AGI and the Knight-Darwin Law: why idealized AGI reproduction requires collaboration

Samuel Allen Alexander$^{1[0000-0002-7930-110X]}$

The U.S. Securities and Exchange Commission samuelallenalexander@gmail.com
https://philpeople.org/profiles/samuel-alexander/publications

Abstract. Can an AGI create a more intelligent AGI? Under idealized assumptions, for a certain theoretical type of intelligence, our answer is: "Not without outside help". This is a paper on the mathematical structure of AGI populations when parent AGIs create child AGIs. We argue that such populations satisfy a certain biological law. Motivated by observations of sexual reproduction in seemingly-asexual species, the Knight-Darwin Law states that it is impossible for one organism to asexually produce another, which asexually produces another, and so on forever: that any sequence of organisms (each one a child of the previous) must contain occasional multi-parent organisms, or must terminate. By proving that a certain measure (arguably an intelligence measure) decreases when an idealized parent AGI single-handedly creates a child AGI, we argue that a similar Law holds for AGIs.

Keywords: Intelligence Measurement · Knight-Darwin Law · Ordinal Notations · Intelligence Explosion

1 Introduction

It is difficult to reason about agents with Artificial General Intelligence (AGIs) programming AGIs[1]. To get our hands on something solid, we have attempted to find structures that abstractly capture the core essence of AGIs programming AGIs. This led us to discover what we call the *Intuitive Ordinal Notation System* (presented in Section 2), an ordinal notation system that gets directly at the heart of AGIs creating AGIs.

[1] Our approach to AGI is what Goertzel [11] describes as the Universalist Approach: we consider "...an idealized case of AGI, similar to assumptions like the frictionless plane in physics", with the hope that by understanding this "simplified special case, we can use the understanding we've gained to address more realistic cases."

We call an AGI *truthful* if the things it knows are true[2]. In [4], we argued that if a truthful AGI X creates (without external help) a truthful AGI Y in such a way that X knows the truthfulness of Y, then X must be more intelligent than Y in a certain formal sense. The argument is based on the key assumption that if X creates Y, without external help, then X necessarily knows Y's source code.

Iterating the above argument, suppose $X_1, X_2, \ldots$ are truthful AGIs such that each X_i creates, and knows the truthfulness and the code of, X_{i+1}. Assuming the previous paragraph, X_1 would be more intelligent than X_2, which would be more intelligent than X_3, and so on (in our certain formal sense). In Section 3 we will argue that this implies it is impossible for such a list $X_1, X_2, \ldots$ to go on forever: it would have to stop after finitely many elements[3].

At first glance, the above results might seem to suggest skepticism regarding the singularity—regarding what Hutter [15] calls *intelligence explosion*, the idea of AGIs creating better AGIs, which create even better AGIs, and so on. But there is a loophole (discussed further in Section 4). Suppose AGIs X and X' collaborate to create Y. Suppose X does part of the programming work, but keeps the code secret from X', and suppose X' does another part of the programming work, but keeps the code secret from X. Then neither X nor X' knows Y's full source code, and yet if X and X' trust each other, then both X and X' should be able to trust Y, so the above-mentioned argument breaks down.

Darwin and his contemporaries observed that even seemingly asexual plant species occasionally reproduce sexually. For example, a plant in which pollen is ordinarily isolated, might release pollen into the air if a storm damages the part of the plant that would otherwise shield the pollen[4]. The Knight-Darwin Law [8], named after Charles Darwin and Andrew Knight, is the principle (rephrased in modern language) that there cannot be an infinite sequence $X_1, X_2, \ldots$ of biological organisms such that each X_i asexually parents X_{i+1}. In other words, if $X_1, X_2, \ldots$ is any infinite list of organisms such that each X_i is a biological parent of X_{i+1}, then some of the X_i would need to be multi-parent organisms. The reader will immediately notice a striking parallel between this principle and the discussion in the previous two paragraphs.

In Section 2 we present the Intuitive Ordinal Notation System.

[2] Knowledge and truth are formally treated in [4] but here we aim at a more general audience. For the purposes of this paper, an AGI can be thought of as knowing a fact if and only if the AGI would list that fact if commanded to spend eternity listing all the facts that it knows. We assume such knowledge is closed under deduction, an assumption which is ubiquitous in modal logic, where it often appears in a form like $K(\phi \rightarrow \psi) \rightarrow (K(\phi) \rightarrow K(\psi))$. Of course, it is only in the idealized context of this paper that one should assume AGIs satisfy such closure.

