

ANDRÁS KORNAI

# Realizing Monads

## ABSTRACT

We reconstruct monads as cyclic finite state automata whose states are what Leibniz calls *perceptions* and whose transitions are automatically triggered by the next time tick he calls *entelechy*. This goes some way toward explaining key aspects of the *Monadology*, in particular, the lack of inputs and outputs (§7) and the need for universal harmony (§59).

## 1. INTRODUCTION

*Monads* are important both in category theory and in programming language design, clearly demonstrating that Leibniz is greatly appreciated as the conceptual forerunner of much of abstract thought both in mathematics and in computer science. Yet we suspect that Leibniz, the engineer, designer of the Stepped Reckoner, would use the resources of the modern world to build a very different kind of monad. Throughout this paper, we confront Leibniz with our contemporary scientific understanding of the world. This is done to make sure that what we say is not “concretely empty” in the sense of Unger (2014). We are not interested in what we can say about Leibniz, but rather in what Leibniz can say about our current world model. Readers expecting some glib demonstration of how our current scientific theories make Leibniz look like a fool will be sorely disappointed.

Looking around, and finding a well-built Calculus Ratiocinator in Turing machines, pointer machines, and similar abstract models of computation, Leibniz would no doubt consider his *Characteristica Universalis* well realized in modern proof checkers such as Mizar, Coq, or Agda to the extent mathematical deduction is concerned, but would be somewhat disappointed in their applicability to natural philosophy in general. He would also, we argue here, be rather dissatisfied with the category-theoretic definition of monads over a category  $\mathcal{C}$ , given by an endofunctor  $T$ , and two natural transformations  $\eta : I_{\mathcal{C}} \rightarrow T$  and  $\mu : T^2 \rightarrow T$  satisfying some identities we need not go into here. (The functional programming definition of a monad involves a further natural transformation ‘tensorial

strength' that satisfies further identities.) It is not so much the fact that the standard definition of monads is complex (in the ordinary sense of the word) that is bothersome here, but that the mathematical definition lacks some of the essential features that Leibniz emphasizes throughout the *Monadology* (which we will quote in the Robert Latta translation, available online at <http://oregon-state.edu/instruct/phl302/texts/leibniz/monadology.html>). In Section 2 we lay the groundwork by showing that attributing a clockwork theory of monads to Leibniz is, at the very least, highly plausible in the light of his other writings, and in Section 3 we explicate this theory in contemporary terms.

## 2. THE CLOCKWORK

Leibniz makes it abundantly clear at the outset (§§1-6) that monads are *atomic* in the original sense of the word. The atoms of science, chemistry in particular, are known not to meet Leibniz's criteria of atomicity, inasmuch as they are built from smaller units, quarks. More important, the quarks themselves are known not to meet these criteria in that they are characterized by attributes (charge, mass, color, spin) that are not constant throughout their lifetimes, can decay, be created out of nothing, and so on. Here we will take an unashamedly Platonic view and assume, without argumentation, that well-defined mathematical objects such as numbers or triangles exist once and for all, and meet the persistence criteria insisted upon by Leibniz. This is not to say that Leibniz himself was a Platonist, though this is a very reasonable assumption, see e.g. Grosholz (1996) or Mercer (2008). Like any original thinker, Leibniz to a remarkable extent defies classification in terms of philosophical traditions, and we use the label Platonism here only as a convenient shortcut avoiding a much longer, and for our purposes unnecessary, ontological discussion concerning the status of mathematical objects. Suffice to say that Leibniz had the contemporary computer scientist's disdain for making strong distinctions between mathematical models and their physical realization.

