<u>Engineering with Irreducibility</u>

Steph Macurdy, Wolfram Blockchain Labs

Abstract

This paper introduces an integrative framework for understanding Emergence. It traces a path from Wolfram's Computation, to Deacon's Teleodynamics, to Nakamoto's Proof-of-Work. Wolfram's framework of computational equivalence and the role of the observer set the stage, highlighting how an epistemological choice—viewing all processes as computations—carries ontological consequences for what we can know about reality, with wide ranging implications about irreducibility and emergence. Deacon's framework articulates how spontaneous, non-spontaneous, and end-directed (purpose-driven) processes emerge through layers of constraints, establishing a hierarchical theory of emergence, with causal explanations for origins of life, consciousness, and symbolic systems. Nakamoto's Bitcoin and consensus mechanism "proof-of-work" establishes the engineering schematics and infrastructure for quantifying various parameters of emergent "symbolic value" in Internet Protocols (TCP/IP). Because blockchain protocols are undergoing such radical transformation through evolutionary trial-and-error, there is a strong expectation that a new protocol will emerge with more precise and scalable properties for quantifying computation, work and "symbolic value." Downstream implications of this possibility are in the fields of complexity science, economics, game theory, distributed computing, quantum cryptography, Artificial Intelligence, smart contracts, and beyond—ultimately making a rigorous and universal connection between the computational, the physical and the social. By connecting these insights from Wolfram's *New Kind of Science*, Deacon's *Incomplete Nature*, and Nakamoto's *Bitcoin*, this thesis offers an articulation of how fundamental computational constraints shape scale-invariant emergence, give rise to symbolic value, and ultimately enhances our scientific understanding of reality.

Introduction

Beginning with an assumption that all processes can be modeled as computations—a perspective that shapes an observer's understanding (epistemology) and the nature of what one can observe (ontology)—we employ Wolfram's **Principle of Computational Equivalence** which states that systems in nature reach a maximal level of computation, rendering them effectively equivalent in sophistication to one another—and to us, their observers. This universal equivalence establishes an ontological boundary, described by Wolfram as "**computational irreducibility**," where beyond a certain threshold the behavior of systems cannot be shortcut by simpler methods; it must be "computed" step by step.

Computational irreducibility generalizes the second law of thermodynamics, restated here as the "**Second Law of Statistical Mechanics**," and offers a broader perspective on what an observer can predict or reduce, invariant to any system of any scale. **Emergence** arises along a spectrum of irreducibility, where the boundary of reducibility-to-irreducibility establishes a constraint that prevents a full reduction of events, yet where *repetitive*, *nested*, *asymmetric, irregular*, and *random* phenomena still materialize. Deacon's hierarchical theory of **Teleodynamic** emergence accounts for transitions between spontaneous, non-spontaneous and end-directed (i.e., purpose-driven) change, highlighting how cause-and-effect relationships unfold in physical reality, which becomes useful in understanding the cause-and-effect relationships of observers in cultural, economic, technological, and symbolic reality.

Finally, these ideas are illustrated with **blockchain** technology. Nakamoto's conception of "proof-of-work" in Bitcoin exemplifies how computational irreducibility enforces a physical and verifiable boundary condition, which ultimately runs the logic that generates the container and the self-organizing infrastructure for facilitating **symbolic value**. In this sense, "proof-of-work" establishes a physical constraint (thermodynamic) and a reducible validation constraint (morphodynamic) which generates an emergent, observer-dependent notion of value that can be used as a downstream constraint in economic, cultural and technological systems (teleodynamic). "Proof of work" consensus establishes a revolutionary new process for understanding and quantifying the end-directed work of observers by connecting irreducibility, work, and symbolic value together in a quantified system. "Quantified" implies measurement capability; as number of hashes per second or kilowatt hours per computation. Where, how and why one would use such a system is partially left unarticulated and open-ended by the author. However, the three most important takeaways claimed in this thesis are: 1) that a rigorous engineering schematic for creating "symbolic value" systems now exists as open-source code (expecting to see many of these experiments) 2) a precision measurement system, akin to the atomic clock for time, telescope for stars, and microscope for bacteria, now exists for "symbolic value" (expecting to see a few of these experiments) and 3) a protocol of physical constraints can be employed to cultivate emergent, symbolic forms of non-organic life; from smart contracts to artificial super intelligence, and constrain their end-goals (hoping to see one or two of these experiments).

Below is an outline that attempts to walk through the basic logic of the emergence framework, from “Observer” through “Symbolic Value.” Feedback is most welcome, my own comments included in the Word file.