[3] This may initially seem to contradict some mathematical constructions [18] [22] of infinite descending chains of theories. But those constructions only work for weaker languages, making them inapplicable to AGIs which comprehend linguistically strong second-order predicates.

[4] Even prokaryotes can be considered to occasionally have multiple parents, if lateral gene transfer is taken into account.

In Section 3 we argue[5] that if truthful AGI X creates truthful AGI Y, such that X knows the code and truthfulness of Y, then, in a certain formal sense, Y is less intelligent than X.

In Section 4 we adapt the Knight-Darwin Law from biology to AGI and speculate about what it might mean for AGI.

In Section 5 we address some anticipated objections.

Sections 2–3 are not new (except for new motivation and discussion). Their content appeared in [4], and was more rigorously formalized there. Sections 4–5 contain this paper's new material. Of this, some was hinted at in [4], and some appeared (weaker and less approachably) in the author's dissertation [2].

2 The Intuitive Ordinal Notation System

If humans can write AGIs, and AGIs are at least as smart as humans, then AGIs should be capable of writing AGIs. Based on the conviction that an AGI should be capable of writing AGIs, we would like to come up with a more concrete structure, easier to reason about, which we can use to better understand AGIs.

To capture the essence of an AGI's AGI-programming capability, one might try: "computer program that prints computer programs." But this only captures the AGI's capability to write *computer programs*, not to write *AGIs*.

How about: "computer program that prints computer programs that print computer programs"? This second attempt seems to capture an AGI's ability to write *program-writing programs*, not to write *AGIs*.

Likewise, "computer program that prints computer programs that print computer programs that print computer programs" captures the ability to write *program-writing-program-writing programs*, not *AGIs*.

We need to short-circuit the above process. We need to come up with a notion X which is equivalent to "computer program that prints members of X".

Definition 1 *(See the following examples) We define the Intuitive Ordinal Notations to be the smallest set $\mathcal{P}$ of computer programs such that:*

- *Each computer program p is in $\mathcal{P}$ iff all of p's outputs are also in $\mathcal{P}$.*

Example 2 *(Some simple examples)*

1. *Let P_0 be "End", a program which immediately stops without any outputs. Vacuously, all of P_0's outputs are in $\mathcal{P}$ (there are no such outputs). So P_0 is an Intuitive Ordinal Notation.*
2. *Let P_1 be "Print('End')", a program which outputs "End" and then stops. By (1), all of P_1's outputs are Intuitive Ordinal Notations, therefore, so is P_1.*
3. *Let P_2 be "Print('Print('End')')", which outputs "Print('End')" and then stops. By (2), all of P_2's outputs are Intuitive Ordinal Notations, therefore, so is P_2.*

[5] This argument appeared in a fully rigorous form in [4], but in this paper we attempt to make it more approachable.

Example 3 *(A more interesting example) Let P_ω be the program:*

Let X = 'End'; While(True) { Print(X); X = "Print('" + X + "')"; }

When executed, P_ω outputs "End", "Print('End')", "Print('Print('End')')", and so on forever. As in Example 2, all of these are Intuitive Ordinal Notations. Therefore, P_ω is an Intuitive Ordinal Notation.

To make Definition 1 fully rigorous, one would need to work in a formal model of computation; see [4] (Section 3) where we do exactly that. Examples 2 and 3 are reminiscent of Franz's approach of "head[ing] for general algorithms at low complexity levels and fill[ing] the task cup from the bottom up" [9]. For a much larger collection of examples, see [3]. A different type of example will be sketched in the proof of Theorem 7 below.

Definition 4 *For any Intuitive Ordinal Notation x, we define an ordinal $|x|$ inductively as follows: $|x|$ is the smallest ordinal α such that $\alpha > |y|$ for every output y of x.*

Example 5

- *Since P_0 (from Example 2) has no outputs, it follows that $|P_0| = 0$, the smallest ordinal.*
- *Likewise, $|P_1| = 1$ and $|P_2| = 2$.*
- *Likewise, P_ω (from Example 3) has outputs notating $0, 1, 2, \ldots$—all the finite natural numbers. It follows that $|P_\omega| = \omega$, the smallest infinite ordinal.*
- *Let $P_{\omega+1}$ be the program "Print(P_ω)", where P_ω is as in Example 3. It follows that $|P_{\omega+1}| = \omega + 1$, the next ordinal after ω.*

The Intuitive Ordinal Notation System is a more intuitive simplification of an ordinal notation system known as Kleene's $\mathcal{O}$.

