CIMC Philosophy Seminar 1: Computational Irreducibility, Probabilistic Generalisations and the Observer

Joscha Bach Paul Borrill Joel Dietz David Kammeyer Hikari Sorensen

Abstract

This paper reports results of a colloquium on computational irreducibility in Wolfram's programme. The colloquium distinguished determinism from efficient predictability, used Rule 110's universality as a benchmark, and treated Rule 30 as an empirically compelling but less formally anchored case of apparent randomness. It further proposed that binary behavioural taxonomies should be supplemented by probabilistic and multiway perspectives, in which uncertainty can be observer-relative. These themes were connected to cryptographic practice, physically rooted device identity, reversible computation and current work on reasoning systems, including tensor-logic unification proposals and recent progress in automated theorem proving. A ranked set of methodological recommendations is presented, prioritised by near-term practical leverage for research and engineering.

Keywords: computational irreducibility; cellular automata; probabilistic cellular automata; multiway systems; reversibility; automated theorem proving

1. Introduction

The colloquium examined a tension between two claims: (i) simple programs often generate behaviour that is hard to predict in practice and (ii) successful science typically offers compressed predictors. Wolfram's notion of *computational irreducibility* was adopted as a statement of the strongest form of this tension: for many systems, later states are not generally obtainable by shortcuts that are substantially cheaper than running the rule (Wolfram 2002).

The organising question was not whether a system is deterministic but whether the *questions posed* about it admit reduction. This emphasis structured the assessment of elementary cellular automata, probabilistic and multiway generalisations, cryptographic practice, reversible computation and current AI reasoning.

2. Irreducibility and the selection effect in scientific modelling

A central thesis was that much of traditional science can be characterised as the pursuit of reductions: equations and models that compress evolution into tractable predictors. This yields a selection effect. Systems that are reducible with respect to observables of interest are disproportionately useful for engineering and therefore disproportionately salient in everyday life. On this view, the apparent predominance of reducible regularities does not by itself indicate that reducibility is typical in the abstract space of programs (Wolfram 2002).

3. Elementary cellular automata as benchmarks

3.1. Rule 110 and universality

Rule 110 was used as the principal benchmark because it has a proof of computational universality (Cook 2004). Universality was taken to anchor the claim that broad behavioural questions can inherit hardness and undecidability phenomena familiar from computability theory (Turing 1936). This does not settle every informal use of 'irreducible', but it supplies a rigorous foothold for arguing that some prediction tasks cannot be uniformly compressed.

3.2. Rule 30 and apparent randomness

Rule 30 was treated as a contrasting case: a deterministic rule that yields strongly random-looking behaviour, including practical uses as a pseudorandom generator. The colloquium's conceptual clarification was that determinism is compatible with effective unpredictability. The relevant issue is whether there exists a systematically cheaper method of predicting the target observable than iterating the rule.

4. Beyond binary classification: probabilistic and multiway generalisations

A recurring thesis was that coarse taxonomies can obscure graded scientific claims.

4.1. Probabilistic cellular automata

Probabilistic cellular automata provide a controlled way to move from single trajectories to distributions over trajectories. They permit study of robustness, sensitivity to stochastic updating and regime dependence in terms not naturally expressed in a purely categorical classification (Mairesse and Marcovici 2014).

4.2. Multiway systems and observer-relative uncertainty

Multiway (branching) rewriting models were treated as a representational shift in which apparent randomness can arise from branching structure plus observer limitations, even if underlying rules are deterministic (Wolfram 2020; Gorard 2020a; Gorard 2020b). The colloquium did not claim that this settles foundational disputes, but it recorded the proposal that many probabilistic descriptions are best understood as statements about coarse-graining and observational perspective.

5. Consequences for cryptography, identity and reversibility

5.1. Cryptographic practice

Cryptographic pseudorandom generation is deterministic in mechanism but assessed in adversarial terms: whether recovering the seed or distinguishing outputs from random is computationally feasible. Standards-driven practice and the historical deprecation of primitives were taken to show that ‘random-looking’ behaviour is not sufficient (NIST 2015).

5.2. Physically rooted identity

Device and link identity were treated as examples where determinism can be security-relevant. Physical one-way functions and PUF-like methods exploit manufacturing variation, supporting authentication and key derivation without requiring metaphysical randomness (Pappu et al. 2002).

5.3. Reversibility and information costs

The colloquium rehearsed the link between logical irreversibility and thermodynamic cost: erasing information dissipates heat, whereas reversible computation can in principle avoid some dissipation (Landauer 1961;

Bennett 1973). This was treated as relevant both to interpretation and to engineering interest in energy efficiency.

6. Reasoning systems and formal methods

Neural networks were treated as systems that aggregate many weak features into robust predictions, especially in overdetermined regimes. The colloquium emphasised that representability results are not sufficient; the practical issue is learnability and optimisation (Cybenko 1989). Evidence that large networks can be pruned after training while retaining performance was used as an illustration (Frankle and Carbin 2019).

The colloquium also recorded interest in unifying formalisms for reasoning and learning (Domingos 2015; Domingos 2025). Formal theorem proving was presented as a domain where correctness is machine-checkable, and recent IMO-level automated theorem proving was cited as evidence that hybrid architectures (informal search plus formal verification) can yield striking performance (Achim et al. 2025).