---

0. Computational Observer

- All processes are viewed as computations
- This is an epistemological decision that determines an ontological perspective
- Computation is a generalization made to enable observers like us to most accurately understand many different kinds of systems and processes.
- This model should offer causal explanatory power greater than other models.
- Assumptions:
    - Observers like us are:
        - 1) computationally bounded
        - 2) perceive persistent state through time

1. The Principle of Computational Equivalence

- The computational sophistication of many different kinds of systems and processes has a maximum upper bound.
- It says that systems found in the natural world can perform computations up to a maximal "universal" level, and that many systems do in fact attain this maximal level of computational power. Consequently, many systems are computationally equivalent.
- For example, the workings of varied systems such as the human brain or the weather can, in principle, compute the same things.
- Computation is therefore simply a question of translating inputs and outputs from one system to another.
- Turing completeness is one way of determining evidence of this across observed systems and simple rules.

2. Resulting Computational Irreducibility

- Past some low threshold, any real system may exhibit essentially the same level of computational sophistication. This implies that observers will tend to be computationally equivalent to the systems they observe.
- This implies that those observers will consider the behavior of such systems as computationally irreducible.
- There won’t be a shortcut to understanding the system beyond the explicit evolution of the system step by step.

- In the end, complexity is a result of computational irreducibility from the principle of computational equivalence, and observations in practice will demonstration that observers are computationally equivalent to systems they're observing.

3. Implications of an Unprecedently Broad Principle of Computational Equivalence—Invariance!

- The benefits of the broadness of the principle of computational equivalence are that it offers a framework of understanding to observers of systems at all scales.
- This implies that computational equivalence, and thus computational irreducibility, are invariant phenomena.
- This implies that observers of systems at different scales are bounded by upper and lower thresholds.
- This is akin to the general boundedness that the 2^{nd} law of thermodynamics implies, but broader and more applicable to computational systems (which will be useful in the teleodynamic analysis).
- The second law of thermodynamics is an articulation downstream of the principle of computational equivalence, not the other way around.
    - Can this be generalized from the second law of thermodynamics to the second law of statistical mechanics?

4. The Fundamental Constraints

- Emergence is a phenomena arising from and defined by boundary conditions, not by the thing in-of-itself. The thing itself wasn't "there" prior, and the rule that produces it doesn't have the outcome embedded (or knowingly embedded) in the rule itself. Therefore, it is not helpful, useful, or productive to define emergence by what it is, rather it is best defined by the constraints defining what it is not.
- Irreducibility is a phenomena that defines the boundary condition for what an observer can know about a computational system. Beyond some threshold is a boundary that no longer contains reducible information. However, if boundary conditions without any reducible information display a signal of information, it means work has occurred and that is a signal defined by the absence of information.
    - Computational Irreducibility – Upper boundaries
        - Generic equivalence in computational sophistication
        - Boundedness of observers (we are finite)
        - Pure Randomness (does pure randomness exist in reality?)
    - Computational Emergence – Lower boundaries
        - Repetitive
        - Nested
        - Random
        - Asymmetric

        - Irregular
    - Irreducibility is a potent engineering mechanism for establishing boundary conditions to cultivate emergence.

5. Cause-and-Effect(s) of Emergence

- What causes events to occur? (Aristotle's *material, efficient, formal* and *final* causality explained by Deacon)
- Some events occur ***spontaneously*** because probabilistically they are so likely to happen as to be considered an inevitable occurrence. Only external work could violate this tendency.
- Some events occur ***non-spontaneously***, meaning they occur in such unlikely probabilities that an explanation requires new understandings or external work imposed on the system.
- Non-spontaneous events may require an explanation inclusive of the observer because said observer is central to the cause-and-effect relationship.
- Terrance Deacon's "Incomplete Nature" articulates a theory of cause-and-effect relationship between *spontaneous*, *non-spontaneous*, and *end-directed* changes which affords greater scientific understanding.
- In this theory of emergence; three hierarchical levels of system processes (or dynamics) are needed to explain the cause-and-effect relationships of any system:
    - 1) thermodynamics (i.e. non-symmetric, spontaneous)
    - 2) morphodynamics (i.e. symmetric, non-spontaneous, self-organization)
    - 3) teleodynamics (i.e. asymmetric, end-directedness).
- Each level is dependent upon the one below it, meaning there can be no teleodynamic process without a morphodynamic process, and no morphodynamic process without a thermodynamic process.
- Observers like us are in pursuit of reducibility (by our nature of "end-directedness"), to save ourselves from expending excess computational work and prolong/preserve ourselves.
- Observers like us will find ways to reduce or equivalence states together to extract a reducible insight that is useful for our end-directedness (like temperature, pressure, price, etc).

