. To our knowledge, this is the only robust automated reasoner for inductive logics at the time the present sentence is being written.

[8]This is what allows us to solve the so-called Gettier Problem. See (Bringsjord et al. 2020).

that the present paper must for economy remain focused, throughout, on deductive logics/reasoning.

Here now is the second axiom from $\mathcal{CA}$ we choose to present:

Introspection (positive)
Intro Humans persons know that they know what they know.

At least the tenor of this axiom is well-known in formal intensional/modal logic because it corresponds to a much-discussed axiom from so-called *alethic* modal logic — an axiom customarily written

$$\Box\,\phi \;\rightarrow\; \Box\Box\phi,$$

when symbolized as the characteristic axiom of the modal logic **S4**, a logic going back to C.I. Lewis. In **S4**, the boxes here are read as 'it's necessary that.' In epistemic logic, we instead read $\Box$ as 'knows that,' denoted by simply '**K**' in our cognitive calculi and in fact in all the intensional logics on the green ray in Fig. 1. A bound $k \in \mathbb{N}$ can be placed on the iteration of **K**, but it would, we think, need to be at least 5 for human-level cognition (for a rationale, see Bringsjord & Ferrucci 2000). The version of the axiom given immediately above has a level of $k = 3$. The axiom here can also be expanded to include provision for negative introspection (i.e., $\neg\mathbf{K}\phi \rightarrow \mathbf{K}\neg\mathbf{K}\phi$), and once again a bound can be placed on the iteration, if desired. Finally, since as we explained in connection with **P2B**, both belief and knowledge have varying strength (or, again, confidence), we have inserted the superscript 'σ' as a variable in the otherwise natural-language version of the axiom **Intro**. In the case of $\mathscr{UCI}$ (and in fact AGI as well; see §7.1), the type of agent a given axiom pertains to doesn't at all have to be those of the human-person variety. But if we stick with **Intro** as we have written it above, and allow '$\mathfrak{a}_{hp}$' to range over human-person agents, and use the epistemic operator **K** for 'knows,' then we have in particular:

$$\forall\phi\ [\mathbf{K}^{\sigma\leq 6}(\mathfrak{a}_{hp},\phi) \;\rightarrow\; \mathbf{K}^{6}(\mathfrak{a}_{hp},\mathbf{K}^{6}(\mathfrak{a}_{hp},\mathbf{K}^{\sigma\leq 6}(\mathfrak{a}_{hp},\phi)))].$$

It is very important for the reader to understand the implications of the formula given immediately above. Specifically, what we have here implies that in our formal approach to $\mathscr{UCI}$, we must minimally use an intensional logic that allows not only standard quantification over object variables and relation variables (and this to some order in higher-order logic), but must also include formalization of meta-logical propositions. We see this here in that there is quantification over ϕ, which is a variable ranging over all well-formed formula in the logic in question. For ease of exposition and to keep the present essay reasonably sized, we refrain from a full discussion of the fact that cognitive calculi include not only object-level formal languages and inference schemata that regiment reasoning over formulae in these languages, but also meta-logical languages, with formulae in them connected to dedicated inference schemata. That said, the reader can see that proposition (+) above is a case in point, since intensional operators range over meta-logical expressions.

2.2. Λ: *Measuring Cognitive Consciousness*

Λ measures consciousness, specifically, as the reader now knows, c-consciousness, based on how the agent in question observably cognizes. Rather than striving to measure phenomenal consciousness (p-consciousness), which is what Φ from Tononi (2012) is supposed to do,[9] Λ explicitly explains and accounts for cognitive consciousness (c-consciousness).

An AI-oriented introduction to Λ is provided in (Bringsjord & Govindarajulu 2020). Later in the present essay, we given a brutally rapid encapsulation of how Λ can be extended to the infinite case, which, as the alert reader by now knows, is certainly needed for the higher reaches of $\mathscr{UCI}$. But in the present section we recapitulate finitary Λ; but, relative to Λ defined in (Bringsjord & Govindarajulu 2020), the version of Λ we explain momentarily has been extended, in that now, here, a crucial part of Λ includes the *justification* for the cognition of the agent whose c-consciousness is being measured.

For setting exposition here, assume we have an agent $\mathfrak{a}$ that acts at discrete timepoints. For some of the agent's actions $\alpha(t)$, the agent outputs a justification/rationale $\mathit{justification}(\mathfrak{a}, \alpha, t)$. Λ is based on the richness of structures found in the justifications produced by the agent. The justification can be a semi-formal structure, and can include a mix of different modalities (non-verbal actions, gestures, written content, etc.). If the structures include references to cognitive states of other agents or the agent itself, we in general assign a high Λ score to the agent at those points in time. Unlike the aforementioned Φ, we don't provide a single Λ value for an agent or system or creature which is to be measured with respect to c-consciousness; rather, Λ consists in a sequence or vector values

[9]For a more technical presentation of the Integrated Information Theory of (p-)consciousness, and Φ, see e.g. (Oizumi, Albantakis & Tononi 2014).

corresponding to the different cognitive verbs discussed above as key, such as *knows* **K**, *believes* **B**, *desires* **D**, *intends* **I**, *communicates* or *says* **S**, temporal structures $\vec{t}$, quantifiers $\forall, \exists, \ldots$, etc. Semi-formally, if we have justification $\mathit{justification}(\mathfrak{a}, \alpha, t)$ produced by an agent $\mathfrak{a}$ for action α at time t, then (see also Fig. 2 for pictorial exposition):

$$\Lambda\big[\mathit{justification}(a, \alpha, t)\big] = \langle \lambda_{\mathbf{B}}, \lambda_{\mathbf{D}}, \lambda_{\mathbf{I}}, \lambda_{\mathbf{K}}, \lambda_{\vec{t}}, \lambda_{\forall}, \lambda_{\exists} \ldots \rangle$$

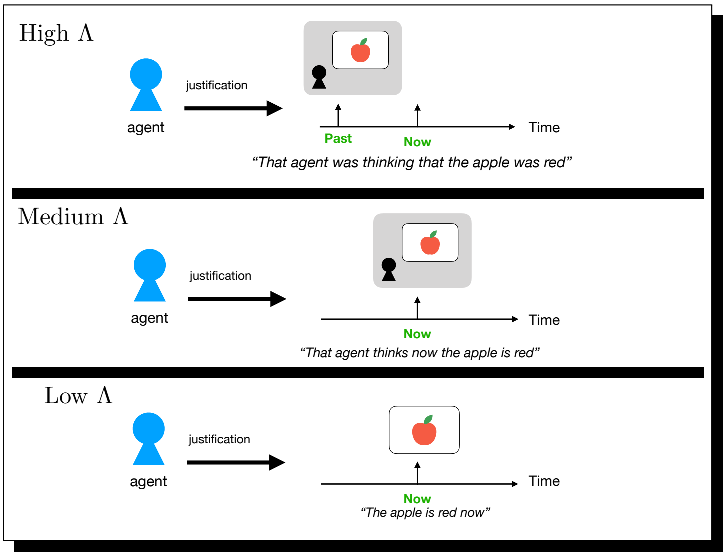

Fig. 2. Λ and Justifications. *We have higher Λ values when the agent has to consider other agents and handle richer temporal structures.*

We end this section with a point about non-modal Λ complexity in formulae and justifications: Because we are herein presupposing the paper in which Λ was introduced (viz. Bringsjord & Govindarajulu 2020), and not fully recapitulating the ins and outs of Λ herein, it may not be evident to some readers of the present essay that differences in the formulae under the scope of modal operators are important determinants with respect to Λ values of formulae. For an example in the present essay, consider the Π_2 formula (2) given later in the paper, quantification in which is somewhat robust. Now consider a simple formula in the propositional calculus, specifically the conditional (denote it by 'c') featured in Test 1, above. Next, we can without loss of generality assume that the reader $\mathfrak{a}_r$ perceives both of these formulae. This means (for us) that both formulae are within the scope of a modal operator $\mathbf{P}$, and there are no other modal operators in play in either case. Yet, clearly Λ applied to $\mathbf{P}(\mathfrak{a}_r(2))$ is significantly greater than Λ applied to $\mathbf{P}(\mathfrak{a}_r c)$. We shall also see in §5.3 that purely extensional elements of formulae inside the scope of modal operators are used to calculate Λ.

3. Why Universal *Cognitive* Intelligence?

An intelligent agent $\mathfrak{a}$ can be artificial. But $\mathfrak{a}$ can also be natural, clearly. After all, since you are currently reading and understanding this sentence (at least up to the right parenthesis that concludes this very remark), you are intelligent — yet you may well not be artificial.[10] You might for instance be a human agent. Therefore, a theory of "universal artificial intelligence" (e.g. that from Hutter 2005) would fail to cover your intelligence (UAI is discussed below, in §7.2). Since we desire to erect a fully comprehensive formal account of cognitive intelligence in agents, whether artificial or natural, finite or infinite, a theory of intelligence that covers only the former is inadequate.[11]

Perhaps, then, we one should seek not a theory of universal *artificial* intelligence, but rather a theory of *computational* intelligence, where 'computational' here is the standard adjective used to cover the processing of information at the level (and below) of a Turing machine and its equivalents.[12] Despite the empirical facts that (i) currently the majority of those working in AI and CogSci restrict their attention to agents capable only of information processing that is indeed merely Turing-level and below (e.g. see the mainstream textbooks, such as Russell & Norvig 2020), and that

[10] If God exists, and you are a human agent created by him, then perhaps it would not be inaccurate to say that from his perspective you are an artifact, and hence artificial. But let us leave this possibility aside, and affirm the customary and convenient terrestrial distinction between natural versus artificial agents.