7. Ranked recommendations and approaches

The colloquium yielded a set of methodological recommendations. The ranking below is editorially synthesised from the colloquium's emphasis and is ordered by near-term practical leverage for research and engineering.

(1) **Specify the observable before arguing about reducibility.** Irreducibility is typically relative to questions asked, not merely to microdynamics (Wolfram 2002).

(2) **Use universal benchmarks to calibrate claims of hardness.** When possible, reduce questions to known universal substrates (Rule 110) to avoid intuition-driven hardness claims (Cook 2004; Turing 1936).

(3) **Move from categorical to graded analysis when evidence is graded.** Prefer probabilistic cellular automata, robustness measures and distributional claims when the phenomenon is regime-dependent (Mairesse and Marcovici 2014).

(4) **Model branching explicitly when uncertainty is observer-relative.** Multiway formalisms provide a structured account of how deterministic updating can yield apparent randomness under coarse-graining (Wolfram 2020; Gorard 2020a).

(5) **In security contexts, privilege adversarial evaluation over phenomenology.** 'Looks random' is

not a security criterion; follow standardised and peer-tested primitives and threat models (NIST 2015).

(6) **Use physical identity roots when replication is the relevant threat.** PUF-like constructions illustrate how determinism and practical unclonability can co-exist (Pappu et al. 2002).

(7) **Treat reversibility as an energy and correctness design axis.** Track where erasure occurs and whether uncomputation is available (Landauer 1961; Bennett 1973).

(8) **Operationalise 'reasoning' with verification when possible.** Formal theorem proving is a concrete arena for measurable reasoning progress and hybrid systems (Achim et al. 2025).

(9) **Exploit over-parameterise-then-prune as an optimisation tool.** Distinguish training-time capacity from deploy-time capacity (Frankle and Carbin 2019).

(10) **Treat unifying formalisms as languages, not guarantees.** Tensor-logic proposals may unify expression even when efficiency still depends on compilation and search (Domingos 2015; Domingos 2025).

8. Conclusion

Three methodological theses were recorded. First, determinism does not entail efficient predictability and computational irreducibility remains a live possibility (Wolfram 2002). Second, classification is often most informative when enriched by probabilistic and multiway formalisms, which represent graded reliability and observer-relative uncertainty (Mairesse and Marcovici 2014; Wolfram 2020). Third, the same distinctions recur in applications, including pseudorandomness, physically rooted identity and reversible computation (NIST 2015; Pappu et al. 2002; Landauer 1961).

References

[Achim et al. 2025] Achim, T., Best, A., Der, K., Fédérico, M., Gukov, S., Halpern-Leistner, D., Henningsgard, K., Kudryashov, Y., Meiburg, A., Michelsen, M., Patterson, R., Rodriguez, E., Scharff, L., Shanker, V., Sicca, V., Sowrirajan, H., Swope, A., Tamas, M., Tenev, V., Thomm, J., Williams, H. and Wu, L. 2025. Aristotle: IMO-level automated theorem proving. *arXiv* 2510.01346. `https://arxiv.org/abs/2510.01346`.

[Bennett 1973] Bennett, C. H. 1973. Logical reversibility of computation. *IBM Journal of Research and Development* 17: 525–32.

[Cook 2004] Cook, M. 2004. Universality in elementary cellular automata. *Complex Systems* 15: 1–40.

[Cybenko 1989] Cybenko, G. 1989. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems* 2: 303–14.

[Domingos 2015] Domingos, P. 2015. *The Master Algorithm*. New York: Basic Books.

[Domingos 2025] Domingos, P. 2025. Tensor logic: the language of AI. *arXiv* 2510.12269. `https://arxiv.org/abs/2510.12269`.

[Frankle and Carbin 2019] Frankle, J. and Carbin, M. 2019. The lottery ticket hypothesis: finding sparse, trainable neural networks. *International Conference on Learning Representations*. `https://arxiv.org/abs/1803.03635`.

[Gorard 2020a] Gorard, J. 2020a. Some quantum mechanical properties of the Wolfram model. *Complex Systems* 29: 537–98.

[Gorard 2020b] Gorard, J. 2020b. Some relativistic and gravitational properties of the Wolfram model. *Complex Systems* 29: 599–654.

[Landauer 1961] Landauer, R. 1961. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development* 5: 183–91.

[Mairesse and Marcovici 2014] Mairesse, J. and Marcovici, I. 2014. Around probabilistic cellular automata. *Theoretical Computer Science* 559: 42–72.

[NIST 2015] National Institute of Standards and Technology. 2015. *Secure Hash Standard (SHS)* (FIPS PUB 180-4). Gaithersburg, MD: NIST. `https://csrc.nist.gov/pubs/fips/180-4/final`.

[Pappu et al. 2002] Pappu, R., Recht, B., Taylor, J. and Gershenfeld, N. 2002. Physical one-way functions. *Science* 297: 2026–30.

[Turing 1936] Turing, A. M. 1936. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society* 42: 230–65.

[Wolfram 2002] Wolfram, S. 2002. *A New Kind of Science*. Champaign, IL: Wolfram Media.

[Wolfram 2020] Wolfram, S. 2020. A class of models with the potential to represent fundamental physics. *Complex Systems* 29: 107–536.