3 Intuitive Ordinal Intelligence

Whatever an AGI is, an AGI should know certain mathematical facts. The following is a universal notion of an AGI's intelligence based solely on said facts. In [4] we argue that this notion captures key components of intelligence such as pattern recognition, creativity, and the ability or generalize. We will give further justification in Section 5. Even if the reader refuses to accept this as a genuine intelligence measure, that is merely a name we have chosen for it: we could give it any other name without compromising this paper's structural results.

Definition 6 *The* Intuitive Ordinal Intelligence *of a truthful AGI X is the smallest ordinal $|X|$ such that $|X| > |p|$ for every Intuitive Ordinal Notation p such that X knows that p is an Intuitive Ordinal Notation.*

The following theorem provides a relationship[6] between Intuitive Ordinal Intelligence and AGI creation of AGI. Here, we give an informal version of the proof; for a version spelled out in complete formal detail, see [4].

[6] Possibly formalizing a relationship implied offhandedly by Chaitin, who suggests ordinal computation as a mathematical challenge intended to encourage evolution, "and the larger the ordinal, the fitter the organism" [7].

Theorem 7 *Suppose X is a truthful AGI, and X creates a truthful AGI Y in such a way that X knows Y's code and truthfulness. Then $|X| > |Y|$.*

Proof. Suppose Y were commanded to spend eternity enumerating the biggest Intuitive Ordinal Notations Y could think of. This would result in some list L of Intuitive Ordinal Notations enumerated by Y. Since Y is an AGI, L must be computable. Thus, there is some computer program P whose outputs are exactly L. Since X knows Y's code, and as an AGI, X is capable of reasoning about code, it follows that X can infer a program P that[7] lists L. Having constructed P this way, X knows: "P outputs L, the list of things Y would output if Y were commanded to spend eternity trying to enumerate large Intuitive Ordinal Notations". Since X knows Y is truthful, X knows that L contains nothing except Intuitive Ordinal Notations, thus X knows that P's outputs are Intuitive Ordinal Notations, and so X knows that P is an Intuitive Ordinal Notation. So $|X| > |P|$. But $|P|$ is the least ordinal $> |Q|$ for all Q output by L, in other words, $|P| = |Y|$. □

Theorem 7 is mainly intended for the situation where parent X creates independent child Y, but can also be applied in case X self-modifies, viewing the original X as being replaced by the new self-modified Y (assuming X has prior knowledge of the code and truthfulness of the modified result).

It would be straightforward to extend Theorem 7 to cases where X creates Y non-deterministically. Suppose X creates Y using random numbers, such that X knows Y is one of $Y_1, Y_2, \ldots, Y_k$ but X does not know which. If X knows that Y is truthful, then X must know that each Y_i is truthful (otherwise, if some Y_i were not truthful, X could not rule out that Y was that non-truthful Y_i). So by Theorem 7, each $|Y_i|$ would be $< |X|$. Since Y is one of the Y_i, we would still have $|Y| < |X|$.

4 The Knight-Darwin Law

> "...it is a general law of nature that no organic being self-fertilises itself for a perpetuity of generations; but that a cross with another individual is occasionally—perhaps at long intervals of time—indispensable." (Charles Darwin)

In his Origin of Species, Darwin devotes many pages to the above-quoted principle, later called the Knight-Darwin Law [8]. In [1] we translate the Knight-Darwin Law into mathematical language.

[7] For example, X could write a general program $Sim(c)$ that simulates an input AGI c waking up in an empty room and being commanded to spend eternity enumerating Intuitive Ordinal Notations. This program $Sim(c)$ would then output whatever outputs AGI c outputs under those circumstances. Having written $Sim(c)$, X could then obtain P by pasting Y's code into Sim (a string operation—not actually running Sim on Y's code). Nowhere in this process do we require X to actually execute Sim (which might be computationally infeasible).

Principle 8 *(The Knight-Darwin Law) There cannot be an infinite sequence $x_1, x_2, \ldots$ of organisms such that each x_i is the lone biological parent of x_{i+1}. If each x_i is a parent of x_{i+1}, then some x_{i+1} must have multiple parents.*

A key fact about the ordinals is they are *well-founded*: there is no infinite sequence $o_1, o_2, \ldots$ of ordinals such that[8] each $o_i > o_{i+1}$. In Theorem 7 we showed that if truthful AGI X creates truthful AGI Y in such a way as to know the truthfulness and code of Y, then X has a higher Intuitive Ordinal Intelligence than Y. Combining this with the well-foundedness of the ordinals yields a theorem extremely similar to the Knight-Darwin Law.