[11] Hutter's (2005) account seems to us inadequate for other reasons. E.g., it fails to insist that no artificial agent can be genuinely intelligent unless it has a large, perhaps even an infinite, amount of declarative knowledge. At a minimum, a cognitively intelligent cognizer must know that it exists, that it has some cognitive states of the formal shape that we have indicated above (e.g., *knowing that it exists*, where here we have instantiation of the schema given above, viz., $\mathfrak{a}$'s V-ing that ϕ, where ϕ is formulae in some formal language for some formal logic/cognitive calculus, and V-ing is cognitive verb in the gerundive-nominal case), etc. As we shall shortly see, intelligent cognizers with high cognitive intelligence have a lot of such knowledge even in the simple case of basic arithmetic. We return to Hutter and his "Universal AI" in §7.2.

[12] Register machines, the λ-calculus, etc. at the level of a Turing machine, and e.g. finite-state automata below this level.

(ii) many who occupy themselves with looking at human brains conceptualize intelligence as constituted and bounded by Turing-level-and-below computation (e.g. Rodriguez & Granger 2016), setting the goal of erecting a comprehensive theory of computational intelligence, with this limited sense of 'computational' affirmed, is an exceedingly bad idea. The reason is obvious: such a theory would fail to cover agents whose information-processing power reaches beyond standard Turing-level machines. For example, an agent able to solve problems that call for *infinite-time* Turing machines (Hamkins & Lewis 2000) would no doubt be rather intelligent, but such an agent stands to a theory of computational intelligence as quantum effects stand to Newtonian mechanics. (As the reader will soon see, $\mathscr{UCI}$ is conceived and designed to be informed by some hierarchies which are mostly about information processing *above* what a Turing machine can do; see, in a look ahead to later in the present paper, Fig. 5. There, both the Arithmetic and Analytic Hierarchies, long established in formal logic/mathematics and theoretical computer science since Turing's dissertation work under Church in the U.S., are mostly super-Turing; and the same holds for the new $\mathfrak{LM}$ hierarchy (see Fig. 7).)

What, then, is to be done? We can erect a formal theory of universal *cognitive* intelligence, $\mathscr{UCI}$, one that can cover all cognizers, whether artificial or natural or artificial-natural hybrids, and whether capable of processing information above the reach of standard Turing machines, or not. And how is $\mathscr{UCI}$ to be erected? Well, many steps need to be taken; and many of the steps are (in part) taken herein for the reader to see. But one key step for us was to observe that Λ, which gives a measure of the level of cognitive consciousness of a given agent $\mathfrak{a}$ at any time t, can easily enough be extended to the infinite case. We efficiently explain this extension below, in Sec. 5.3, but first we explain in broad terms another key step on the road to $\mathscr{UCI}$, which is to note that so-called *Psychometric AI* can be extended to allow tests of cognitive power that exceed the Turing Limit on multiple fronts. We explain this step next, in brief.

4. Psychometric AI to Universal Psychometrics

We have earlier introduced a form of AI that is based on the measurement of cognitive ability possessed by given agents. This form of AI is known as, again, *Psychometric AI* (PAI), which was first introduced in (Bringsjord & Schimanski 2003), with subsequent treatment and expansion for example in (Chapin, Szymanski, Bringsjord & Schimanski 2011, Bringsjord 2011,

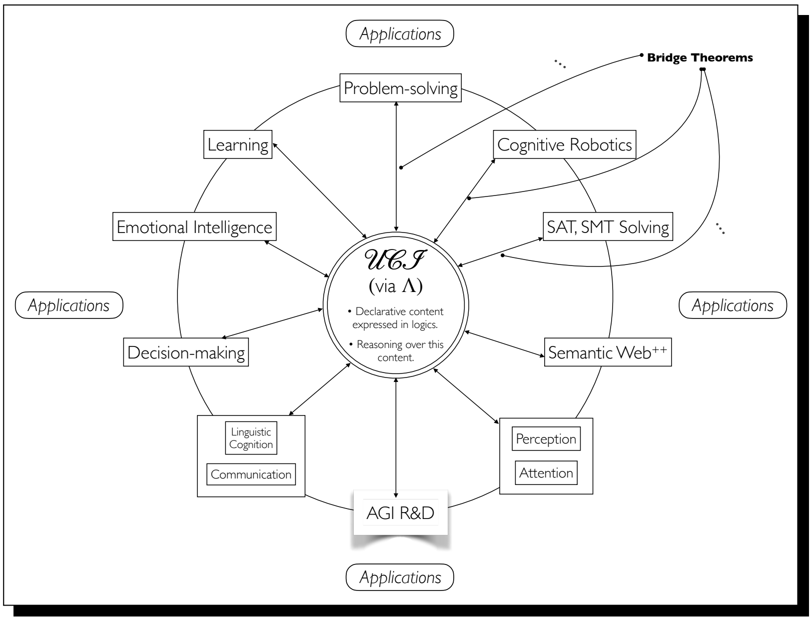

Fig. 3. Universal Cognitive Intelligence as the Core. *All the areas of AI, AGI, and computational CogSci can be reduced to the logicist formalisms and automated reasoners posited by* $\mathscr{UCI}$*, via bridge theorems.*

Bringsjord & Licato 2012, Klenk, Forbus, Tomai & Kim 2011). PAI holds that AI is, or at least ought to be, the field devoted to the conception, design, and implementation of artificial agents able to excel on tests of cognitive ability and skill. Here, we re-affirm the prior commitment to this doctrine — but extend this orientation to agents that are not artificial, not at present terrestrial, not finitary, and not merely Turing-level. For instance, alien creatures that are natural might have exceedingly high levels of cognitive intelligence in connection with arithmetic, as evidenced by their performance on tests of arithmetic. Crucial for understanding this possibility is Fig. 4, which the reader should now consult.

This figure restricts PAI and its tests to be exclusively about arithmetic, but extended to the infinite and super-Turing. All test questions must be restricted to arithmetic involving only the simple arithmetic functions over $\mathbb{N}$: addition, multiplication, subtraction, with relations for $>, <, \geq, \leq$. To put the matter in science-fiction terms, we believe that if alien agents from another planet show up on Earth in a spaceship, they will as a matter of cognitive necessity have had to pass from the innermost circle shown in

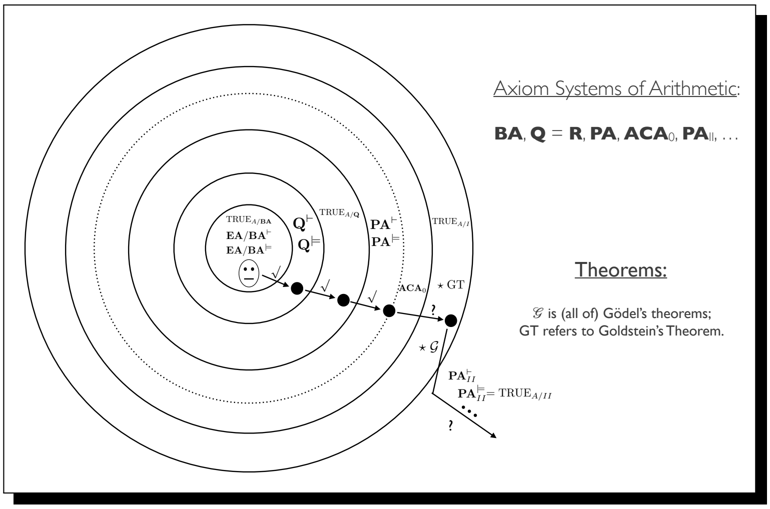

Fig. 4. The Reverse-Mathematics Basis for Universal Psychometrics (Arithmetic Case). *We assume that all agents with significant cognitive intelligence must "travel" from the innermost circle here outwards, as they master more and more arithmetic. A number of axiom systems for arithmetic are referred to here.* ***PA*** *is Peano Arithmetic; this will be familiar to all readers well-versed in formal logic and/or theoretical computer science. Closer to the center of the circles brings one to less powerful systems of arithmetic (for instance "Baby Arithmetic" (**BA**), nicely covered in (Smith 2013)), and "Robinson Arithmetic" (**Q**), nicely covered in (Boolos et al. 2003); farther out from* ***PA*** *brings one to more powerful systems of arithmetic (e.g., second-order* ***PA****). The key thing is the tracking of agents that, as they travel outwards, still understand matters well enough to answer "test questions." For example, and this is a very important feature of the figure, $\mathcal{G}$ denotes Gödel's First Incompleteness Theorem, and a test question that any agent with impressive cognitive intelligence should be able to answer, and justify with a supporting proof, is: "Using the techniques of finitary deduction for $\mathscr{L}_1$, is there in the abstract a proof or disproof of every assertion about arithmetic in the theory of Peano Arithmetic =* $\boldsymbol{PA}^{\vdash \mathscr{L}_1}$ *?"*

Fig. 4 at least out to the minimal circle that includes both $\mathbf{PA}^{\vdash}$ (which by default is the deductive closure of the axiom system under first-order inference) and $\mathbf{PA}^{\models}$ (closure under standard Tarskian model theory), which is distinguished by an understanding of the theory of arithmetic based upon the axiom system **PA**.[13] In the case of human agents, this first happened, at the latest, by 1930.

[13]Excellent introductory coverage of **PA**, sufficient for understanding Fig. 4, is provided in a book the lead author treasures for pedagogy in starting mathematical logic: (Ebbinghaus, Flum & Thomas 1994).

5. Universal Cognitive Intelligence

We now do three things to help further convey to the reader the nature of universal cognitive intelligence, $\mathscr{UCI}$, respectively: We build upon Test 1, given to the reader in our introduction, and issue an additional test (and after that point to further tests), to help the reader better understand at least the early levels in the progression of cognitive power at the heart of $\mathscr{UCI}$ (§5.1); then we venture the start, and *only* the start, of a full, formal definition of an arbitrary level of $\mathscr{UCI}$ (§5.1); and finally, as promised above, we quickly explain how Λ can be extended to the infinite case (§5.3).

