Why traditional reinforcement learning will not lead to AGI: Archimedean measurements considered poor for non-Archimedean structures

Samuel A. Alexander*†
The U.S. Securities and Exchange Commission

2020

Abstract

After generalizing the Archimedean property of real numbers in such a way as to make it adaptable to non-numeric structures, we demonstrate that the real numbers cannot be used to accurately measure non-Archimedean structures. We argue that, since an agent with Artificial General Intelligence (AGI) should have no problem engaging in tasks that inherently involve non-Archimedean rewards, and since traditional reinforcement learning rewards are real numbers, therefore traditional reinforcement learning cannot lead to AGI. We indicate two possible ways traditional reinforcement learning could be altered to remove this roadblock.

1 Introduction

Whenever we measure anything using a particular number system, the corresponding measurements will be constrained by the structure of that number system. If the number system has a different structure than the things we are measuring with it, then our measurements will suffer accordingly, just as if we were trying to force square pegs into round holes.

For example, the natural numbers make lousy candidates for measuring distances in a physics laboratory. Distances in the lab have properties such as, for example, the fact that for any two distinct distances, there is an intermediate distance strictly between them. The natural numbers lack this property. Imagine the poor physicist, brought up in a world of only natural numbers, scratching his or her head upon encountering a rod with length strictly between two rods of length 1 and 2.

*Email: samuelallenalexander@gmail.com
†2010 Mathematics Subject Classification: 97R40 (Primary), 26E35 (Secondary)

It's tempting to think of the real numbers $\mathbb{R}$—i.e., the unique complete ordered field—as a generic number system with whatever structure suits our needs. But it's important to remember that the real numbers do have a very specific structure. That structure is flexible enough to accomodate many needs, but we shouldn't just take that for granted. One particular property of the real numbers is the following.

Lemma 1. (The Archimedean Property[1]) Let $r > 0$ be any positive real number. For every real number y, there is some natural number n such that $nr > y$.

Rather than directly prove Lemma 1, we will prove a generalized result more adaptable to other structures.

Lemma 2. (The Generalized Archimedean Property) Let $r > 0$ be any positive real number. For any $x, y \in \mathbb{R}$, say that y is *significantly greater* than x if $y \geq x + r$. If $x_0, x_1, x_2, \ldots$ is any infinite sequence of real numbers, where each x_{i+1} is significantly greater than x_i, then for every real number y, there exists some i such that x_i is significantly greater than y.

Proof. If not, then there is some y such that $y + r > x_i$ for all i. Thus, $X = \{x_0, x_1, x_2, \ldots\}$ has an upper bound. By the completeness of $\mathbb{R}$, X must have a least upper bound $z \in \mathbb{R}$. Since z is the least upper bound for X, $z - r$ is not an upper bound for X, so there is some i such that $x_i > z - r$. By assumption, $x_{i+1} > x_i + r$, so $x_{i+1} > z$, contradicting the choice of z. □

Lemma 1 follows from Lemma 2 by letting $x_i = ir$.

The above property is automatically inherited by subsystems of the reals, such as the rational numbers $\mathbb{Q}$, the natural numbers $\mathbb{N}$, the integers $\mathbb{Z}$, or the algebraic numbers. All inherit the Generalized Archimedean Property in obvious ways.

Lemma 2 allows us to adapt the notion of Archimedeanness to other things than real numbers, even to things for which there is no notion of multiplication

[1]The Archimedean property is named after Archimedes of Syracuse. A similar property appears as the fifth axiom in his *On the Sphere and Cylinder* [4]:

> Further, of unequal lines, unequal surfaces, and unequal solids, the greater exceeds the less by such a magnitude as, when added to itself, can be made to exceed any assigned magnitude among those which are comparable with [it and with] one another.

Note that Archimedes specifically restricts his statement to lengths, surface areas and volumes, in fact going out of his way to limit the magnitudes to which said length/area/volume can be made to exceed (he could have saved some ink by stopping his sentence at "...can be made to exceed any assigned magnitude", if that were his intention).