Theorem 9 *(The Knight-Darwin Law for AGIs) There cannot be an infinite sequence $X_1, X_2, \ldots$ of truthful AGIs such that each X_i creates X_{i+1} in such a way as to know X_{i+1}'s truthfulness and code. If each X_i creates X_{i+1} so as to know X_{i+1} is truthful, then occasionally certain X_{i+1}'s must be co-created by multiple creators (assuming that creation by a lone creator implies the lone creator would know X_{i+1}'s code).*

Proof. By Theorem 7, the Intuitive Ordinal Intelligence of $X_1, X_2, \ldots$ would be an infinite strictly-descending sequence of ordinals, violating the well-foundedness of the ordinals. □

It is perfectly consistent with Theorem 7 that Y might operate faster than X, performing better in realtime environments (as in [10]). It may even be that Y performs so much faster that it would be infeasible for X to use the knowledge of Y's code to simulate Y. Theorems 7 and 9 are profound because they suggest that descendants might initially appear more practical (faster, better at problem-solving, etc.), yet, without outside help, their knowledge must degenerate. This parallels the *hydra game* of Kirby and Paris [16], where a hydra seems to grow as the player cuts off its heads, yet inevitably dies if the player keeps cutting.

If AGI Y has distinct parents X and X', neither of which fully knows Y's code, then Theorem 7 does not apply to X, Y or X', Y and does not force $|Y| < |X|$ or $|Y| < |X'|$. This does not necessarily mean that $|Y|$ can be arbitrarily large, though. If X and X' were themselves created single-handedly by a lone parent X_0, similar reasoning to Theorem 7 would force $|Y| < |X_0|$ (assuming X_0 could infer the code and truthfulness of Y from those of X and X')[9].

In the remainder of this section, we will non-rigorously speculate about three implications Theorem 9 might have for AGIs and for AGI research.

[8] This is essentially true by definition, unfortunately the formal definition of ordinal numbers is outside the scope of this paper.

[9] This suggests possible generalizations of the Knight-Darwin Law such as "There cannot be an infinite sequence $x_1, x_2, \ldots$ of biological organisms such that each x_i is the lone grandparent of x_{i+1}," and AGI versions of same. This also raises questions about the relationship between the set of AGIs initially created by humans and how intelligent the offspring of those initial AGIs can be. These questions go beyond the scope of this paper but perhaps they could be a fruitful area for future research.

4.1 Motivation for Multi-agent Approaches to AGI

If AGI ought to be capable of programming AGI, Theorem 9 suggests that a fundamental aspect of AGI should be the ability to collaborate with other AGIs in the creation of new AGIs. This seems to suggest there should be no such thing as a *solipsistic* AGI[10], or at least, solipsistic AGIs would be limited in their reproduction ability. For, if an AGI were solipsistic, it seems like it would be difficult for this AGI to collaborate with other AGIs to create child AGIs. To quote Hernández-Orallo et al: "The appearance of multi-agent systems is a sign that the future of machine intelligence will not be found in monolithic systems solving tasks without other agents to compete or collaborate with" [12].

More practically, Theorem 9 might suggest prioritizing research on multi-agent approaches to AGI, such as [6], [12], [14], [17], [19], [21], and similar work.

4.2 Motivation for AGI Variety

Darwin used the Knight-Darwin Law as a foundation for a broader thesis that the survival of a species depends on the inter-breeding of many members. By analogy, if our goal is to create robust AGIs, perhaps we should focus on creating a wide variety of AGIs, so that those AGIs can co-create more AGIs.

On the other hand, if we want to reduce the danger of AGI getting out of control, perhaps we should *limit* AGI variety. At the extreme end of the spectrum, if humankind were to limit itself to only creating one single AGI[11], then Theorem 9 would constrain the extent to which that AGI could reproduce.

4.3 AGI Genetics

If AGI collaboration is a fundamental requirement for AGI "populations" to propagate, it might someday be possible to view AGI through a genetic lens. For example, if AGIs X and X' co-create child Y, if X runs operating system O, and X' runs operating system O', perhaps Y will somehow exhibit traces of both O and O'.