5.1. *More Tests to Convey the Idea*

Let us return to teacher Alice from the introduction to the present paper, who now is going to issue Test 2 to you. Alice is actually herself a mathematician, a number theorist, specifically. And here's the background to the new test that she gives you, and we quote what she says to you:

"Recall again the familiar set $\mathbb{N}$ of natural numbers, that is $\{0, 1, 2, \ldots\}$, where of course the ellipsis here indicates that the progression continues infinitely. You have of course long been acquainted with this set; your exposure started in the first grade. Given the natural numbers as our domain of quantification, now consider the arithmetical proposition ν, which I assert to be true, that there is a number $n^\star \in \mathbb{N}$ which is such that, if $n^\star$ is nice, every number is nice. Very well, now your new two-part test question follows immediately."

Test 2

Is Alice's assertion correct? Prove that your answer is correct.

As in the beginning of the present essay, we encourage you to take a minute to reflect, with pen and paper if that is handy for you. ... If you managed to pass Test 2, and you can pass all such tests for arbitrary assertions about the natural numbers, you have cognitive intelligence at a fairly impressive level. This level, in terms of logic-programming, exceeds what Prolog can enable, since the inference schema and process at the heart of Prolog is only guaranteed to succeed when what is to be decided is in fact a theorem. Since $\mathscr{L}_1$ is only semi-Turing-decidable, we would already have to place your level of cognitive intelligence above Turing machines. By the way, now that we have delayed a bit and thereby given you more time, we get around to informing you that Alice is correct — though we leave the discovery of a proof to you, if you don't have one yet.[14] Expressed in $\mathscr{L}_1$, and using obvious symbols for logicizing Alice's claim, we have:

$$\nu : \qquad \exists n[N(n) \rightarrow \forall y N(y)].$$

If we collect the background information available to prove ν into Φ, then, in a direct parallel to what we ascribed to you in the case of Test 1 from the introduction to the present essay, we are here ascribing to you the cognitive intelligence needed to decide:

$$\Phi \vdash_{\mathscr{L}_1} \nu.$$

That level of intelligence, again, can't be achieved by any Prolog program, nor for that matter by any standard automated deductive reasoner operating at the level of $\mathscr{L}_1$.

For a book-length treatment of $\mathscr{UCI}$, we would now continue to give you — in keeping with the avowed psychometric orientation of $\mathscr{UCI}$; recall §4 — tests that see how far up the Arithmetic (which is limited to $\mathscr{L}_1$) and Analytic (which is limited to $\mathscr{L}_2$) Hierarchy you can go, and then we would shift to tests involving intensional operators, such as third-, fourth-, fifth-order ... false-belief tests (Bringsjord, Govindarajulu & C. 2019). In other words, we would be asking you to climb the left-branch progression for Cognitive Calculi shown in Fig. 7.

5.2. *A Full, Formal Definition of* $\mathscr{UCI}$*?*

Currently, we cannot provide a full, formal definition of $\mathscr{UCI}$. Doing so would mean supplying a formal definiens for at least the following definiendum on the left side of the biconditional:

Def$_\Lambda$ Agent $\mathfrak{a}$ is cognitive-intelligent at level L if and only $\phi(\Lambda)$.

In this definiens, ϕ is itself a formula, one in which 'Λ' occurs, at least once; and L is integer, real, or infinite ordinal.[15] Obviously it's going to be a challenge to fully flesh out Def$_\Lambda$, in subsequent work and corresponding publication. The present essay serving merely as an introduction to $\mathscr{UCI}$,

[14] Hint: Channel the mind of Euclid and approach indirectly, by assuming the opposite of Alice's claim, then use/prove quantifier shift to go from $\neg\exists n \ldots$ to $\forall n \neg \ldots$ in search of your needed contradiction.

[15] Even if we wished to stick with ascending levels of ever-higher cognitive intelligence that climb only extensional logics, we will need to move to transfinite numbers: infinitary logics are *based* upon such numbers; see e.g. (Dickmann 1975), and note the subscript in '$\mathscr{L}_{\omega_1\omega}$,' the one infinitary logic we use in the present essay (§5.3).

the burden of providing a fully instantiated Def_Λ is one we fortunately do not bear at present. But we hope the reader understands the structure of the definition.

5.3. *Expansion of Λ for $\mathscr{UCI}$*

Where $\mathfrak{a}$ is, as above, any agent (natural, biological, artificial, supernatural, alien, finite, infinite, divine, etc.) that can consume one or more test questions in the form of a set i of formulae (object-level or meta-formulae), the level of resulting cognitive consciousness is given as Λ, which, as explained in (Bringsjord & Govindarajulu 2020), and reviewed above, is a matrix $\mathbf{X}$. But we only looked at the finitary case for such matrices before, that is, in (Bringsjord & Govindarajulu 2020). Let's look briefly turn now, as promised above, to what the infinitary case looks like:

Example 1

Our agent $\mathfrak{a}$ here knows a basic number-theoretic base formulae expressed in the formal infinitary language[16] for the extensional infinitary logic $\mathcal{L}_{\omega_1\omega_1}$ (the smallest infinitary logic), viz.

$$\phi \coloneqq \exists x_1 \exists x_2 \ldots (x_2 > x_1 \land x_3 > x_2 \land x_4 > x_3 \land \ldots)$$

Assuming that ϕ pertains to the positive integers, this formula can be taken to express in the style of finitary formulae in the axiom system **PA** that integer 2 is greater than integer 1, that integer 3 is greater than integer 2, and so on *ad infinitum.*

Now, what are our Λ measures? Let us have the following four from the scheme of (Bringsjord & Govindarajulu 2020):

- μ^1 The "Boolean rank" of a base formula.
- μ^2_i The amount of occurrences of a relation R_i in a base formula.
- μ^3 The number of distinct relations in a base formula.
- μ^4 The number of quantifiers in a base formula.

Then we have an extension to the infinitary case, unprecedented for Λ relative to earlier work, but simple and easily understood nonetheless:

$$\mathbf{X}_\mathfrak{a} = \begin{pmatrix} \aleph_0 & \aleph_0 & 1 & \aleph_0 \end{pmatrix}$$

[16]The language allows infinite disjunctions and conjunctions of the cardinality of $\mathbb{N}$, but restricts the number of quantifiers used in any formula to a particular $k \in \mathbb{N}$. Efficient coverage is provided in (Ebbinghaus et al. 1994).

6. $\mathscr{UCI}$ and the Formal Hierarchies

$\mathscr{UCI}$ harmonizes strikingly well with the standard, formal hierarchies of computational[17] power, including such power above standard Turing machines, because such hierarchies are all based upon ascending complexity and depth of formulae in formal logics.[18] We obviously can't herein systematically canvass these hierarchies, which include the Polynomial, Chomsky, Arithmetic, and Analytic. (As we said earlier in the present essay, we don't include a pictorial overview of the first of these, but such a description of the latter three can, again, be found in Fig. 5.) Of these, which we now proceed to synoptically discuss, we focus upon the Arithmetic, in part because most readers can be counted upon to have an understanding of first-order logic $= \mathscr{L}_1$, which is reasonably well-covered in all the major, comprehensive AI textbooks of today (this holds e.g. for the dominant such textbook: Russell & Norvig 2020).

6.1. *Cognitive Intelligence & the Arithmetic Hierarchy*

In his teaching, the first author has found that it's actually quite easy to build before one's eyes the Arithmetic Hierarchy (AH), and thereby understand it.[19] To do so, one can start with a Turing-decidable relation R to get the climb up AH going. For example, suppose that R logicizes the property/relation of a particular Turing machine $\mathfrak{m}$ taking some particular input a in and, after executing its program, giving some particular b as output in $k \in \mathbb{N}$ steps. Formally:

$$(1) \qquad R(\mathfrak{m}, a, b, k).$$

In terms of cognitive intelligence, you the reader, and indeed all neurologically normal, educated human agents have at least a level of cognitive intelligence sufficient for deciding whether or (1) holds, for any particular quadruple of assignments to the constants the relation is to hold over. (**Proof**: You can simulate the Turing machine in question on the input

[17]Those who insist upon reserving 'computation' and 'computational' as terms for Turing-level and sub-Turing-level information processing can simply view the hierarchies in question as describing *information processing*.

[18]Finitary measurement via Λ of the amount of cognitive consciousness in an agent, and as such of the cognitive intelligence in an agent, link directly to these hierarchies, because the formulae that are scored by Λ can be placed within the hierarchies. But we must leave details aside here in the interest of economy.

[19]A conventional, non-do-it-yourself introduction to AH is provided in (Davis, Sigal & Weyuker 1994).

for k steps. ■) We can also say, from the perspective of Λ, that a human who carries out the cognizing here reaches some particular level of cognitive consciousness = cognitive intelligence. But this level is quite humble. We can climb up in AH, rather quickly, to a level, specifically Π_2, that is much greater cognitive intelligence, and that is above anything a Turing machine can do. A specific case that is a favorite of the first author is the cognitive intelligence needed to create valid computer programs. We assume that in order to do this, the agent in question must be able to judge whether two programs compute the same functions or not (= whether two programs, while syntactically divergent, nonetheless are equivalent). If we assume that this capacity holds in the general case, then using the very same relation R we have just allowed ourselves for (1), we can logicize to obtain this formula:

$$(2) \qquad \forall u \forall v[\exists k[R(\mathfrak{m}_1, u, v, k) \leftrightarrow \exists k' R(\mathfrak{m}_2, u, v, k')].$$

Formula (2), in which of course both $\mathfrak{m}_1$ and $\mathfrak{m}_2$ are free variables, has a leading sequence of universal quantifiers, and — if converted to prenex-normal form so that existential quantifiers are moved to the left — then a sequence of existential quantifiers, and this pattern is what makes it Π_2.[20] An agent able to ascertain whether (2) holds for a pair of Turing machines (or computer programs) has a level of cognitive intelligence at Π_2, and once again we could employ Λ to obtain a score of cognitive consciousness/intelligence as well.