There is one assumption we will need to make: that even seemingly complex mathematical objects, such as cyclic groups or prime order, are in fact atomic in the sense relevant to Leibniz. This of course leaves a huge variety of mathematical structures to choose from, very much including the current (Godement/Mac Lane) definition of monad, so it will take special effort to show that in the entire mathematical bestiary it is cyclic groups of prime order that we should concentrate on, especially as group theory postdates Leibniz by more than a century. We begin with a simple form of the Chinese Remainder Theorem, that  $p_1, p_2, \dots, p_k$  for primes and arbitrary remainders  $r_1, r_2, \dots, r_k$  satisfying  $0 \leq r_i < p_i$  there is exactly one number  $x$  between 0 and  $p_1 p_2 \cdots p_k$  such that dividing  $x$  by  $p_i$  leaves remainder  $r_i$ .

The *Monadology* dates from 1698, and as it happens, both 1697 and 1699 are primes. If we set up two clock dials with 1697 and 1699 angular steps respectively, and each with a hand advancing once a second, each hand will make its round in less than half an hour (1800 seconds), but it will take over 33 days for the pair of hands to revisit the same position. With six dials and six primes, say 1693, 1697, 1699, 1783, 1787, and 1789, each dial still revisits the same position within half an hour, yet the total configuration is not repeated for over 881 billion years, over 60 times the estimated lifetime of the universe. Leibniz, of course, was quite familiar with the Chinese Remainder Theorem, and as the designer of the Stepped Reckoner, was no doubt aware of these implications.

Further, since Leibniz was a prominent supporter of the *machina mundi*, the clockwork theory of the universe, our main thesis concerning the *Monadology*, that he sought to model the world as being composed of elementary clocks, is not at all implausible, even though the proper conceptual machinery to make this stick, the theory of cyclic groups, has not been invented at the time. As is often the case, Leibniz's views are best characterized by his opponents, in this case Clarke (1715) who writes (First Reply, §4):

The Notion of the World's being a great Machine, going on without the Interposition of God, as a Clock continues to go without the Assistance of a Clockmaker; is the Notion of Materialism and Fate, and tends [...] to exclude Providence and God's Government in reality out of the World.

It is a fascinating aspect of the true history of ideas that Newton, whose universe will be taken by subsequent generations to be an absolutely deterministic clockwork model, was himself a theist, very open to the idea of God's direct intervention (Dr. Clarke is generally assumed to be Newton's mouthpiece in the Clarke-Leibniz debate), and it is Leibniz who is the forerunner of the more deist position of the early Enlightenment, Kant in particular. Coffa (1991.14) actually sees Leibniz as the pivotal figure between modern logical semantics and the Aristotelian tradition:

How do we determine a constituency of a concept; what criteria determine whether a concept B is "in" the concept A? When Kant started to think about this question, there were two standard answers, one emerging from a long and venerable tradition, the other first put forth by Leibniz. The Leibniz-Arnauld correspondence clearly displays the conflict between these two standpoints. With his characteristic blend of genius and insanity, Leibniz had conceived a project in which the simple concepts would be represented by prime numbers and their composition by multiplication. From the Chinese number theorem (and certain assumptions about the nature of truth) he inferred that – given this "perfect language" all matters of truth could be resolved by appeal to the algorithm of division. "For example", he explained,

if the symbolic number of man is assumed to be 6 and that of ape to be 10, it is evident that neither does the concept of ape contain the concept of man, nor does the converse hold ... If, therefore, it is asked whether the concept of the wise man is contained in the concept of the just man... we have only to examine whether the symbolic number of the just man can be exactly divided by the symbolic number of the wise man (Leibniz 1966. 22).

Thanks to Boole and Gödel, today we immediately recognize the semantic theory put forth by Leibniz: concepts are to be analyzed into primitives, each primitive is to be assigned a prime number, conjunctive concepts are to be assigned the product of the component concepts' numbers, and logical derivation is to be done by arithmetic. What Leibniz is doing here (already the subject of his habilitation thesis, *Dissertatio de arte combinatoria*) is devising a mapping from concepts, both simple and compound, to numbers, and modeling the calculus of concepts on Euclid's theory of primes.<sup>1</sup> This presupposes precisely the theory of qualities expressed in §8: *For what is in the compound can come only from the simple elements it contains, and the Monads, if they had no qualities, would be indistinguishable from one another, since they do not differ in quantity.*