5 Discussion

In this section, we discuss some anticipated objections.

5.1 What does Definition 6 really have to do with intelligence?

We do not claim that Definition 6 is the "one true measure" of intelligence. Maybe there is no such thing: maybe intelligence is inherently multi-dimensional.

[10] That is, an AGI which believes itself to be the only entity in the universe.

[11] Or to perfectly isolate different AGIs away from one another—see [25].

Definition 6 measures a type of intelligence based on mathematical knowledge[12] closed under logical deduction. An AGI could be good at problem-solving but poor at ordinals. But the broad AGIs we are talking about in this paper should be capable (if properly instructed) of attempting any reasonable well-defined task, including that of notating ordinals. So Definition 6 does measure one aspect of an AGI's abilities. Perhaps a word like "mathematical-knowledge-level" would fit better: but that would not change the Knight-Darwin Law implications.

Intelligence has core components like pattern-matching, creativity, and the ability to generalize. We claim that these components are needed if one wants to competitively name large ordinals. If p is an Intuitive Ordinal Notation obtained using certain facts and techniques, then *any* AGI who used those facts and techniques to construct p should also be able to iterate those same facts and techniques. Thus, to advance from p to a larger ordinal which not just *any* p-knowing AGI could obtain, must require the creative invention of some new facts or techniques, and this invention requires some amount of creativity, pattern-matching, etc. This becomes clear if the reader tries to notate ordinals qualitatively larger than Example 3; see the more extensive examples in [3].

For analogy's sake, imagine a ladder which different AGIs can climb, and suppose advancing up the ladder requires exercising intelligence. One way to measure (or at least estimate) intelligence would be to measure how high an AGI can climb said ladder.

Not all ladders are equally good. A ladder would be particularly poor if it had a top rung which many AGIs could reach: for then it would fail to distinguish between AGIs who could reach that top rung, even if one AGI reaches it with ease and another with difficulty. Even if the ladder was infinite and had no top rung, it would still be suboptimal if there were AGIs capable of scaling the whole ladder (i.e., of ascending however high they like, on demand)[13]. A good ladder should have, for each particular AGI, a rung which that AGI cannot reach.

Definition 6 offers a good ladder. The rungs which an AGI manages to reach, we have argued, require core components of intelligence to reach. And no particular AGI can scale the whole ladder[14], because no AGI can enumerate all

[12] Wang has correctly pointed out [23] that an AGI consists of much more than merely a knowledge-set of mathematical facts. Still, we feel mathematical knowledge is at least one important aspect of an AGI's intelligence.

[13] Hibbard's intelligence measure [13] is an infinite ladder which is nevertheless short enough that many AGIs can scale the whole ladder—the AGIs which do not "have finite intelligence" in Hibbard's words (see Hibbard's Proposition 3). It should be possible to use a *fast-growing hierarchy* [24] to transfinitely extend Hibbard's ladder and reduce the set of whole-ladder-scalers. This would make Hibbard's measurement ordinal-valued (perhaps Hibbard intuited this; his abstract uses the word "ordinal" in its everyday sense as synonym for "natural number").

[14] Thus, this ladder avoids a common problem that arises when trying to measure machine intelligence using IQ tests, namely, that for any IQ test, an algorithm can be designed to dominate that test, despite being otherwise unintelligent [5].

the Intuitive Ordinal Notations: it can be shown that they are not computably enumerable[15].

5.2 Can't an AGI just print a copy of itself?

If a truthful AGI knows its own code, then it can certainly print a copy of itself. But if so, then it necessarily cannot know the truthfulness of that copy, lest it would know the truthfulness of itself. Versions of Gödel's incompleteness theorems adapted [20] to mechanical knowing agents imply that a suitably idealized truthful AGI cannot know its own code and its own truthfulness.

5.3 Prohibitively expensive simulation

The reader might object that Theorem 7 breaks down if Y is prohibitively expensive for X to simulate. But Theorem 7 and its proof have nothing to do with simulation. In functional languages like Haskell, functions can be manipulated, filtered, formally composed with other functions, and so on, without needing to be executed. Likewise, if X knows the code of Y, then X can manipulate and reason about that code without executing a single line of it.