6.2. *The New Logic-Machines Hierarchy $\mathfrak{LM}$*

The "engine" that generates the new logic-machines hierarchy $\mathfrak{LM}$ (see Fig. 7) is *pure general logic programming*, or just PGLP (see Fig. 6), a new logic-programming paradigm for the Bringsjord-found universal rational calculus sought by Leibniz throughout his lifetime.[21] The calculus in

[20]We skip some niceties for economy: For example, some readers may ask: "Where's the *set* to be decided? I thought what gets decided (or not) in these hierarchies are sets." The answer is that the free variables give us a set, and we can easily make this explicit by set-builder notation, e.g. by:

$$\{\mathfrak{m}_1, \mathfrak{m}_2 : (2)\}.$$

[21]In a 2016 celebration of Leibniz's death after 200 years, at the University of Turin, Bringsjord announced in his invited lecture that he had found Leibniz's universal rational calculus — and that he would wait to see if anyone could discover his discovery. Later, an in-print declaration was made in (Bringsjord et al. 2021). Many have over centuries pondered and sought what it is that Leibniz sought, and some have proposed some things

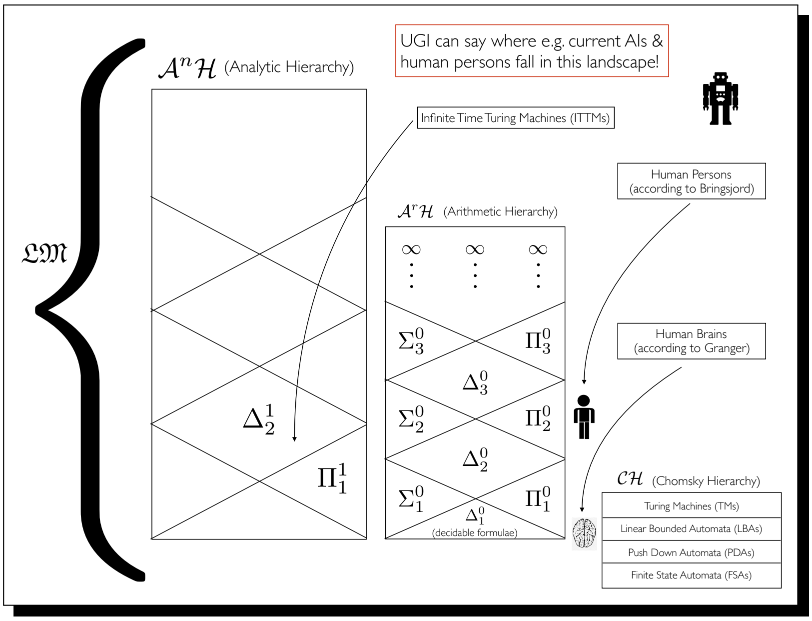

Fig. 5. The Main Logic-based Hierarchies. *The Chomsky Hierarchy will be familiar to most readers; its max is a full Turing machine. The symbols 'Π' and 'Σ' here just indicate that nature of the starting block of quantifiers, either universal in the case of* Π *or existential in the case of* Σ. *For coverage of the Arithmetic Hierarchy in purely formal terms, see (Davis et al. 1994), and in connection with AI, see (Bringsjord & Zenzen 2003).*

---

that at least genuinely relate to what Leibniz dreamed of — but they have missed the mark, for many reasons. E.g., consider:

> Today, the best candidates to be considered universal formal languages are the higher-order logics based on type theories, which form the basis of proof assistants such as Coq (Paulin-Mohring, 2015) and Isabelle (Wenzel, 2015). The universality of these logics, from both theoretical and practical points of view, is evidenced by their ability to embed/encode other logics (and even simulate Turing machines) and by their application in many different domains, including Mathematics, Software and Hardware Verification, and even Metaphysics. (Woltzenlogel Paleo 2016, p. 316).

The same position, that higher-order logics at the very least come close to the arrival of Leibniz's dream, is articulated by Benzmüller (2017). Unfortunately, higher-order logics, which form the center trunk/branch in Fig. 7, are subsumed in but absolutely cannot possibly be the universal rational calculus or language Leibniz wanted, for many reasons; here are three: **One**, they are deductive, and as such include no inference question is the language for expressing and defining all cognitive calculi in the family of such, as they have been defined in numerous publications authored by Bringsjord (with others, often), and the framework for *generating* a given cognitive calculus for a given domain of application to be reasoned in and about. PGLP is the *calculus ratiocinator*, the machine or mechanical system that brings the universal rational calculus, or *characteristica universalis*, to concrete, implemented life. But let us dispense with further discussion of such matters, which are best suited for considering Leibniz and his legacy in other venues, and turn to PGLP in earnest, independent of these deeper issues.

PGLP can be viewed abstractly as conforming to the annotated Fig. 6. The three main elements shown in this figure are: $\mathbb{P}$, a program; $\mathbb{R}$, a reasoner; and $\mathbb{C}$, a checker. Territory above the horizontal line is purely specificational;[22] L is simply the background formal language in which both the program $\mathbb{P}$ and query `q` are expressed. This formal language is bounded only by the bounds of formal logic/meta-logic, whatever they might be. This means, for example, that L might be the object-level formal language underlying the infinitary logic $\mathcal{L}_{\omega_1\omega}$ we used above; in this case, wffs would be allowed to be (countably) infinitely long. The symbols `Y`, `N`, and `U` correspond to "Yes," "No," and "Undetermined." Usually, $\mathbb{P}$ consists simply of a set of formulae against which the query is issued. Programming in PGLP consists in setting this up to issue such a query, and then initiating the activity of the reasoner and the checker. Further details regarding PGLP are beyond the scope of the present essay.

---

schemata in *in*ductive logic, which in its philosophical tradition has informal versions of such schemata (see e.g. (Johnson 2016)), nor do they include provision for probability or likelihood, such as is seen in *pure inductive logic* (Paris & Vencovská 2015); **two**, they are purely symbolic/linguistic, and thus include no inference schemata for diagrammatic/visual information such as are specified in (Arkoudas & Bringsjord 2009); and **three**, they leave aside infinitary logics and infinitary reasoning, since for starters each formulae in higher-order logics is a finite string. The meta-language of cognitive calculi, which is Leibniz's universal rational calculus/language, has this trio of missing elements, and more.

[22]PGLP renders program verification an incalculably easier affair than the well-known burden created by programs in the procedural/imperative, object-oriented, functional, or even standard logic-programming paradigms. The reason is that since specifications *are* programs, program verification reduces to proof/argument checking; this approach was first indicated in (Arkoudas & Bringsjord 2007), and then in the subsequent (Bringsjord 2015) explicitly set out and proposed.

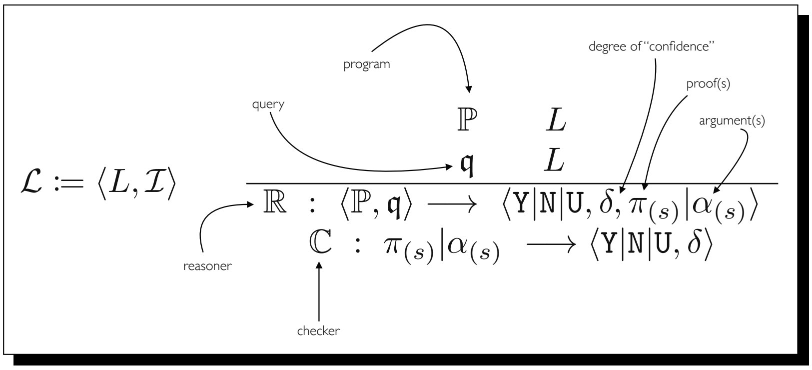

Fig. 6. The Annotated "Engine" for the Hierarchy $\mathfrak{LM}$: PGLP. *Pure General Logic Programming is what generates the hierarchy. In the simple case of PGLP subsuming Proplog and Prolog, the background logic is $\mathscr{L}_{PC}$ and $\mathscr{L}_1$, resp. The hierarchy ascends by virtue of the fact that the power of the combination of logic and corresponding automated reasoner increase.*

6.3. *What About Real-Number-Based Hierarchies?*

Theoretical computer science has traditionally been based upon discrete structures, and the established hierarchies we have visited above are no exception. The reader will have noticed early on that it's the natural numbers which has been a cornerstone herein for most of our presentation of $\mathscr{UCI}$, and of course specifically for Λ (in both the finitary and infinitary cases) and the established hierarchies. There is for example nothing whatsoever "contaminating" AH from the space of uncountable sets, such as the reals = **R**. What, then, is the relationship between $\mathscr{UCI}$ and the reals (and, indeed, above)? We offer but two remarks, under space limitations. One, there are in fact real-based hierarchies, and we are working out the relationship between $\mathscr{UCI}$ and one of them, one given in (Mycka & Costa 2007). Two, while we have for ease of exposition and focus restricted our attention to ℕ, and to basic arithmetic over this set (recall Fig. 4 and the discussion revolving around it), formal logic can of course be used to capture continuous mathematics, without issue. This we have learned from the sub-part of the discipline of formal logic known as *reverse mathematics*.[23] In general then, universal psychometrics extended to tests regarding such branches of mathematics as analysis, and reasoning over the axiom systems we know to be sufficient to give us continuous mathematics, will be possible in the future to incorporate into $\mathscr{UCI}$.

[23] A wonderful starting point for the interested and formally inclined is (Simpson 2010), which covers the reach of various axiom system into mathematics in general.