Again, this must be contrasted with our current understanding of qualities, even at the price of being completely ahistoric. Leibniz, erstwhile secretary of the Nuremberg alchemical society (Ross 1974), was keenly aware of the conceptual problems attendant to transmutation, especially as these issues, transubstantiation in particular, occupied central position in scholastic debates. Consider the property or quality of tasting salty. Today we know that salt is NaCl, and neither constituent tastes salty. Similarly, water is liquid at room temperature, but neither H nor O are. We derive salty taste as an essentially geometric property, with NaCl molecules having the shape (including distribution of electrical charges) that makes them fit well with certain receptors on the tongue, and the same kind of explanation, attraction of water molecules to one another, is used in computing the liquid state.

Leibniz, while obviously not aware of the specifics that will take centuries from alchemy to quantum chemistry to develop, was nevertheless very clear in his support for configurationally emerging qualities: §64. *Thus the organic body of each living being is a kind of divine machine or natural automaton, which infinitely surpasses all artificial automata. For a machine made by the skill of man is not a machine in each of its parts. For instance, the tooth of a brass wheel has parts or fragments which for us are not artificial products, and which do not have the special characteristics of the machine, for they give no indication of the use for which the wheel was intended. But the machines of nature, namely, living bodies, are still machines in their smallest parts ad*

<sup>1</sup> Subsequently, he will refine the system by introducing a separate list for negative properties, akin to the denominator in fractions, but we will not follow this development here.

*infinitum. It is this that constitutes the difference between nature and art, that is to say, between the divine art and ours.*

As this and many other sections of the *Monadology* make clear, a literal reading of the work no longer makes sense in the light of subsequent developments. Cell theory, and the doctrine that each cell is formed of other cells, *omnis cellula e cellula*, was not properly articulated until the 19th century. The full synthesis of a living cell from inorganic materials is still some ways off (see the 2010 JCVI FAQ), but by now the scientific consensus treats life itself as an emergent property. That said, a less literal reading, wherein life is not attributed to any particular monad, but to the configuration as a whole, is still feasible, indeed preferred. Since the genetic configuration in question is very complex (560 kilobases, slightly over one megabit, not counting the empty cell in which it is placed), it is currently beyond our materials science modeling ability. This leaves some maneuvering room for hard-line skeptics who deny that life can emerge from purely inorganic matter. But clearly, with the advent of replicating artificial organisms, empirical study of entelecheia is slowly becoming feasible.

### 3. CYCLIC FLUENTS

At this point, we have read about as much into the *Monadology* as we plausibly can. What remains is to provide a contemporary technical apparatus built from parts such as *cyclic groups* and *fluents* that were not at Leibniz's disposal, and see to what extent his statements make sense in terms of this apparatus. Again, our goal is to confront our contemporary scientific understanding of the world with the aid of these newly reconstructed monads, treating the philological problem "is this *really* what he meant?" as secondary compared to the philosophical problem of "is this *really* the way things are". That said, it takes no special effort to read the *Monadology* in the manner indicated here, and time and again we will call attention to sections that would, under other readings, remain highly challenging or downright mysterious.

In situation calculus (McCarthy and Hayes 1969) and other versions of action logic, *fluents* are simply variables sensitive to the passage of time. In the AI/planning literature the background is almost uniformly assumed to be absolute, Newtonian time. (In the past decade relative, Einsteinian time is beginning to get more attention, see e.g. Andréka et al. 2012, but this work also assumes an underlying real closed field, elementary equivalent to  $R$ .) Our current nuanced understanding of the real line  $R$  as a suitable theory of both one-dimensional Euclidean space and of the passing of (absolute or relative) time is a relatively new development, starting with Cauchy, Weierstrass, Bolzano, and still significantly incomplete until the deeper foundational work of Tarski (1951) and Robinson (1966). We see Leibniz as operating in an entirely different regimen, again more