The Archimedean property is also closely related to Definition 4 of Book V of Euclid's *Elements* [8]:

> (Those) magnitudes are said to have a ratio with respect to one another which, being multiplied, are capable of exceeding one another.

Many math historians speak of the modern-day Archimedean property, Archimedes' 5th axiom, and Euclid's axiom as being identical, but in fact they are all subtly different from one another, see [5].

by natural numbers (Lemma 1 would not adapt to such things). All we need is a notion of "significantly greater than". For any set of things, some of which are "significantly greater than" others, we can ask whether or not the property in Lemma 2 holds. We will make this formal in Section 2.

Example 3. (Fuzzy widgets) Suppose we have some fuzzy widgets, and we observe that certain widgets are fuzzier than others. Naturally, we are inclined to quantify the fuzziness of the widgets, assigning them numerical fuzziness measures from some number system. Nine times out of ten, we choose to use the real numbers, or a subsystem thereof, often without a second thought. But suppose among these widgets, there happen to be widgets $w_1, w_2, \ldots$ such that each w_{i+1} is significantly fuzzier than w_i, and another widget w_∞ which is significantly fuzzier than all the w_i's. Suddenly, our decision to use real numbers puts us in a bind. It is impossible to assign real number fuzziness measures to our widgets in such a way that significantly fuzzier widgets get significantly greater real number measures. That would contradict Lemma 2.

Note that the above example does not require us to have any notion of multiplying fuzziness by a natural number n (as we would need to have if we wanted to adapt Lemma 1). This illustrates the enhanced adaptability of Lemma 2.

The structure of this paper is as follows.

- In Section 2 we formally adapt Lemma 2 to obtain a notion of Archimedeanness for non-numerical structures, and demonstrate that non-Archimedean such structures cannot accurately be measured using the real numbers.
- In Section 3 we argue that traditional reinforcement learning will not lead to AGI because its rewards are overly constrained.
- In Section 4 we discuss non-traditional variations of reinforcement learning that avoid the problem of overly constrained rewards.
- In Section 5 we summarize and make concluding remarks.

2 Generalized Archimedean Structures

The real numbers possess the Archimedean property, but other structures may or may not. To make this more precise, we introduce the following formalism, adapting from Lemma 2.

Definition 1. A *significantly-ordered structure* is a set X with an ordering $\ll$. For $x_1, x_2 \in X$, we say x_2 is *significantly greater* than x_1 if $x_1 \ll x_2$. A significantly-ordered structure is *Archimedean* if it has the following property: for every X-sequence $x_0 \ll x_1 \ll x_2 \ll \cdots$, for every $x_\infty \in X$, there is some i such that $x_\infty \ll x_i$.

For any real number $r > 0$, a prototypical example of an Archimedean significantly-ordered structure is the real numbers with $\ll$ defined such that $x \ll y$ if and only if $y \geq x + r$.

Definition 2. Suppose $(X, \ll)$ is a significantly-ordered structure. A function $f : X \to \mathbb{R}$ is said to *respect* $\ll$ if there is some real $r > 0$ such that the following requirement holds:

- For all $x, y \in X$, $x \ll y$ if and only if $f(y) \geq f(x) + r$.

The following proposition formalizes the dilemma we illustrated in Example 3 (think of X as a set of things we want to measure).

Proposition 4. (Inadequacy of the reals for non-Archimedean structures) Suppose $(X, \ll)$ is a significantly-ordered structure. If X is non-Archimedean, then no function $f : X \to \mathbb{R}$ respects $\ll$.