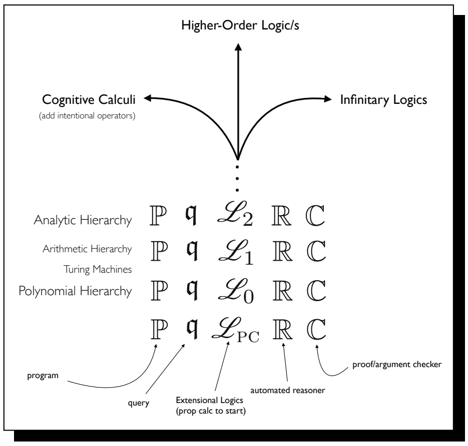

Fig. 7. An Impressionistic View of $\mathfrak{LM}$ Hierarchy. *When one is referring only to purely extensional, finitary logics, a one-dimensional hierarchy is easy to build, and — at least initially — achieve results with respect to, and then catalogue those results. This would be for the trunk of $\mathfrak{LM}$, and its center branch in the figure here. But there are two other branches, one for intensional logics of every greater complexity, and one for infinitary extensional logics. The figure here gives merely an impressionistic view because in point of fact many additional branches would need to be specifically shown and charted. For instance, is it clear from the logic given in (Arkoudas & Bringsjord 2009) that any finitary extensional logic can be extended into a* ***heterogeneous*** *logic, that is, a logic that includes not only a linguistic/symbolic formal language, but provision also for diagrams/pictures. Our presentation of $\mathfrak{LM}$ in the present paper suppresses explicit calling out of such nuances.*

7. Related Work: Some Remarks

7.1. *Artificial General Intelligence (AGI)*

AGI, artificial general intelligence, is, at least as some of AGI researchers view their discipline and research, quite relevant to $\mathscr{UCI}$.[24] To many readers, this is probably plain, to the point that they would fully expect the present section.

[24]Outside of AGI, we are not aware of related attempts to erect a comprehensive framework along at the general lines of $\mathscr{UCI}$. This is why the present section is limited to discussion of AI under related-to-$\mathscr{UCI}$ work. Of course, one can view the construction and exploration of the standard hierarchies (Polynomial, Arithmetic, and Analytic), and perhaps other non-standard ones (e.g., for quantum computation and real-number-based computation) as quite relevant. (We are not aware of anyone who has done this, but our knowledge is of course limited.) Such a view is, as we have made clear, exactly our own — but at any rate $\mathscr{UCI}$ subsumes all these by virtue of the new $\mathfrak{LM}$ hierarchy.

A subfield of Artificial Intelligence (AI), Artificial General Intelligence (AGI) can generally be classified as the field that explores the creation of computational agents that possess some level of *general intelligence*: the ability to exhibit complex problem-solving capabilities in an arbitrary environment, akin to the ability of humans (but not necessarily at the same level as humans) (Goertzel 2014, Goertzel 2015, Wang 2019). As AGI focuses on a broad overarching goal, inevitably there are many camps in AGI, each based upon its own approach to the problem (Duch, Oentaryo & Pasquier 2008, Goertzel 2015). These groupings are key to understanding links between $\mathscr{UCI}$ and AGI, as $\mathscr{UCI}$ can be better applied to some groups than to others. Obviously, camps that are not overtly logicist bear little connection to $\mathscr{UCI}$. Here's a simple but nonetheless triadic breakdown of approaches in AGI, the first of which relates most closely to $\mathscr{UCI}$:

- **The Symbolic Approach**: Here logic is in fact the basis for memory and reasoning. Knowledge in these systems consist of statements from which new knowledge can be derived by logical reasoning. New statements may also be added by way of fully logic-based perception see (Wang 2013*b*). Different approaches use different ontologies and different logics with different properties to optimize for the type of reasoning to be done (Gust, Krumnack, Schwering & Kühnberger 2009). Invariably, at least so far, relative to the logics upon which $\mathscr{UCI}$ is based, logics in this approach to AGI are inexpressive, and reasoning is correspondingly simple. In particular, anything represented in and reasoned over in this AGI approach can be reduced to information and processing in $\mathscr{UCI}$ (perhaps with tailor-made inference schemata as needed) at the level of only $\mathscr{L}_1$, augmented perhaps with a few intensional operators. Some notable members of the symbolic camp are Wang's NARS (Wang 2013*a*) system and Shaprio's SNePS and GLAIR architectures (Shapiro & Bona 2010), all of which encode symbolic representations of knowledge into a graph representation.
- **The Emergent Approach**: This approach focuses on creating agents whose memory and learning take the form of connectionist systems. The emergent approach assumes, naturally enough, an emergent hypothesis: that symbolic reasoning and learning can emerge from basic connections and interactions between nodes, as they perhaps do (at least in part) in the human brain. "Knowledge" in emergent systems is encoded within the weights and connections between nodes of a network, which may

evolve over time for "Learning." $\mathscr{UCI}$ can only subsume emergent approaches indirectly, via reasoning over the declarative content that axiomatizes and thereby captures all connectionist systems. (In this regard, see (Bringsjord 1991).) Direct translation between the sub-symbolic content in such systems to declarative content expressed in one or more logics is impossible, and unwanted. In addition, $\mathscr{UCI}$ would in principle have linkages to processing in artificial neural networks that are more powerful than Turing machines (e.g. analog chaotic neural nets; see Siegelmann 1999).

- **Hybrid Approaches**: Hybrid AGI systems aim to combine emergent and symbolic approaches. According to Duch et al. (2008), hybrid approaches suffer from the same shortcomings as emergent approaches: they have "difficulty in realizing higher-order cognitive functions" such as reasoning over arbitrarily complex/iterated propositional attitudes, which as we have seen are pivotal to cognitive consciousness/TCC. Despite this shortcoming, there is some promise that $\mathscr{UCI}$ could obtain symbolic representations for things at the sub-symbolic level, which would secure the kind of direct connection to $\mathscr{UCI}$ impossible for the emergent approach.

AGI stands in stark contrast to today's mainstream "narrow" AI systems, usually machine-learning models which are trained on massive datasets to excel in one particular task. For $\mathscr{UCI}$ it is also important to contextualize "human-level" AI in the context of AGI. "Human-level" AI can be thought of as a goal of AGI, but it is only a point on a spectrum of general intelligence that AGI agents fall on. This means that AGI researchers of either a thoroughgoingly or substantive logicist bent can presumably locate their ambitions for future AGI systems in the $\mathscr{UCI}$ space. Unlike other measurements of intelligence, we can quantify very well where humans fall when measured with respect to Λ. Due to the inherently cognitive nature of AGI systems, we think It should be fully feasible, in the future, that any symbolic or hybrid approach in/to AI can be placed within some level/s of $\mathscr{UCI}$. This is indicated by the reference to the relevant bridge theorems between AGI and $\mathscr{UCI}$ in Fig. 3.

One particular point worth nothing is that $\mathscr{UCI}$ stands in contrast to Gortezel's conception of intelligence (Goertzel 2021, p. 5), since he writes that "Intelligence in general must be considered as an open-ended phenomenon without any single scalar or vectorial quantification." This of course runs completely counter to the spirit and specifics of $\mathscr{UCI}$. The fact is, in common everyday language we often compare the intelligence of human and nonhuman animal cognitive agents in line with how Λ works. Consider for instance a person "on the street" who states: "Humans are more intelligent than dogs." Clearly intelligence in the sense used here is as a single scalar. However, if our man is asked why he believes humans are more intelligent than dogs, he is likely to resort to informal correlates of measures that are at the heart of Λ. For instance, will our representative human here fail to agree that Fido believes that humans believe that Fido and other canines have no beliefs about the future 10 years hence? Probably not.

7.2. *Universal Artificial Intelligence (Hutter)*

Marcus Hutter has introduced the concept of "universal artificial intelligence" in his eponymously titled monograph *Universal Artificial Intelligence* (Hutter 2005). His overarching computational model — the *AIXI Model* — features an agent seeking to maximize the reward that it obtains in an unknown environment. The purpose of Hutter's model is to provide a formal, unbounded model for AI. We rest content here with but a few remarks upon Hutter's model. A simpler description than what is provided in the monograph cited immediately above can be found in (Legg & Hutter 2007), and more recently in our own overview of AI (Bringsjord & Govindarajulu 2018) we provide an entire sub-section on AIXI.[25]

Hutter's AIXI model is a formal one of an agent operating in an unknown environment. Given certain reasonable assumptions, this model is universally optimal across all possible environments and has certain provably optimality properties. This model builds upon Solomonoff's Theory of Induction (Solomonoff 1978) and Sequential Decision theory (Sutton & Barto 1998).

More specifically, the AIXI model is an agent-based model in which an agent $\mathfrak{a}$ performs an action in an environment E at time time t. The environment responds with a unique percept and a reward. This continues until some time termed the *horizon* or the *lifetime* of the agent.

AIXI does have at least one property that in general accords with the fact, presented, explained, and discussed above, that $\mathscr{UCI}$, for the most part, is based upon logics whose formulae are allowed to be uncomputable (e.g., infinitely long conjunctions cannot fit on any finite square

[25] Go to: https://plato.stanford.edu/entries/artificial-intelligence/aixi.html.

of a Turing-machine tape), and whose corresponding automated reasoners in $\mathfrak{LM}$ are allowed to exceed what a Turing machine can do: AIXI is Turing-uncomputable. At an intuitive level, which is sufficient here, the uncomputability of AIXI is due to the involvement of Turing-uncomputable Kolmogorov complexity and an infinite sum in the specification of AIXI's optimal action. Put intuitively, AIXI is Turing-uncomputable only because it "backs into" Turing-uncomputability. After all, the model rests directly and in no small part on a multi-tape Turing as a fixed part of the theory. In stark contrast, there are levels of cognitive intelligence in $\mathscr{UCI}$ that front and center rest on information-processing that is invariably and essentially beyond the operation of a Turing machine. E.g. consider the cognitive intelligence required to decide Π_2 problems in the Arithmetic Hierarchy, mentioned above. Another example of the fundamental divergence between AIXI and $\mathscr{UCI}$ is clearly seen when one simply takes note of the fact that Kolmogorov complexity, crucial to AIXI, by definition treats *programs* as standard programs only. PGLP programs, in contrast, as we have seen, are arbitrary collections of formulae or meta-formulae constrained only by the formal language of the formal logics or logics selected. In short, and in sum, universal artificial intelligence is certainly not universal cognitive intelligence; and specifically, an artificial agent at the higher infinitary reaches of $\mathscr{UCI}$ far exceeds the cognitive power of any agent in the AIXI framework.