characteristic of computer science, where discrete time is assumed as a matter of course. We will be working with fluents whose underlying time parameter is taken from a discrete set. If we use an elementary time tick  $t$ , either we obtain  $Z$  or, once we have  $t \cdot n = 0$  but no  $t \cdot m = 0$  for  $m < n$ , we have the cyclic group  $Z_n$  as our basic clock wheel. §11 *It follows from what has just been said, that the natural changes of the Monads come from an internal principle, since an external cause can have no influence upon their inner being.* §12. *But, besides the principle of the change, there must be a particular series of changes, which constitutes, so to speak, the specific nature and variety of the simple substances.* §13. *This particular series of changes should involve a multiplicity in the unit or in that which is simple. For, as every natural change takes place gradually, something changes and something remains unchanged; and consequently a simple substance must be affected and related in many ways, although it has no parts.*

The simplest case, a binary ticker, has two states 0 and 1, and corresponds in linear time  $Z$  to the predicate *even* or its negation *odd*. Inspecting such a ticker at random intervals may show it more often, say with probability  $p > 1/2$ , in state 0 than in state 1. In such case we can even it out by the same trick that we use to construct a fair coin from an inherently biased one, redefining “tick” as change from 0 to 1 or change from 1 to 0. This way, the two kinds of newly defined ticks become equiprobable independent of the value of  $p$ , so our clock ticks somewhat more evenly. There is still no guarantee that the absolute time spent between two ticks is exactly the same on each occasion, nor is there a guarantee that the “up” tick would fall exactly halfway between two “down” ticks. Nevertheless, we can now have a *virtual* binary ticker, defined by the *average* of the time spent between up- and down-ticks.

There are hard conceptual and practical problems here, e.g. protecting the clock from external effects (mechanical clocks are very sensitive to temperature, and we don’t know too much about the effects of e.g. strong magnetic or gravitational fields on atomic clocks) and from slow drift (which may affect the physical constants the clock is based on). We will steer clear of these problems here, assuming simply that there *are* elementary clocks of sufficient precision, a move Leibniz would no doubt approve of. The easiest way to build something more complex from elementary clocks (cyclic groups  $C_i$ ) is a clock  $C = \prod_i C_i$  containing these. This is simply the direct (Cartesian) product of the components, with no provision that the  $C_i$  start off in the same state. Indeed, we can arbitrarily designate one step, any step, as the starting point on every face.

We stay a step removed from the contemporary theory of cyclic groups in that powers of the same number must be disregarded: once we have clock  $C_p$  for some prime  $p$  there is no value in forming  $C_p \times C_p$  as this will measure the exact same time  $p$  – only relative primes need apply. Leibniz would no doubt be seriously challenged in extending the *Monadology* to cases of allotropy brought forth by Berzelius a century later, but we are not particularly interested in the outcome, since by now we are reasonably certain that in contemporary materials

science we already have the right theory of allotropy. What we are interested in is applying the *Monadology* to those cases where our best contemporary theories are still fraught with difficulty, in particular, the study of entelechy.

*Why* do the planets revolve? Why do people strive, and if they strive so mightily, why do they die anyway? Leibniz is generally viewed as the person who took up the key issues of Aristotelian dynamics, energy and entelechy, and provided the philosophical underpinnings of the transition to classical physics. Yet his physics (for a summary see McDonough 2014) is developed in directions completely orthogonal to the direction that classical mechanics took: there is little in his work that truly anticipates d'Alembert, Fourier, Lagrange, and Hamilton, and a great deal that concerns itself with properties such as elasticity that remained outside the purview of classical mechanics for centuries after him. When Leibniz rejects Newton's greatest invention, gravity, this is based neither on personal animus nor on metatheoretical considerations such as rejecting action at a distance (though in his third letter to Dr. Clarke he certainly compares this to occult powers), but rather on a strongly felt need to provide a theory of the soul.