Proof. Assume, for sake of a contradiction, that some $f : X \to \mathbb{R}$ exists which respects $\ll$. Thus there is some real $r > 0$ such that for all $x, y \in X$, $x \ll y$ if and only if $f(y) \geq f(x) + r$. Since X is non-Archimedean, there is some X-sequence $x_0 \ll x_1 \ll x_2 \ll \cdots$ and some $x_\infty \in X$ such that there is no i such that $x_\infty \ll x_i$. By choice of r, each $f(x_{i+1}) \geq f(x_i) + r$ and there is no i such that $f(x_i) \geq f(x_\infty)$. This contradicts Lemma 2. $\square$

Proposition 4 tells us that we cannot faithfully measure non-Archimedean structures using real numbers. Any attempt to do so will necessarily be misleading, because ordering relationships among the non-Archimedean structures will fail to be reflected by the real-number measurements given to them. We will inevitably end up like the puzzled physicist brought up in a world of only natural numbers, confronted by a rod of length 1.5.

Example 5. (Examples of non-Archimedean structures)

- (Sets) Say that set U is significantly greater than set V if there is an injective function from V into U but there is no bijective function from V onto U. It is easy to show there are sets $U_0, U_1, \ldots$, with each U_{i+1} significantly greater than U_i, and $U_\infty = \cup_{i=0}^{\infty} U_i$ is a set which is significantly greater than all of them. Thus, sets are non-Archimedean. In the field of *set theory*, logicians measure the size of sets using Georg Cantor's famous non-Archimedean number system, the cardinal numbers.

- (Logical theories) It is not difficult to come up with (for example) true theories $T_0, T_1, \ldots$ (in the language of arithmetic) such that each T_{i+1} proves the consistency of T_i, and an additional true theory T_∞ (in the language of arithmetic) which proves the consistency of $\cup_{i=0}^{\infty} T_i$. In a sense, then, each T_{i+1} is significantly stronger than T_i (see Gödel's incompleteness theorems), and T_∞ is significantly stronger than all of them. In this sense, logical theories are non-Archimedean. In the field of *proof theory* [14], logicians measure the logical strength of theories not using real numbers but using computable ordinal numbers, another non-Archimedean number system.

- (Probability) The measurable subsets of $\mathbb{R}$ are non-Archimedean if we define $\ll$ such that for all measurable subsets X, Y of $\mathbb{R}$, $X \ll Y$ if and only if $X \subseteq Y$ and $|Y - X| = \infty$. Nevertheless, probability theorists routinely measure such sets using the real number system. The inadequacy expressed in Proposition 4 manifests itself in well-known paradoxes such as the fact that an event can have probability 0 and yet still be possible, or can have probabity 1 and yet not be guaranteed, or that an infinite event can have probability 0. See [6] for an alternative non-Archimedean approach to probability.

Example 6. (Speculative examples of potentially non-Archimedean structures) Certain structures might plausibly be non-Archimedean, but it is a difficult question to say whether they truly are or not. The reader could come up with such examples in great abundance.

- (Musical beauty) It is plausible that music might be non-Archimedean, in the following sense: there might be songs $S_0, S_1, \ldots$ such that each S_{i+1} is significantly more beautiful than S_i, and another song S_∞ which is significantly more beautiful than all the S_i's.
- (Literature) It is plausible that there might be stories $S_0, S_1, \ldots$ such that each S_{i+1} is significantly more entertaining than S_i, and another story S_∞ which is significantly more entertaining than all the S_i's.
- (Video game difficulty) It is plausible that there might be video games $V_0, V_1, \ldots$ such that each V_{i+1} is significantly more difficult than V_i, and another video game V_∞ which is significantly more difficult than all the V_i's.
- (AGI) It is plausible (and, in this author's opinion, very likely) that there are[2] AGIs $A_0, A_1, \ldots$ such that each A_{i+1} is significantly more intelligent than A_i, and another AGI A_∞ which is significantly more intelligent than all the A_i's. We first pointed this out in [2], where we propose measuring the intelligence of mechanical knowing agents using computable ordinals, the same non-Archimedean number system which proof theorists use to measure logical strength of mathematical theories. Incidentally, if AGI intelligence is non-Archimedean, then Proposition 4 shows it is impossible to measure machine intelligence using real numbers without some of those measurements being misleading[3].