6. 'Via Negativa'

- Constraint is a negative property of a collection or ensemble. It is a way of referring to what is not exhibited, hence 'negativa', but could have been, at least under some circumstances.
- The term *constraints* denotes the property of being restricted or being less variable than possible, all other things being equal, and irrespective of why it is restricted. Something will tend to be assessed as being more orderly if it reflects more constraint. We tend to

describe things as more orderly if they are more predictable, more symmetric, more correlated, and thus more redundant in some features.
- To the extent that constraint is “reduced variety,” there will be more redundancy in attributes. The advantage of this negative way of assessing order is that it does not imply any modeled conception of order, regularity, or predictability.
- To avoid a modeled conception of order, regularity, or predictability in a physical substance or substrate, as in a CA rule or in a material, we eliminate the hard problem of explaining where order physically exists. (Best explained by constraints and absence).

7. A New Kind of Science

- While discrete, bottom-up physics preserves emergence, it loses precision simply because of the implication of the Principle of Computational Equivalence. Meaning, the modeling of the intricate molecular details are inevitably lost at some threshold relative to bounded computational systems of observers like us.
    - The implied mission of the field of Complexity Science is to fill the gap between irreducibly complex systems and non-emergent mathematics.
- Therefore, any model, innovation, or technology that increases the capacity to reduce or equivalence states is work that observers will inevitably exploit.
- The belief that observers can preserve emergence or repeatedly produce the effects of emergence in a way that ultimately reduces their work will explain why observers will exploit models, innovations and technologies. In Wolfram’s terms, “in a sea of irreducibility there will be pockets of reducibility.” Just like the thermometer for temperature, barometer for pressure, or price for markets, these aspects are useful for observers which afford something useful in it’s reduction or equivalence.
- Reduction or equivalencing occurs at all scales, it is invariant in its application.

8. Example: Blockchains and “Proof-of-Work” in Bitcoin

- Wolfram has previously pointed out that Bitcoin's "Proof of Work," while juvenile in its construction, employs computational irreducibility to establish a boundary condition beyond which "work" can be provably verified to have occurred.
- This proof, while expensive in its production, potentially significantly reduces the total amount of "work" that all participating downstream constituents of the protocol require, indefinitely into the future.
- The connection between the “expensive” physical-computational “work” and the resultant “cheap” validation proof, with the intrinsically produced token, is where a container of symbolic value emerges.
- The symbolic value container is defined solely by the network's constraints; specifically computational irreducibility.
- This is a specific kind of "value," which will be interpreted differently by each observer who reads this.

- In this case, “value” is the description of the empty container unit AND the physical infrastructure/processes which facilitates the container unit.
- Ultimately, the facilitation infrastructure can imbue the value container with additional constraints, determined by end user’s goals/objectives.

9. Predictions and Conclusions

- Causal invariance will appear in blockchain consensus mechanisms.
- "Symbolic value" will become a well-defined object in a NKS and Teleodynamic framework; not by the physical substance it's composed of, nor the physical substrate that facilitates it, but by the features left absent by the physically overlapping constraints imposed upon it.
- This story begins with the fundamental constraint defined by the principle of computational equivalence.
- In conclusion, this is a thesis that spans the range from the very bottom of Computation to the very top of Symbolic Value in physical systems that we as observers find meaningful. It offers causally powerful explanations for why and how emergence can occur at various levels of scale and uses the fundamental constraint of computational irreducibility to get us there. Blockchains are an example of this framework and are used to demonstrate the invariance of this explanation at a different scale. This thesis also makes a claim that "symbolic value" in economic space will be most rigorously be defined by a blockchain system, and therefore, should be the center of our exploration. My prediction is that as observers of the kind we are, we'll see scientifically rigorous accounts of "value" in the evolution of blockchain systems which will help unlock our understanding of the connection between Computation, Physics and Economics in the world.

10. Experiments

1. Cellular automata random initial state, signal detection of entropy reduction, database storage of signals detected, state of system updates, database updates, system returns to random state. Rerun experiment
2. Multi-way graph plus a concept of “fuel.” How many cycles can you try to find the shortest path with finite “fuel.”
3. Return to “Game of Life” and include components for absence + work. (Need to find efficient pathways through causal space to reduce “work”).
4. Video game designed for various agents to explore the causal space with finite “fuel/cycles.”
5. How do we know we have exhausted the search for easily reduce CA? Bittensor
6. Describe CMBR and LIGO detection in terms of error reduction leading to signal detection.