The fact of the matter is that in critical respects *De Anima* is impossible to make sense of, at least in the form we have it. To Leibniz the missing part is the *realization*: how do you construct a soul? Now that we are building all kinds of soulless automata, ones with remarkable sensory and motion capabilities, the question is even more acute. According to Iltis (1971)

What is real in nature for Leibniz is primitive force or striving, and this was developed by him in the succeeding years as the essence of the monad.

Planets, as far as Leibniz is concerned, will follow their prescribed orbits just by virtue of conservation of momentum. Unlike Newton, whose primary accomplishment is to unify the sublunary and the superlunary spheres by means of gravity, Leibniz is interested in separate theories of the two, with the *Monadology* to fill the gap left by *Discours de Métaphysique*. When we construct a soul, we need good building materials, not just for immortality—obviously a nontrivial design goal—but also for free will. Newton himself may have believed that planets have souls (schoolbook presentations in subsequent centuries have carefully suppressed the occult elements in Newton's thinking) but these are not souls with free will.

Today, computer science offers a satisfactory mathematical model of free will in nondeterministic automata (Floyd 1967), and the very idea of using purely mathematical entities such as finite groups is sufficient to guarantee a solid foundation for immortality. The central contribution of the *Monadology* to the construction of a soul is the theory of entelechy it offers. §14. *The passing condition, which involves and represents a multiplicity in the unit or in the simple substance,*

*is nothing but what is called Perception, which is to be distinguished from Apperception or Consciousness (...) In this matter the Cartesian view is extremely defective, for it treats as non-existent those perceptions of which we are not consciously aware. This has also led them to believe that minds [esprits] alone are Monads, and that there are no souls of animals nor other Entelechies. Thus, like the crowd, they have failed to distinguish between a prolonged unconsciousness and absolute death, which has made them fall again into the Scholastic prejudice of souls entirely separate [from bodies], and has even confirmed ill-balanced minds in the opinion that souls are mortal.*

Planets will continue to revolve simply because there is no dissipative force acting on them, but why do people continue to go on in the face of difficulties? The answer, we propose here, is that they are driven by entelechy, an internal clockwork that will go on even when they are asleep or unconscious for a prolonged time. §15. *The activity of the internal principle which produces change or passage from one perception to another may be called Appetition. It is true that desire cannot always fully attain to the whole perception at which it aims, but it always obtains some of it and attains to new perceptions.*

This provides a new answer to concerns recently raised in regards to *wire-heading* (Yampolskiy 2014), the situation when rats, humans, or artificial general intelligences (AGIs) gain direct access to their reward mechanism and can freely stimulate themselves. For Leibniz, this would amount to continual resetting of the internal clock, with the subject being stuck in the same (pleasurable) state. As a practical matter, AGIs currently have no access to their utility functions (Lenat 1983), but the matter becomes increasingly urgent as our understanding of the human mind and body, and the concomitant potential for self-modification, grows.

#### 4. CONCLUSIONS

The main design lessons to be learned from the *Monadology* are rather clear. A basic, maybe *the* basic, defining element of the human condition is awareness of our mortality. If life is an emergent property, so is death. Of the many parallel working internal clocks providing entelechy, there need not be any single one responsible for the main cycle of our lives. Sometimes we may get a glimpse, as with the discovery of the mutation responsible for Huntington's disease, a situation that Ridley (2006) describes as follows:

Either you have the Huntington's mutation and will get the disease or not. This is determinism, predestination and fate on a scale of which Calvin never dreamed.



But in general, we are still in the dark. Individual experience is practically worthless (premotions, if true, appear to come too late), and collective experience provides only crude statistical guidelines. The desire to tamper with one's internal clock and stay forever young is universal, and if life expectancy statistics are to be believed, humanity is actually making noticeable gains in this regard. However, if we take matters into our own hands (as we should), the task is much larger than better understanding the genetic code or hyperfine transitions (the current basis of our atomic clocks). This is because for entelechy to work, we need §59 (...) *universal harmony, according to which every substance exactly expresses all others through the relations it has with them.*