8. Objections; Replies

8.1. *Objection #1: What about intentionality?*

The first objection can be encapsulated as follows: "$\mathscr{UCI}$ inherits from TCC and Λ the centrality of the phenomenon known by philosophers as *intentionality*, a term coined by Brentano.[26] Roughly, as you no doubt well know, intentionality is the 'aboutness' of at least some mental states, especially the very states that you have placed at the heart of $\mathscr{UCI}$. I refer for example to epistemic c-conscious states such as *Selmer's believing that Naveen believes that Selmer believes that human persons are immaterial.* Here, Selmer, somehow, has a belief that is genuinely *about* Naveen — and figuring out what it is that makes this the case is an enduring philosophical conundrum. Moreover, some have claimed that intentionality is itself an

[26] An overview of intentionality with a level of detail more than sufficient for the present paper is provided in (Jacob 2003/2019).

immaterial, or non-physical, phenomenon, and that as a result human persons are *themselves* immaterial. Surely you don't want $\mathscr{UCI}$ to be saddled with a seemingly insoluble problem, let alone with the long-dead substance dualism of Descartes."

Our rejoinder is short, simple, and decisive: $\mathscr{UCI}$ will no doubt raise for many philosophers the spectre of intentionality. But no matter, and no problem; for we are concerned with the abstract *structure* of c-consciousness, and the all-the-more abstract concept of how to measure it (via tests, and Λ, of course). We take an intensional approach[27] to formalizing those states that are (for many philosophers of mind) intentional, but we premeditatedly leave behind the baggage that usually shows up on the doorstep of such philosophers. Agents blessed with high levels of $\mathscr{UCI}$ enter into robust c-conscious states by definition,[28] and such agents can reason with and over such states represented in intensional logics, but that in no ways forces us to affirm even a shred of the positions advocated by Brentano and followers (such as Roderick Chisholm).

8.2. *Objection #2: $\mathscr{UCI}$ is Too Abstract, From the* Practice *of AI*

The objection that $\mathscr{UCI}$ is hopelessly abstract has been pressed against us in more than a few conversations (with antecedents of this objection being made against us in connection with the TCC/$\mathcal{CA}$/Λ bases of $\mathscr{UCI}$, and we bring it to the reader's attention here for two reasons: viz., because it stands to reason that some of our readers, too, will and be inclined to object along this line, and because the objection, while in and of itself anemic, triggers a rebuttal that is important. The rebuttal is expressed, pictorially, in Fig. 3. This shows our vision that AI of today, AI pursued by *practitioners*, can (and as far as we're concerned, should) be carried out on the basis of $\mathscr{UCI}$. What do we have in mind? The basic idea behind the picture of Fig. 3 is actually quite simple, as soon as one understands the "bridge" theorems the figure so prominently in it. To understand these theorems, let's

[27] Since we use intensional logics. All our cognitive calculi are intensional logics.

[28] For the record, and to erase any confusion, the reason is simply that if an agent is at a time or over an interval highly cognitively intelligent, that is because, fundamentally, put non-technically, the cognitive depth of their states in this period of time is high. But that that depth is high is precisely what makes their level of cognitive consciousness is high. If for instance an agent provides, with justification, a correct answer to the fourth-order false-belief task (recall above), they have an impressive iterated belief, and a justification in support of it, and by Λ that belief and justification will be high — which is to say that the agent's cognitive consciousness is high.

consider one of the sub-areas of AI shown in Fig. 3: "Linguistic Cognition" and "Communication,' each of which have their own box, with both boxes within a larger one. From a practitioner point of view, this larger box can be identified with natural-language processing (NLP), which is composed in AI by natural-language understanding (NLU), and by natural-language generation (NLG). In the case of NLU, the core challenge is to design and implement an artificial agent that, taking in natural language, for instance English, can *understand* that language. Given this, and given the assumption that the type of NLU that stands to be most relevant to $\mathscr{UCI}$ at least takes declarative knowledge seriously, it's not difficult to see how bridge theorems can allow NLU to be reduced to the formal languages, inference schemata, and automated reasoning that stand at the heart of $\mathscr{UCI}$. A great example is the peerless work on knowledge-rich NLU by McShane & Nirenburg (2021). The reason is the the format in which they represent knowledge, and the processes use to exploit that knowledge in service of NLU, can all be recast as into use of a cognitive calculus that has, as its finitary extensional component, $\mathscr{L}_1$ and automated reasoning for it.[29]

We would be remiss if we didn't share that our response to the $\mathscr{UCI}$-is-too-abstract objection includes that we are actively attempting to build a physical robot (PERI.2[30]) with unparalleled language-mediated manipulation capacity, and high levels of cognitive consciousness according to Λ, and correspondingly high levels of cognitive intelligence. This engineering is in line with robots and consciousness as described in (Chella, Cangelosi, Metta & Bringsjord 2019), and is designed to take account of the coverage of artificial consciousness achieved in (Chella & Manzotti 2007). More specifically, and connected to concepts covered in the present essay, in general conformity with the kind of hybrid architecture presented in (Chella, Frixione & Gaglio 2000), PERI.2 combines logicist representation and reasoning ability (including the treatment of visual information) provided by our cognitive calculi, but perceptual capability, enabled by ARCADIA (Lovett, Bridewell & Bello 2021), that is outside logicist information and processing, but nonetheless tightly integrated with this processing.

[29]Of course, stating and proving the bridge theorem here from frame-based knowledge representation and reasoning to the formal logics and automated reasoners of $\mathscr{UCI}$ is beyond scope, and will have to wait for another day.

[30]The predecessor to PERI.2 was PERI, the robots concretized Psychometric AI/PAI discussed herein in §4. PERI and some of its feats are presented in (Bringsjord & Schimanski 2003).

8.3. *Objection #3: Gamez*

$\mathscr{UCI}$ runs deeply counter to the claims of Gamez (2020), since, as we have explained at some length, $\mathscr{UCI}$ is (i) based upon the explicit measurement, through Λ, of the complexity of reasoning (which certainly can be counted as a test) on the part of a given agent, and (ii) also overtly upon performance on tests of cognitive ability and skill (as e.g. in the case of testing whether a given agent has deeper and deeper understanding of arithmetic/number theory; recall Fig. 4). For confirmation of the position of Gamez, we for example have this from him:

> In humans we use batteries of tests to indirectly measure intelligence. This approach breaks down when we try to apply it to radically different animals and to the many varieties of artificial intelligence. (Gamez 2020, 51)

As confirmed in this quote, Gamez reasons from two premises to the rejection of approaches like ours that seek to measure arbitrarily high levels of intelligence in machines (i.e. in AIs or artificial agents). We reject both of these premises, as we now briefly explain.

Gamez's first premise is that measuring the intelligence of (nonhuman) animals is problematic. We happily concede this for the sake of argument, despite our knowing that more than a few cognitive scientists have administered cognitive tests to animals, with results, as these researchers see things, confirming that some nonhuman animals are capable of performance on these tests that implies an appreciable level of intelligence (and, indeed, what would in fact be an appreciable level of cognitive intelligence on the scales/spectra of $\mathscr{UCI}$) (e.g. see Taylor, Hunt, Medina & Gray 2008).[31] Nothing follows with regard to $\mathscr{UCI}$ and its formal bases and psychometric roots for the simple reason that it doesn't concern itself with near-zero cognitive intelligence. No nonhuman animals can comprehend domain-independent abstract inference schemata, so for that reason alone they are outside our concern. Recall that we started with Test 1, issued to the reader at the very outset. This test is one the understanding of which requires a prior understanding of domain-independent inference schemata — but we all know very well that such understanding is absent in the case of nonhuman animals. In fact, such creatures have no capacity in natural language at the level of humans (i.e. at the top of the Chomsky

[31] Our view, for what it's worth, is that these investigations are intrinsically inconclusive. We point out that others claim poor performance on the very same class of tasks by monkeys; see e.g. (Visalberghi & Limongelli 1994).

Hierarchy), so they can't even read the simplest of the very tests in the $\mathscr{UCI}$ paradigm.

And what of the second premise affirmed by Gamez? Perhaps this one gives rise to an objection that isn't merely a non-starter. This premise is that tests cannot be applied to AIs. But this premise renders any argument such as his circular. What is at issue is whether levels of cognitive consciousness can rationally and otherwise acceptably be mapped to levels of cognitive intelligence to yield $\mathscr{UCI}$, in symbiotic conjunction with tests of cognitive ability and skill. His second premise is simply the denial of part of the very foundation upon which the $\mathscr{UCI}$ edifice rests. What Gamez needs to provide is some separate, standalone argument as to why cognitive consciousness can't simply be directly, formally connected to assessment via Λ and, in general, tests of cognitive ability and skill along the lines we have presented above (e.g., Is this axiom system for arithmetic consistent?). This will be an acute uphill battle for Gamez, for the simple reason that AI itself, as concretely practiced, places tests front and center in the field, which is why the landmark achievements in AI have so often been constituted by success on tests, usually game-based ones (for chess, checkers, *Jeopardy!*, Go, etc.).

8.4. *Objection #4: But Emotional, Artistic, . . . Intelligence?*

At least some of those skeptical about our logicist approach to cognitive intelligence can, in our experience, be counted upon to object that the emotions, and ergo emotional *intelligence*, is beyond the reach of $\mathscr{UCI}$. The critic here seems not to have appreciated that according to TCC, cognitive consciousness is reduced to the *formal structure* of cognition in a a given agent — and why shouldn't TCC be able to capture such structure in the case of emotions? It is able to do this. Witness that the present objection is exploded in large part by prior work of others making use only of fragments of the highly expressive cognitive calculi we have invented, continue to invent, and avail ourselves of when engineering intelligent machines. A nice example of such prior work is (Adam, Herzig & Longin 2009), in which it is shown that, with the exception of *love*, all the emotions in the dominant account of emotions in cognitive science [the so-called "OCC theory," provided in (Ortony, Clore & Collins 1988)] can be logicized. Additionally, in work along this logicist line, but based upon our cognitive calculi, the second author has led the capture of sophisticated emotions that underlie virtuous behavior; see for example (Govindarajulu, Bringsjord, Ghosh & Sarathy 2019).