[2] As hinted by Protagoras, assuming Protagoras's own intelligence stays constant and remains higher than the intelligence of his student: "The very day you start, you will go home a better man, and the same thing will happen the day after. Every day, day after day, you will get better and better." [13]

[3] This would solve an open problem implicitly stated by Legg and Hutter [11] when they said of their real-number universal intelligence measure: "...none of these people have been able to communicate why the work [on measuring universal intelligence using real numbers] is so obviously flawed in any concrete way ... If anyone would like to properly explain their position to us in the future, we promise not to chase you down the street!"

- (Nonstandard cosmologies) Some authors [1] [3] [15] [16] have even speculated about the nature of non-Archimedean space and/or time.

3 Reinforcement learning

In reinforcement learning (RL), an agent interacts with an environment, taking actions from a fixed set of possible actions. With every action the agent takes, the environment responds with a new observation and with a reward. In traditional RL, these rewards are real numbers (many authors further constrain them to be rational numbers).

By restricting rewards to be real (or rational) numbers, we unconsciously constrain RL to only be applicable toward tasks of an inherently Archimedean nature. For example, Wirth et al point out [19] that in tasks related to cancer treatment [20], "the death of a patient should be avoided at any cost. However, an infinitely negative reward breaks classic reinforcement learning algorithms and arbitrary, finite values have to be selected." This problem could be avoided if instead of real numbers, rewards were drawn from a suitable non-Archimedean number system containing negative infinites, for example, from the so-called surreal numbers (which we will discuss further on). Doing so would, however, be a departure from traditional RL.

To give an intuitive example, assume that musical beauty is non-Archimedean, as in Example 6. We can imagine environments where the RL agent is tasked with composing songs. For example, the possible actions the agent is allowed to rake might include one action for each piano key, plus an additional "stand and bow" action to signal that a song is finished. Whenever the agent stands and bows, the agent is awarded with applause based on the beauty of the song the agent composed[4]. Assuming musical beauty is non-Archimedean, such an environment falls outside the possibility of traditional RL. By Lemma 4, there is no way to assign real number rewards to songs without misleading the agent. If $S_0, S_1, S_2, \ldots$ are songs where each S_{i+1} is significantly more beautiful than S_i, and S_∞ is a song significantly more beautiful than all the S_i's, then there is no way to assign real-valued rewards to these songs such that each S_{i+1} gets significantly more reward than S_i and such that S_∞ gets significantly more reward than all the S_i's.

Or, to re-use the cancer example, assume there are certain bad procedures the robotic surgeon could take, each one significantly worse than the previous, but all still significantly better than killing the patient. There is no way to assign real-valued rewards to these actions, and to killing the patient, in such a way that each bad action gets punished significantly harsher than the previous, but still significantly more forgiving than the punishment for killing the patient.

The reader might object by challenging the non-Archimedeanness of music and of medical procedures. But we only used those to make the examples more

[4]To quote Wang and Hammer: "Decision makings often do not happen at the level of basic operations, but at the level of composed actions, where there are usually infinite possibilities." [18]

intuitive. If the reader insists, we can resort to mathematical tasks. For example, imagine that instead of playing piano, the agent is tasked with typing up mathematical theories, and when the agent stands and bows, the agent is rewarded with applause based on the proof-theoretical strength of the theory (or hit with tomatoes if the theory is inconsistent). In Example 5 we noted that proof-theoretical strength of theories is non-Archimedean. There exist theories $T_0, T_1, \ldots$, each significantly proof-theoretically stronger than the previous, and another theory T_∞, significantly proof-theoretically stronger than all the T_i's. We cannot possibly assign real-valued rewards to these theories without misleading the agent.

The reader might object to the above example because judging the proof-theoretical strength of a theory is inherently non-computable anyway. The example could be modified so that instead of typing up mathematical theories, the agent has to type up mathematical subtheories in (say) the language of Peano arithmetic, accompanied by consistency proofs in (say) ZFC. It can be shown that the proof-theoretical strength of mathematical theories is still non-Archimedean, even when restricted to subtheories of arithmetic whose consistency can be proven in ZFC.

