

Flavours of Self-Replication

Barbora Hudcová¹, Jordan Cotler², and Clément Hongler^{3,1}

¹ EPFL `barbora.hudcova@epfl.ch`

² Harvard University

³ Xent Labs

Abstract. Kleene’s recursion theorems imply that every Turing-complete language admits quines: programs that reproduce their own source code. Such programs have often served as a basis for examples of artificial self-replication. In contrast stands von Neumann’s seminal construction of a non-trivial self-replicator in a cellular automaton, a result that anticipated key abstract principles of biological replication before the discovery of DNA’s structure and helped establish cellular automata as a widely studied family of dynamical systems. This dichotomy raises fundamental questions about the nature of self-replication: what conditions must be satisfied for self-replication to be non-trivial? Does every Turing-complete system satisfy them, and hence automatically facilitate non-trivial self-replication? And in what sense is von Neumann’s self-replicator stronger than a quine? In this talk, I answer these questions based on joint work with Jordan Cotler and Clément Hongler [1].

Keywords: Self-replication · Turing completeness · Cellular automata

1 Introduction

A consequence of Kleene’s recursion theorems [6] is that, for every universal Turing machine $\mathcal{U}$, there exists a Turing machine $\mathcal{Q}$ such that $\mathcal{U}$ on input $([\mathcal{Q}], \emptyset)$ halts with output $([\mathcal{Q}], [\mathcal{Q}])$. We call $\mathcal{Q}$ a quine, and interpret it as a program that “self-replicates with respect to $\mathcal{U}$ ”. Such self-referential programs underlie many virtual worlds designed to study evolving self-replicators [4,5].

In contrast stands von Neumann’s work [2,3]: the design of a cellular automaton together with an initial pattern $\mathcal{P}$ in a quiescent background, whose evolution contains progressively more copies of $\mathcal{P}$ in disjoint regions. The pattern $\mathcal{P}$ encodes both a machine and its tape in cellular space. Crucially, von Neumann’s construction revealed the tape’s double use, for transcription and translation, before the discovery of DNA’s structure. The machine acting on the tape is both computation- and construction-universal, meaning it can construct arbitrary admissible patterns. These properties are typically used to argue for the non-triviality of von Neumann’s self-replicator (VNSR), in contrast to a simple replication process such as crystal growth.

Despite extensive work on CA self-replicators, until recently two key questions remained open:

- Q1** In what precise sense, if any, is von Neumann’s CA self-replicator stronger than a quine?
- Q2** Does every dynamical system capable of universal computation support non-trivial self-replication?

Here we build intuition for these questions and for the technical details we refer the reader to our recent paper [1].

2 From Self-Reference to Self-Replication

To our knowledge, every known self-replicator, biological or artificial, relies on two interacting layers: an informational layer, such as DNA, storing data needed for replication; and a functional layer, such as ribosomes and other cellular machinery, acting on this data to realize translation (constructing a new functional layer) and transcription (copying the informational layer). This dichotomy appears both in biological cell replication and in von Neumann’s work. In both cases, the layers interact through the underlying physical laws to produce a new organism containing a full copy of both layers.

For Turing machines, this separation is less apparent and raises the question of what should count as the organism and what as the physical laws. We use the classical interpretation of a Turing machine as a dynamical system with a “moving head”, whose position and state are explicitly represented in the tape configuration [7]. Under this interpretation, the tape content plays the role of the informational layer, while the moving head plays the role of the functional layer. Genuine self-replication should then require replication of both: without a copied head, the offspring cannot itself replicate. This is not supported by the standard Turing-machine architecture, but one can easily modify the dynamics so that, when a head halts, it creates a second head, with the two heads moving to their respective tape regions. In this sense, the system could be said to contain a genuine self-replicator. Yet this replication of the functional layer is fully “baked into the physical laws”, and so captures only half of the phenomenon. It differs sharply from von Neumann’s construction, where the local rule is chosen not to perform replication directly, but to support higher-level building blocks that constitute the functional layer. In VNSR, both layers are built from elementary parts, and their replication emerges at a higher level. This gives one answer to **Q1**: von Neumann’s self-replicator is stronger because it replicates not only information, but also the machinery that acts on it in a non-trivial way.

Turning to **Q2**, it remains a challenging open problem to formally characterize genuine self-replication. We therefore isolate a necessary condition.

Informal Definition 1. *A CA $\mathcal{A} = (S^{\mathbb{Z}^d}, F)$ with an admissible-background pattern $\mathcal{P}$ satisfies the necessary condition for self-replication if copies of $\mathcal{P}$ appear in a growing number of disjoint regions of the grid, and $\mathcal{P}$ undergoes nontrivial temporal evolution rather than remaining static.*

The precise definition is given in [1]. This condition already yields a negative answer to both **Q1** and **Q2** through the following theorem proved therein.

Theorem 1. *There exists a Turing-complete cellular automaton that does not satisfy the necessary condition for self-replication.*

Intuitively, this CA simulates a universal Turing machine, and hence contains quines, but does not support robust replication of the Turing-machine heads. The functional layer is too fragile to sustain genuine self-replication.

3 Future Work

We conclude with open problems, including the search for progressively stronger necessary conditions for self-replication, a precise distinction between replication emerging from elementary parts and replication “hard-baked into the dynamics”, and quantitative notions of replicator complexity.

Acknowledgments & Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article. The work was supported by NCCR SwissMAP.

References

1. Cotler, J., Hongler, C., Hudcová, B.: Self-replication and computational universality. arXiv preprint arXiv:2510.08342 (2025). <https://doi.org/10.48550/arXiv.2510.08342>
2. von Neumann, J.: The general and logical theory of automata. In: Jeffress, L.A. (ed.) *Cerebral Mechanisms in Behavior: The Hixon Symposium*, pp. 1–41. John Wiley & Sons, New York, NY (1951)
3. von Neumann, J.: *Theory of Self-Reproducing Automata*. Burks, A.W. (ed.) University of Illinois Press, Urbana (1966)
4. Ray, T.S.: An approach to the synthesis of life. In: Langton, C.G., Taylor, C., Farmer, J.D., Rasmussen, S. (eds.) *Artificial Life II*, pp. 371–408. Addison-Wesley, Redwood City, CA (1991)
5. Adami, C., Ofria, C., Collier, T.C.: Evolution of biological complexity. *Proceedings of the National Academy of Sciences* **97**(9), 4463–4468 (2000). <https://doi.org/10.1073/pnas.97.9.4463>
6. Kleene, S.C.: *Introduction to Metamathematics*. North-Holland, Amsterdam (1952)
7. Kůrka, P.: On topological dynamics of Turing machines. *Theoretical Computer Science* **174**(1–2), 203–216 (1997). [https://doi.org/10.1016/S0304-3975\(96\)00025-4](https://doi.org/10.1016/S0304-3975(96)00025-4)