9. Two Confessions and Conclusion

Alert readers will doubtless have noticed that nothing above addresses unconscious or generally non-occurrent mental states. Assuming that such states exist in the human case (for our money, they certainly exist in in the abstract logico-mathematical space of cognizers in general), the question (Q) arises for us, viz.,

> (Q) Can c-conscious states be non-occurrent?

(Q) arises because in fact it's hard to deny that much human activity occurs beneath the surface of awareness, and is subject at least to quasi-rational constraints; consider the simple activities of reaching or walking. In particular for the first author of the present essay, (Q) not only arises, but is pressing, because he has defended at length the proposition that in the human case cognitive intelligence exceeds the Turing Limit = Σ_1 (Bringsjord & Zenzen 2003),[32] and the information-processing that enables that level of $\mathscr{UCI}$ is thought by him to be at least in significant part unconscious processing. Nonetheless, and here the first confession, we at this point understand c-consciousness, and $\mathscr{UCI}$, to exclusively pertain to occurrent mental states; we thus simply answer (Q) in the negative. Put in terms of problem-solving, and the tests we discussed above (along with many additional tests along the same line), our current assumption about the human case is that unconscious processing only has "intelligence value" insofar as that processing provides to occurrent states and processing outputs that foster answers that are consciously understood and provided as ultimate outputs. Even the first of the tests we presented above, Test 1, may trigger in our readers unconscious processing that yields the answer, but that answer then must be wordsmithed occurrently. Future development of the $\mathscr{UCI}$ paradigm will include deeper consideration and analysis of the unconscious direction.

And now for the second of our confessions: As by now the reader well knows, $\mathscr{UCI}$ is to a significant degree based on TCC and Λ, which, as we've seen, entails that $\mathscr{UCI}$ is based on c-consciousness, which in turn entails that p-consciousness is completely excluded. Putting the matter starkly:

[32]See also (Bringsjord & Arkoudas 2004, Bringsjord, Kellett, Shilliday, Taylor, van Heuveln, Yang, Baumes & Ross 2006).

A monumentally intelligent agent by $\mathscr{UCI}$, say one that cognizes in the paradise of transfinite levels of Λ courtesy of effortless spinning infinitely long proofs in $\mathscr{L}_{\omega_1\omega}$ its mind, can at the same time nonetheless be an agent wholly bereft of qualia. This implies that we ought to make a concession, viz. that if there are some types or levels of cognitive intelligence that exploit p-conscious states, $\mathscr{UCI}$ is incomplete. We do not assert *here* that such types or levels exist; but elsewhere one of us has issued such an assertion: (Bringsjord, Noel & Ferrucci 2002). That assertion revolves around what suffices here to be classified as only a concern: the concern, specifically, that *creativity* falls outside of $\mathscr{UCI}$. Future work on $\mathscr{UCI}$ must include systematic investigation of this concern.

Acknowledgements

The invention and introduction of $\mathscr{UCI}$ has been made possible by the generosity of many supporters. In particular, funding from AFOSR (award #FA9550-17-1-0191) and from ONR (in multiple awards) to Bringsjord and Govindarajulu has been indispensable. An anonymous reviewer insightfully pointed out a number of weak spots in the $\mathscr{UCI}$ conception; we are most thankful, and have done our best to fortify the conception accordingly. Inadequacies that might remain are avowedly our fault. Bringsjord is deeply thankful for being a researcher in the ANR-sponsored PROGRAMme project "What is a (computer) program? Historical and Philosophical Perspectives," participation in which has enabled a number of the ideas presented herein (e.g. PGLP and $\mathfrak{LM}$) to be conceived and — at least to a degree — refined. Leadership of PROGRAMme by Liesbeth De Mol and Tomas Petricek is specifically much appreciated. Finally, we are grateful for the ongoing wisdom, patience, and guidance of Antonio Chella, which has been and continues to be crucial to our work, and for his longstanding and continued leadership in the intersection of AI and consciousness.

References

Adam, C., Herzig, A. & Longin, D. (2009), 'A Logical Formalization of the OCC Theory of Emotions', *Synthese* **168**(2), 201–248.

Arkoudas, K. & Bringsjord, S. (2007), 'Computers, Justification, and Mathematical Knowledge', *Minds and Machines* **17**(2), 185–202.
URL: *http://kryten.mm.rpi.edu/ka_sb_proofs_offprint.pdf*

Arkoudas, K. & Bringsjord, S. (2009), 'Vivid: An AI Framework for Heterogeneous Problem Solving', *Artificial Intelligence* **173**(15), 1367–1405.
URL: *http://kryten.mm.rpi.edu/KA_SB_Vivid_offprint_AIJ.pdf*

Arora, S. & Barak, B. (2009), *Computational Complexity: A Modern Approach*, Cambridge University Press, Cambridge, UK.

Ashcraft, M. & Radvansky, G. (2013), *Cognition*, Pearson, London, UK. This is the 6th edition.

Benzmüller, C. (2017), 'Universal Reasoning, Rational Argumentation and Human-Machine Interaction'.
URL: *https://arxiv.org/abs/1703.09620*

Block, N. (1995), 'On a Confusion About a Function of Consciousness', *Behavioral and Brain Sciences* **18**, 227–247.

Boolos, G. S., Burgess, J. P. & Jeffrey, R. C. (2003), *Computability and Logic (Fourth Edition)*, Cambridge University Press, Cambridge, UK.

Bringsjord, S. (1991), 'Is the Connectionist-Logicist Clash one of AI's Wonderful Red Herrings?', *Journal of Experimental & Theoretical AI* **3.4**, 319–349.

Bringsjord, S. (1997), 'Consciousness by the Lights of Logic and Common Sense', *Behavioral and Brain Sciences* **20**(1), 227–247.

Bringsjord, S. (1999), 'The Zombie Attack on the Computational Conception of Mind', *Philosophy and Phenomenological Research* **59**(1), 41–69.

Bringsjord, S. (2007), 'Offer: One Billion Dollars for a Conscious Robot. If You're Honest, You Must Decline', *Journal of Consciousness Studies* **14**(7), 28–43.
URL: *http://kryten.mm.rpi.edu/jcsonebillion2.pdf*

Bringsjord, S. (2008), Declarative/Logic-Based Cognitive Modeling, *in* R. Sun, ed., 'The Handbook of Computational Psychology', Cambridge University Press, Cambridge, UK, pp. 127–169. This URL goes to a preprint only.
URL: *http://kryten.mm.rpi.edu/sb_lccm_ab-toc_031607.pdf*

Bringsjord, S. (2011), 'Psychometric Artificial Intelligence', *Journal of Experimental and Theoretical Artificial Intelligence* **23**(3), 271–277.

Bringsjord, S. (2015), 'A Vindication of Program Verification', *History and Philosophy of Logic* **36**(3), 262–277. This url goes to a preprint.
URL: *http://kryten.mm.rpi.edu/SB_progver_selfref_driver_final2_060215.pdf*

Bringsjord, S. & Arkoudas, K. (2004), 'The Modal Argument for Hypercomputing Minds', *Theoretical Computer Science* **317**, 167–190.

Bringsjord, S., Bello, P. & Govindarajulu, N. (2018), Toward Axiomatizing Consciousness, *in* D. Jacquette, ed., 'The Bloomsbury Companion to the Philosophy of Consciousness', Bloomsbury Academic, London, UK, pp. 289–324.

Bringsjord, S. & Ferrucci, D. (2000), *Artificial Intelligence and Literary Creativity: Inside the Mind of Brutus, a Storytelling Machine*, Lawrence Erlbaum, Mahwah, NJ.

Bringsjord, S., Giancola, M. & Govindarajulu, N. S. ((forthcoming)), Logic-Based Modeling of Cognition, *in* R. Sun, ed., 'The Handbook of Computational Psychology', Cambridge University Press, Cambridge, UK.
URL: *http://kryten.mm.rpi.edu/Logic-basedComputationalModelingOfCognition.pdf*

Bringsjord, S. & Govindarajulu, N. (2020), 'The Theory of Cognitive Consciousness, and Λ (Lambda)', *Journal of Artificial Intelligence and Consciousness* **7**(1), 155–181. The URL here goes to a preprint of the paper.
URL:
http://kryten.mm.rpi.edu/sb_nsg_lambda_jaic_april_6_2020_3_42_pm_NY.pdf

Bringsjord, S., Govindarajulu, N. & Giancola, M. (2021), 'Automated Argument Adjudication to Solve Ethical Problems in Multi-Agent Environments', *Paladyn, Journal of Behavioral Robotics* **12**, 310–335. The URL here goes to a *rough*, *uncorrected*, *truncated* preprint as of 071421.
URL: *http://kryten.mm.rpi.edu/AutomatedArgumentAdjudicationPaladyn071421.pdf*

Bringsjord, S. & Govindarajulu, N. S. (2018), Artificial Intelligence, *in* E. Zalta, ed., 'The Stanford Encyclopedia of Philosophy'.
URL: *https://plato.stanford.edu/entries/artificial-intelligence*

Bringsjord, S., Govindarajulu, N. S. & C., E. (2019), Logicist Computational Cognitive Modeling of Infinitary False-Belief Tasks, *in* A. Goel, C. Seifert & C. Freksa, eds, 'Proceedings of the 41st Annual Conference of the Cognitive Science Society', Cognitive Science Society, Montreal, QB, pp. 43–45.