Key Questions

- How do I get from Observer, to Computational Equivalence, to Second Law of Statistical Mechanics, to Teleodynamics, to Symbolic Value?
- Do we need an account of the world that explicitly includes symbolic expressions in physical cause-and-effect relationships that can be rigorously measured so that our fundamental theories of the physics can include the non-spontaneous change that we as observers are intentionally capable of. If so, this requires a new kind of science AND a teleodynamic framework!
- I need more intuition for SW ideas:
    - How do we know we're starting from random experiments?
    - How are CA structured in a way that we know it's random?
    - Are we starting in highly constained environments with CA experiments? Should we assume significant order from the start?
    - How do we know we have exhausted the search for easily reduce CA?
    - How do we know by visually inspecting CA that we're detecting complexity?
- How is computational irreducibility seen or defined as a constraint?
- How is computational irreducibility defined relative to the second law of thermodynamics?
- Is computational irreducibility a generalized version of the second law of thermodynamics, best stated as the second law of statistical mechanics?
- How does one explain non-spontaneous end-directed cause and effect relationships?

Asymmetric

A property of processes or structures that do not exhibit symmetry; that is, having parts that fail to correspond to one another in shape, size, or arrangement. Asymmetrical systems naturally arise in nature, like prime numbers, or rule 110 cellular automata.

Blockchain

A distributed ledger technology where data are grouped into blocks that are cryptographically linked. Blockchains rely on consensus mechanisms (like proof-of-work) that leverage computational irreducibility to validate proofs which guarantee trust and security without central authority.

Boundedness of Observers

Refers to the finite or limited capacity of observers—whether in computational power, memory, or time. Because observers cannot out-compute certain systems, they become subject to irreducibility constraints.

Causal Invariance

A principle suggesting that the ultimate outcomes of certain computational or physical processes remain consistent despite variations in their intermediate sequences or paths. In blockchain context, it implies consistency of state despite different validation or block order sequences.

Computational Emergence

A lower bound phenomenon (in this framework) indicating that new, higher-level structures, patterns, or properties arise out of boundary conditions in complex systems, rather than being explicitly encoded in the initial rules or components. Examples include repetitive, nested, asymmetric, irregular, or random patterns.

Computational Irreducibility

The notion that past some threshold, a system's behavior cannot be shortcut or predicted by any simpler computational method than actually simulating the system step by step. In other words, there is no "closed-form" or faster algorithm to determine its future states (as known in mathematics)—this applies universally to many natural and artificial systems.

Computational Observer

An entity (like a human or machine) that treats all processes as computations. This perspective shapes both how the observer acquires knowledge (epistemology) and what can be known about the world (ontology). Such an observer is themselves computationally bounded.

Constraint

Any factor that limits or bounds the behavior of a system. In computational and physical contexts, constraints can be physical, energetic, or logical/informational—preventing certain states while allowing others. Constraints often give rise to emergent properties by defining the boundary conditions within which a system can evolve.

Epistemology

The branch of philosophy concerned with the theory of knowledge—how we know what we know, what justifies our beliefs, and what the limits of knowledge may be. In the context of computational observers, epistemology focuses on how observers acquire, structure, and validate information about the systems they study.

Equivalence of State Through Time

An idea that observers maintain some form of continuity or consistent memory across time, allowing them to identify patterns or regularities, yet remain constrained by computational bounds.

Invariance

A principle or property remaining unchanged under certain transformations or across different contexts/scales. In this framework, **causal invariance** is one example; it suggests that the final outcome of a process is the same even if the intermediate steps are permuted. More broadly, **computational irreducibility** is considered invariant across different scales and systems, which ultimately generalizes the second law of thermodynamics into the law of statistical mechanics.

Irreducibility (as an Engineering Mechanism)

Using irreducibility as a tool to define boundary conditions or constraints in systems. By intentionally imposing aspects of a system that are impossible to shortcut, it is possible to cultivate boundaries for emergent behaviors (e.g., proof-of-work schemes).

Morphodynamics

A level of emergent behavior in Terrance Deacon's hierarchy. It focuses on self-organizing processes that emerge from lower-level thermodynamic processes but are not themselves spontaneous. Morphodynamics show symmetry and pattern formation (e.g., crystal growth). Morphodynamics can only emerge from thermodynamics, and can give rise to a higher level called Teleodynamics.

New Kind of Science (NKS)

Stephen Wolfram's computational framework suggests that simple rules can yield the entires spectrum of simple-to-irreducibly-complex behaviors, illuminating a key insight into complexity science, scale invariance, the second law of statistical mechanics, observer theory, metamathematics, physics, and a wide variety of other fields including philosophy, biology, economics and foreign intelligences.