The reader might object that the above theories-with-proofs example is contrived. But if an AGIs has human or superhuman intelligence, then even such contrived tasks should pose no problem for that AGI. When we prove that the Halting Problem is unsolvable, we do so by considering contrived programs that we could write if the Halting Problem were solvable. The contrivedness of those programs does not invalidate the proof of the unsolvability of the Halting Problem. Again, when we prove that C++ templates are Turing complete [17], we do so by considering extremely bizarre C++ templates that would never arise naturally in a software studio. This does not invalidate the proof that C++ templates are Turing complete. Reinforcement Learning is very useful for many practical tasks, but at least in its traditional flavor, it is too constrained (by its arbitrary choice of number system for its rewards) to apply to certain non-Archimedean tasks[5], which, however contrived they are, could certainly be attempted by an AGI. Traditional reinforcement learning will not lead to AGI.

4 Non-traditional reinforcement learning

We have argued that traditional RL cannot lead to AGI, because an AGI is capable of attempting non-Archimedean tasks whose rewards are too rich to be expressed using real numbers. There are at least two potential ways to change RL so as to make it applicable to such tasks and, thus, at least potentially capable of leading to AGI.

[5]Perhaps explaining why "despite almost two decades of RL research, there has been little solid evidence of RL systems that may one day lead to [AGI]" [12].

4.1 Preference-based reinforcement learning

A lot of exciting research has been done on non-traditional variations of RL where, instead of giving the agent numerical rewards for taking actions, one instead informs the agent about the relative preference of various actions or action-sequences. See [19] for a survey. This very nicely side-steps the problems from this paper.

4.2 Reinforcement learning with other number systems

The most obvious way to modify RL to avoid the problems presented in this paper is to change which number system is used. As far as this author is aware, the choice to use real (or rational) numbers for rewards was not made based on any fundamental criteria. The real (or rational) numbers are currently a useful pragmatic choice because they are easy to compute with using 21st century software and 21st century school curricula, but that's hardly relevant in the field of genuine AGI. One might say the real numbers were a good choice because they are familiar, but even that is arguable: in general, students are not taught what the real numbers *actually are*, unless they major in pure mathematics at the university level. Anyway, the familiarity argument is totally irrelevant in the field of AGI.

Various non-Archimedean number systems exist, but the choice is simplified because we clearly want to choose the most general possible number system (unless we have some compelling reason to believe rewards should be more structurally limited then necessary, which, in the context of AGI, seems doubtful). All of the well-known non-Archimedean extensions of $\mathbb{R}$ are subsystems of the so-called *surreal numbers* [7] [10]. The surreal numbers were initially discovered during John Conway's attempts to study two-player combinatorial games like Go and Chess, so it would not be surprising if they turn out to be important in the eventual development of AGI.

The surreal numbers are defined as the union of a hierarchy S_α of subsystems where α ranges over the ordinal numbers. In practice, any particular AGI might not be capable of comprehending the entire class of all surreal numbers, instead only being able to comprehend those subsystems S_α such that α is a computable ordinal which has some code which the AGI knows is the code of a computable ordinal. At least within the constraints of the Church-Turing thesis, no individual AGI can know such codes for all computable ordinals, because the set of codes of computable ordinals is badly non-computable. Any individual RL environment might have rewards limited to S_α for a specific computable ordinal α, and an AGI too weak to comprehend α would be unable to participate in that environment[6]. We would submit this as evidence in favor of our thesis [2] that a machine's intelligence ought to be measured in terms of the computable ordinals which the machine comprehends. Anyway, this points at a possible joint path toward AGI incorporating both machine learning and symbolic logic, perhaps a much-needed reconciliation of these two approaches.

[6]This situation is reminiscent of [9].