6 Conclusion

The Intuitive Ordinal Intelligence of a truthful AGI is defined to be the supremum of the ordinals which have Intuitive Ordinal Notations the AGI knows to be Intuitive Ordinal Notations. We argued that this notion measures (a type of) intelligence. We proved that if a truthful AGI single-handedly creates a child truthful AGI, in such a way as to know the child's truthfulness and code, then the parent must have greater Intuitive Ordinal Intelligent than the child. This allowed us to establish a structural property for AGI populations, resembling the Knight-Darwin Law from biology. We speculated about implications of this biology-AGI parallel. We hope by better understanding how AGIs create new AGIs, we can better understand methods of AGI-creation by humans.

Acknowledgments

We gratefully acknowledge Jordi Bieger, Thomas Forster, José Hernández-Orallo, Bill Hibbard, Mike Steel, Albert Visser, and the reviewers for discussion and feedback.

[15] Namely, because if the set of Intuitive Ordinal Notations were computably enumerable, the program p which enumerates them would itself be an Intuitive Ordinal Notation, which would force $|p| > |p|$.

References

1. Alexander, S.A.: Infinite graphs in systematic biology, with an application to the species problem. Acta Biotheoretica **61**, 181–201 (2013)
2. Alexander, S.A.: The theory of several knowing machines. Ph.D. thesis, The Ohio State University (2013)
3. Alexander, S.A.: Intuitive ordinal notations (IONs). GitHub repository, https://github.com/semitrivial/ions (2019)
4. Alexander, S.A.: Measuring the intelligence of an idealized mechanical knowing agent. In: CIFMA (2019)
5. Besold, T., Hernández-Orallo, J., Schmid, U.: Can machine intelligence be measured in the same way as human intelligence? KI-Künstliche Intelligenz **29**, 291–297 (2015)
6. Castelfranchi, C.: Modelling social action for AI agents. AI **103**, 157–182 (1998)
7. Chaitin, G.: Metaphysics, metamathematics and metabiology. In: Hector, Z. (ed.) Randomness through computation. World Scientific (2011)
8. Darwin, F.: The Knight-Darwin Law. Nature **58**, 630–632 (1898)
9. Franz, A.: Toward tractable universal induction through recursive program learning. In: ICAGI. pp. 251–260 (2015)
10. Gavane, V.: A measure of real-time intelligence. JAGI **4**, 31–48 (2013)
11. Goertzel, B.: Artificial general intelligence: concept, state of the art, and future prospects. JAGI **5**, 1–48 (2014)
12. Hernández-Orallo, J., Dowe, D.L., España-Cubillo, S., Hernández-Lloreda, M.V., Insa-Cabrera, J.: On more realistic environment distributions for defining, evaluating and developing intelligence. In: ICAGI. pp. 82–91 (2011)
13. Hibbard, B.: Measuring agent intelligence via hierarchies of environments. In: ICAGI. pp. 303–308 (2011)
14. Hibbard, B.: Societies of intelligent agents. In: ICAGI. pp. 286–290 (2011)
15. Hutter, M.: Can intelligence explode? JCS **19**, 143–166 (2012)
16. Kirby, L., Paris, J.: Accessible independence results for Peano arithmetic. Bulletin of the London Mathematical Society **14**, 285–293 (1982)
17. Kolonin, A., Goertzel, B., Duong, D., Ikle, M.: A reputation system for artificial societies. arXiv preprint arXiv:1806.07342 (2018)
18. Kripke, S.A.: Ungroundedness in Tarskian languages. JPL **48**, 603–609 (2019)
19. Potyka, N., Acar, E., Thimm, M., Stuckenschmidt, H.: Group decision making via probabilistic belief merging. In: 25th IJCAI. AAAI Press (2016)
20. Reinhardt, W.N.: Absolute versions of incompleteness theorems. Nous **19**, 317–346 (1985)
21. Thórisson, K.R., Benko, H., Abramov, D., Arnold, A., Maskey, S., Vaseekaran, A.: Constructionist design methodology for interactive intelligences. AI Magazine **25**, 77–90 (2004)
22. Visser, A.: Semantics and the liar paradox. In: Handbook of philosophical logic, pp. 149–240. Springer (2002)
23. Wang, P.: Three fundamental misconceptions of artificial intelligence. Journal of Experimental & Theoretical Artificial Intelligence **19**, 249–268 (2007)
24. Weiermann, A.: Slow versus fast growing. Synthese **133**, 13–29 (2002)
25. Yampolskiy, R.V.: Leakproofing singularity-artificial intelligence confinement problem. JCS **19** (2012)