Observer-Dependent Notions of Value

The idea that "value" is context-specific and emerges from the interplay between physical constraints and the subjective interpretations of those constraints by observers. For instance, Bitcoin's proof-of-work mechanism creates a container for "symbolic value," judged differently by each participant.

Ontology

The branch of philosophy that deals with the nature of reality or being—what fundamentally exists in the world and how those entities relate to one another. In a computational framework, ontology pertains to how the observer's approach (e.g., treating all processes as computations) defines what "is" within the scope of that model.

Principle of Computational Equivalence (PCE)

Wolfram's hypothesis that many systems—biological, physical, or technological—attain the same maximal computational sophistication. Once a system passes a low-level threshold of complexity, it becomes universal in computational power, evidenced by Turing-completeness, which implies an equivalence to those systems.

Proof-of-Work (PoW)

A consensus mechanism (e.g., in Bitcoin) that uses computational irreducibility as a boundary condition. Nodes must perform costly computational "work," which can then be verified cheaply by the validating network, forming the basis for trust and security in a decentralized environment.

Randomness

The aspect of a system's behavior or initial conditions that is unpredictable or cannot be trivially compressed into a simpler rule. Randomness can be part of both emergent processes and irreducible computations, making outcomes difficult or impossible to predict.

Second Law of Statistical Mechanics

A broader interpretation of the second law of thermodynamics, applied to general computational or informational contexts. To the bounded observer, the second law of thermodynamics arises. A computationally bounded observer observing thermodynamic, morphodynamic or teleodynamic work may not necessarily use the historical interpretation of thermodynamics to comprehend work occurring a levels higher in the hierarchy of work and information.

Symbolic Value

Value that is defined not by the physical substrate but by the constraints or rules that limit its production, validation or use. In proof-of-work blockchain systems, "symbolic value" emerges in a functional container unit from the validation of a process which cannot be spoofed. The system facilitating the functional container unit does not define it's symbolic value, but the absence created by the container is established to be valued by the end-directedness of the end-user.

Teleodynamics

The highest-level process in Terrance Deacon's hierarchy, referring to systems that exhibit end-directed behaviors (i.e., purpose-driven or goal-driven processes). Teleodynamic processes depend on morphodynamic and thermodynamic layers to exist and cannot be explained solely by spontaneous or self-organizing phenomena. Teleodyanmic work helps explain cultural, economic, technological and spiritual work in human systems and offers explanation through physical processes to validate that work.

Thermodynamics

Traditionally, the branch of physics dealing with heat, energy, and work. In Deacon's framework, thermodynamics represents the most basic, spontaneous level of physical processes—necessary for higher-level emergent phenomena to build upon. This background constraining layer is more precisely articulated by Wolfram's computational irreducibility and establishes a signal detection layer for work.

Turing-Completeness

A measure of a system's computational universality indicating it can perform any computation that a Turing machine can, given enough time and resources. In this context, Turing-completeness signifies that many natural and artificial systems reach a universal threshold of computation.

Scale Invariance

A specific type of invariance where the system's underlying properties or behaviors remain consistent when viewed at different magnitudes (e.g., microscopic vs. macroscopic levels). For instance, certain fractal patterns or statistical distributions in nature exhibit scale invariance, meaning their complexity remains similar regardless of the scale at which they are measured.

Wolfram's Four (or Five?) Classes of Complexity

Stephen Wolfram's classification of cellular automaton behaviors, generalized here to describe complexity in natural or computational systems. The four classes are:

1. **Repetitive**: The system quickly settles into a stable, repetitive pattern (e.g., a fixed point or cycle).
2. **Nested**: The system produces self-similar, fractal-like structures that replicate at multiple scales.
3. **Random**: The system's outcomes appear disordered or pseudo-random, with no obvious repeating pattern.
4. **Irregular** (sometimes referred to as "complex"): The system exhibits complicated structures or behaviors that are neither fully repetitive nor purely random, often with emergent features.
5. **Asymmetric:** Asymmetrical systems naturally arise in nature, like prime numbers, or rule 110 cellular automata. One-way hash functions like SHA-256 as example.

Work

In thermodynamics and physics, "work" is the energy transferred to or from a system when a force is applied over a distance. In computational or informational contexts, it refers to the demonstrable expenditure of resources—time, energy, processing—for a specific purpose. Proof-of-work blockchains, for example, require participants to perform and show evidence of computational "work" that expends real energy.