This is a tall order. The next question, which we cannot in any detail discuss here, is whether we can build something that provides a reasonable approximation of universal harmony the same way as computers are reasonable approximations of Turing Machines. The Stepped Reckoner had problems with decimal carry that took a century to debug through Payen's Arithmomètre and the Tarrant-Felt Comptometer that stayed in use until the advent of electronic calculators. Here we are faced with a somewhat similar problem: §63. *The body belonging to a Monad (which is its entelechy or its soul) constitutes along with the entelechy what may be called a living being, and along with the soul what is called an animal. Now this body of living being or of an animal is always organic; for, as every Monad is, in its own way, a mirror of the universe, and as the universe is ruled according to a perfect order, there must also be order in that which represents it, i.e. in the perceptions of the soul, and consequently there must be order in the body, through which the universe is represented in the soul.*

In contemporary terms, what we need is some form of knowledge representation that is as tightly coupled to the motivational system (entelechy) as to the sensory system. Without this, automata cannot be animated.

#### ACKNOWLEDGEMENT

We thank András Máté (ELTE) for pointing out Leibniz's pioneering treatment of negation.

## REFERENCES

- A Collection of Papers, Which passed between the late Learned Mr. Leibnitz, and Dr. Clarke, In the Years 1715 and 1716 (London: 1717).* <http://www.newtonproject.sussex.ac.uk/view/texts/normalized/THEM00227>, 1715.
- Anon., J. Craig Ventner Institute. *First self-replicating synthetic bacterial cell FAQ*. <http://www.jcvi.org/cms/research/projects/first-self-replicating-synthetic-bacterial-cell/faq>, 2010.
- Coffa, J. Alberto. *The Semantic Tradition from Kant to Carnap: To the Vienna Station*. Cambridge University Press, 1991.
- Floyd, Robert W. Nondeterministic algorithms. *Journal of the ACM (JACM)*, 14(4):636–644, 1967.
- Grosholz, Emily. Plato and Leibniz against the materialists. *Journal of the History of Ideas*, 57(2):255–276, 1996.
- Hajnal, Andréka, Madarász, Judit X., Némethi, István, and Székely, Gergely. A logic road from special relativity to general relativity. *Synthese*, 186(3):633–649, 2012.
- Iltis, Carolyn. Leibniz and the vis viva controversy. *Isis*, 62(1):21–35, 1971.
- Leibniz, Gottfried Wilhelm. *Philosophical Papers*, ed. George Henry Radcliffe Parkinson. Clarendon Press, 1966.
- Lenat, Douglas B. EURISKO: A program that learns new heuristics and domain concepts. *Artificial Intelligence*, 21:61–98, 1983.
- McCarthy, John and Hayes, Patrick J. Some philosophical problems from the standpoint of artificial intelligence. In Meltzer, B. and Michie, D., editors, *Machine Intelligence 4*, pages 463–502. Edinburgh University Press, 1969.
- McDonough, Jeffrey K. Leibniz’s philosophy of physics. In Zalta, Edward N., editor, *The Stanford Encyclopedia of Philosophy*. Spring 2014 edition, 2014.
- Mercer, Christia. The Platonism at the core of Leibniz’s philosophy. In Hedley, Douglas and Hutton, Sarah, editors, *Platonism at the Origins of Modernity*, volume 196 of *International Archives of the History Of Ideas*, pages 225–238. Springer Netherlands, 2008.
- Ridley, Matt. *Genome*. Harper, 2006.
- Robinson, Abraham. *Non-standard Analysis*. North-Holland, 1966.
- Ross, G M. Leibniz and the Nuremberg alchemical society. *Studia Leibnitiana*, 6: 222–248, 1974
- Tarski, Alfred. *A decision method for elementary algebra and geometry*. Berkeley and Los Angeles, 1951.
- Unger, Peter. *Empty Ideas*. Oxford University Press, 2014.
- Yampolskiy, Roman V. Utility function security in artificially intelligent agents. *Journal of Experimental & Theoretical Artificial Intelligence*, 26:1–17, 2014.