Theorem (Finite Self-Reference Impossibility)

Author: Pedro R. Andrade
Date: 05JAN2026

Definitions

Definition 1 (Finite Dynamical System).
A *finite dynamical system* R is a system whose complete state space and evolution rule admit a finite description. Let $MDL(R)$ denote the minimum description length required to encode the full state, dynamics, and update rules of R.

Definition 2 (Internal Model).
An *internal model* $M(R)$ of a system R is any encoding instantiated entirely within R whose description length is bounded above by the description length of the system that instantiates it:

$$MDL(M(R)) < MDL(R)$$

Definition 3 (Lossless Self-Representation).
An internal model $M(R)$ is *lossless* iff the exact trajectory of R can be reconstructed from M(R) without access to any information external to R.

Theorem

Let R be a finite dynamical system capable of recursive self-reference.

There exists no internal model $M(R)$ such that $M(R)$ is both lossless and fully determines the behavior and evolution of R.

Proof

Assume, for contradiction, that such a model $M(R)$ exists.

Because $M(R)$ is lossless, it must encode sufficient information to reconstruct:

1. the complete state of R,
2. the dynamical evolution rule of R,
3. the conditions under which reconstruction occurs.

Therefore, the description length of $M(R)$ must satisfy:

$$MDL(M(R)) \geq MDL(R)$$

However, since $M(R)$ is an internal model instantiated within R, by Definition 2 it must satisfy:

$$MDL(M(R)) < MDL(R)$$

This is a contradiction.

Hence, no internal, lossless, complete self-representation of a finite self-referential system exists.

Immediate Consequence

All internal self-models of finite self-referential systems are necessarily partial, approximate, or lossy.

Interpretive Remark

This impossibility is structural, not epistemic. It follows solely from finiteness, internality, and self-reference, independent of substrate, semantics, or implementation.