5 Conclusion

In traditional reinforcement learning, utility-maximizing agents interact with environments, receiving real (or rational) number rewards in response to actions, and using those rewards to update their behavior. We have argued that the decision to limit rewards to real numbers is inappropriate in the context of AGI, because the real numbers have the Archimedean property, which makes it impossible to use them to accurately portray the value of actions when a task involves inherently non-Archimedean rewards. Thus, we argue, traditional RL cannot possibly lead to AGI, because a genuine AGI should have no trouble at least making attempts to engage in tasks that inherently involve non-Archimedean structures. We suggested two possible ways to modify traditional reinforcement learning to fix this bug: switch to preference-based reinforcement learning, or else generalize reinforcement learning to allow rewards from a non-Archimedean number system (such as the surreal numbers).

References

[1] Haidar Al-Dhalimy and Charles J Geyer. Surreal time and ultratasks. *The Review of Symbolic Logic*, 9(4):836–847, 2016.

[2] Samuel A Alexander. Measuring the intelligence of an idealized mechanical knowing agent. In *Cognition, Interdisciplinary Foundations, Models, and Applications (CIFMA)*. Springer, 2019.

[3] Hajnal Andréka, Judit X Madarász, István Németi, and Gergely Székely. A logic road from special relativity to general relativity. *Synthese*, 186(3):633–649, 2012.

[4] Archimedes. On the sphere and cylinder. In Thomas Heath, editor, *The works of Archimedes*. Cambridge University Press, 1897.

[5] Jacques Bair, Piotr Błaszczyk, Robert Ely, Valérie Henry, Vladimir Kanovei, Karin U Katz, Mikhail G Katz, Semen S Kutateladze, Thomas McGaffey, David M Schaps, David Sherry, and Steven Shnider. Is mathematical history written by the victors? *Notices of the American Mathematical Society*, 60(7):886–904, 2013.

[6] Vieri Benci, Leon Horsten, and Sylvia Wenmackers. Non-archimedean probability. *Milan Journal of Mathematics*, 81(1):121–151, 2013.

[7] John H Conway. *On Numbers and Games*. CRC Press, 2nd edition, 2000.

[8] Euclid. Book v: Theory of proportion. In John Casey, editor, *First Six Books of the Elements of Euclid*. Project Gutenburg, 2007.

[9] Bill Hibbard. Measuring agent intelligence via hierarchies of environments. In *International Conference on Artificial General Intelligence*, pages 303–308. Springer, 2011.

[10] Donald Ervin Knuth. *Surreal numbers: a mathematical novelette*. Addison-Wesley, 1974.

[11] Shane Legg and Marcus Hutter. Universal intelligence: A definition of machine intelligence. *Minds and machines*, 17(4):391–444, 2007.

[12] Scott Livingston, Jamie Garvey, and Itamar Elhanany. On the broad implications of reinforcement learning based agi. In *International Conference on Artificial General Intelligence*, pages 478–482, 2008.

[13] Plato. Protagoras. In John M Cooper, Douglas S Hutchinson, et al., editors, *Plato: complete works*. Hackett Publishing, 1997.

[14] Wolfram Pohlers. *Proof theory: The first step into impredicativity*. Springer, 2008.

[15] Patrick F Reeder. *Infinitesimals for Metaphysics: Consequences for the Ontologies of Space and Time*. PhD thesis, The Ohio State University, 2012.

[16] Elemer E Rosinger. Cosmic contact: To be, or not to be archimedean? *arXiv preprint physics/0702206*, 2007.

[17] Todd L. Veldhuizen. C++ templates are turing complete. Technical report, Indiana University, 2003.

[18] Pei Wang and Patrick Hammer. Assumptions of decision-making models in agi. In *International Conference on Artificial General Intelligence*, pages 197–207. Springer, 2015.

[19] Christian Wirth, Riad Akrour, Gerhard Neumann, and Johannes Fürnkranz. A survey of preference-based reinforcement learning methods. *The Journal of Machine Learning Research*, 18(1):4945–4990, 2017.

[20] Yufan Zhao, Michael R Kosorok, and Donglin Zeng. Reinforcement learning design for cancer clinical trials. *Statistics in medicine*, 28(26):3294–3315, 2009.