Bringsjord, S., Govindarajulu, N. S., Licato, J. & Giancola, M. (2020), Learning *Ex Nihilo*, *in* 'GCAI 2020. 6th Global Conference on Artificial Intelligence', Vol. 72 of *EPiC Series in Computing*, International Conferences on Logic and Artificial Intelligence at Zhejiang University (ZJULogAI), EasyChair Ltd, Manchester, UK, pp. 1–27.
URL: *https://easychair.org/publications/paper/NzWG*

Bringsjord, S., Kellett, O., Shilliday, A., Taylor, J., van Heuveln, B., Yang, Y., Baumes, J. & Ross, K. (2006), 'A New Gödelian Argument for Hypercomputing Minds Based on the Busy Beaver Problem', *Applied Mathematics and Computation* **176**, 516–530.

Bringsjord, S. & Licato, J. (2012), Psychometric Artificial General Intelligence: The Piaget-MacGuyver Room, *in* P. Wang & B. Goertzel, eds, 'Foundations of Artificial General Intelligence', Atlantis Press, Amsterdam, The Netherlands, pp. 25–47. This url is to a preprint only.
URL: *http://kryten.mm.rpi.edu/Bringsjord_Licato_PAGI_071512.pdf*

Bringsjord, S., Noel, R. & Ferrucci, D. (2002), Why Did Evolution Engineer Consciousness?, *in* J. Fetzer & G. Mulhauser, eds, 'Evolving Consciousness', Benjamin Cummings, San Francisco, CA, pp. 111–138.

Bringsjord, S. & Schimanski, B. (2003), What is Artificial Intelligence? Psychometric AI as an Answer, *in* 'Proceedings of the 18^{th} International Joint Conference on Artificial Intelligence (IJCAI–03)', Morgan Kaufmann, San Francisco, CA, pp. 887–893.
URL: *http://kryten.mm.rpi.edu/scb.bs.pai.ijcai03.pdf*

Bringsjord, S. & Zenzen, M. (2003), *Superminds: People Harness Hypercomputation, and More*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Chapin, N., Szymanski, B., Bringsjord, S. & Schimanski, B. (2011), 'A Bottom-Up Complement to the Logic-Based Top-Down Approach to the Story Arrangement Test', *Journal of Experimental and Theoretical Artificial Intelligence* **23**(3), 329–341.

Chella, A., Cangelosi, A., Metta, G. & Bringsjord, S., eds (2019), *Consciousness in Humanoid Robots*, Frontiers, Lausanne, Switzerland. ISSN 1664-8714; DOI 10.3389/978-2-88945-866-0.

Chella, A., Frixione, M. & Gaglio, S. (2000), 'Understanding Dynamic Scenes', *Artificial Intelligence* **123**, 89–132.

Chella, A. & Manzotti, R., eds (2007), *Artificial Consciousness*, Imprint Academic, Exeter, UK.

Davis, M., Sigal, R. & Weyuker, E. (1994), *Computability, Complexity, and Languages: Fundamentals of Theoretical Computer Science*, Academic Press, New York, NY. This is the second edition, which added Sigal as a co-author.

Dean, W. (2019), 'Computational Complexity Theory and the Philosophy of Mathematics', *Philosophia Mathematica* **27**(3), 381–439.

Dickmann, M. A. (1975), *Large Infinitary Languages*, North-Holland, Amsterdam, The Netherlands.

Duch, W., Oentaryo, R. J. & Pasquier, M. (2008), Cognitive Architectures: Where do we go from here?, *in* 'AGI', pp. 122–136.

Ebbinghaus, H. D., Flum, J. & Thomas, W. (1994), *Mathematical Logic (second edition)*, Springer-Verlag, New York, NY.

Fitting, M. (2015), Intensional Logic, *in* E. Zalta, ed., 'The Stanford Encyclopedia of Philosophy'.
URL: *https://plato.stanford.edu/entries/logic-intensional*

Gamez, D. (2020), 'The Relationships Between Intelligence and Consciousness in Natural and Artificial Systems', *Journal of Artificial Intelligence and Consciousness* **7**(1), 51–62.

Gardner, M. (1958), *Logic Machines and Diagrams*, McGraw-Hill, New York, NY.

Goertzel, B. (2014), 'Artificial General Intelligence: Concept, State of the Art, and Future Prospects', *Journal of Artificial General Intelligence* **0**, 1–48.

Goertzel, B. (2015), 'Artificial General Intelligence', *Scholarpedia* **10**(11), 31847. revision #154015.

Goertzel, B. (2021), 'The General Theory of General Intelligence: A Pragmatic Patternist Perspective', *ArXiv* **abs/2103.15100**, 1–64.

Govindarajulu, N. & Bringsjord, S. (2017), On Automating the Doctrine of Double Effect, *in* C. Sierra, ed., 'Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)', International Joint Conferences on Artificial Intelligence, pp. 4722–4730.
URL: *https://doi.org/10.24963/ijcai.2017/658*

Govindarajulu, N. S., Bringsjord, S., Ghosh, R. & Sarathy, V. (2019), Toward the Engineering of Virtuous Machines, *in* V. Conitzer, G. Hadfield & S. Vallor, eds, 'Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (AIES 2019)', ACM, New York, NY, pp. 29–35.

Gust, H., Krumnack, U., Schwering, A. & Kühnberger, K.-U. (2009), 'The Role of Logic in AGI Systems: Towards a Lingua Franca for General Intelligence', pp. 126–131.

Hamkins, J. D. & Lewis, A. (2000), 'Infinite Time Turing Machines', *Journal of Symbolic Logic* **65**(2), 567–604.

Hutter, M. (2005), *Universal Artificial Intelligence: Sequential Decisions Based on Algorithmic Probability*, Springer, New York, NY.

Jacob, P. (2003/2019), Intentionality, *in* E. Zalta, ed., 'The Stanford Encyclopedia of Philosophy'.
URL: *https://plato.stanford.edu/entries/intentionality*

Johnson, G. (2016), *Argument & Inference: An Introduction to Inductive Logic*, MIT Press, Cambridge, MA.

Klenk, M., Forbus, K., Tomai, E. & Kim, H. (2011), 'Using Analogical Model Formulation with Sketches to Solve Bennett Mechanical Comprehension Test Problems', *Journal of Experimental and Theoretical Artificial Intelligence* **23**(3), 299–327.

Legg, S. & Hutter, M. (2007), 'Universal intelligence: A Definition of Machine Intelligence', *Minds and Machines* **17**(4), 391–444.

Lovett, A., Bridewell, W. & Bello, P. (2021), Selection, Engagement, & Enhancement: A Framework for Modeling Visual Attention, *in* 'Proceedings of the 43rd Annual Conference of the Cognitive Science Society', Cognitive Science Society, Vienna, Austria, pp. 1893–1899.

McShane, M. & Nirenburg, S. (2021), *Linguistics for the Age of AI*, MIT Press, Cambridge, MA.

Mycka, J. & Costa, J. F. (2007), 'A New Conceptual Framework for Analog Computation', *Theoretical Computer Science* **374**, 277–290.

Newell, A. & Simon, H. (1956), 'The Logic Theory Machine: A Complex Information Processing System', *P-868 The RAND Corporation* pp. 25–63. An almost exactly similar version of this paper can be found in *IRE Transactions on Information Theory*, vol **2**, pages 61–79.

Oizumi, M., Albantakis, L. & Tononi, G. (2014), 'From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0', *Computational Biology* **5**(10), 1–25.

Ortony, A., Clore, G. L. & Collins, A. (1988), *The Cognitive Structure of Emotions*, Cambridge University Press, Cambridge, UK.

Paleo, B. W. (2016), Leibniz's Characteristica Universalis and Calculus Ratiocinator Today, *in* C. Tandy, ed., 'Death And Anti-Death, Volume 14: Four Decades After Michael Polanyi, Three Centuries After G. W. Leibniz', Ria University Press, pp. 313–332.

Paris, J. & Vencovská, A. (2015), *Pure Inductive Logic*, Cambridge University Press, Cambridge, UK.

Rodriguez, A. & Granger, R. (2016), 'The Grammar of Mammalian Brain Capacity', *Theoretical Computer Science* **633**, 100–111.

Russell, S. & Norvig, P. (2020), *Artificial Intelligence: A Modern Approach*, Pearson, New York, NY. Fourth edition.

Shapiro, S. & Bona, J. (2010), 'The GLAIR Cognitive Architecture', *International Journal of Machine Consciousness* **02**, 144–152.

Siegelmann, H. T. (1999), *Neural Networks and Analog Computation: Beyond the Turing Limit*, Birkhäuser, Boston, MA.

Simpson, S. (2010), *Subsystems of Second Order Arithmetic*, Cambridge University Press, Cambridge, UK. This is the 2nd edition.

Smith, P. (2013), *An Introduction to Gödel's Theorems*, Cambridge University Press, Cambridge, UK. This is the second edition of the book.

Solomonoff, R. (1978), 'Complexity-based Induction Systems: Comparisons and Convergence Theorems', *IEEE Transactions on Information Theory* **24**(4), 422–432.

Sutton, R. & Barto, A. (1998), *Reinforcement Learning*, MIT Press.

Taylor, A., Hunt, G., Medina, F. & Gray, R. (2008), 'Do New Caledonian Crows

Solve Physical Problems Through Causal Reasoning?', *Proceedings. Biological sciences / The Royal Society* **276**, 247–54.

Tononi, G. (2012), *Phi: A Voyage from the Brain to the Soul*, Pantheon, New York, NY.

Visalberghi, E. & Limongelli, L. (1994), 'Lack of Comprehension of Cause-Effect Relations in Tool-Using Capuchin Monkeys (*Cebus apella*)', *Journal of Comparative Psychology* **108**(1), 15–22.

Wang, P. (2013*a*), *Non-Axiomatic Logic, A Model of Intelligent Reasoning*, Sciendo.

Wang, P. (2013*b*), Proceedings of the 6th International Conference on Artificial General Intelligence, *in* 'Artificial General Intelligence', pp. 160–169.

Wang, P. (2019), 'On Defining Artificial Intelligence', *Journal of Artificial General Intelligence* **10**, 1 – 37.