Coherence-Rupture-Regeneration and Solomonoff Induction: A Mathematical Comparison

Mathematical Analysis

January 6, 2026

Abstract

This document presents a mathematical analysis comparing the Coherence-Rupture-Regeneration (CRR) framework with Solomonoff induction. We establish correspondence theorems between the two frameworks and identify what each provides. The main findings are: (1) CRR surprisal corresponds to accumulated conditional Kolmogorov complexity; (2) CRR rupture implements MDL model switching; (3) CRR regeneration implements a soft MDL evidence-weighting mechanism related to, but distinct from, Solomonoff's universal prior; (4) the frameworks are compatible, with CRR providing operational details that Solomonoff leaves unspecified. We carefully distinguish the pairwise switching threshold from the global temperature parameter.

Contents

1 Introduction

1.1 Scope

This document compares two frameworks for inductive inference:

(i) **Solomonoff Induction** (1964): A theoretical framework for sequence prediction based on algorithmic information theory and Kolmogorov complexity. It provides optimality guarantees but is incomputable.

(ii) **Coherence-Rupture-Regeneration (CRR)**: A framework describing how bounded systems maintain identity through discontinuous change, grounded in Bayesian model comparison.

We address three questions:

1. What is the mathematical relationship between CRR and Solomonoff induction?
2. Are the two frameworks compatible?
3. What does each framework provide that the other does not?

1.2 Summary of Findings

Main Findings

What Solomonoff Induction Provides:

- The optimal prior over hypotheses: $P(h) = 2^{-K(h)}$
- The optimal posterior: $P(h|y) \propto P(y|h) \cdot 2^{-K(h)}$
- Proof of optimality (dominates all computable predictors)

What Solomonoff Does Not Provide:

- When to act on accumulated evidence
- A computable implementation
- The mechanism of model change
- Tunable parameters for different contexts

What CRR Adds:

- Decision rule: switch when $\mathcal{S}_m - \mathcal{S}_{m'} > \Delta_{m,m'}$
- Computable approximation via temporal surprisal accumulation
- Regeneration operator with temperature parameter Ω
- Separation of switching threshold Δ from selection sharpness Ω

1.3 The Three-Framework Synthesis

Framework	Specifies	Mathematical Form
Solomonoff	What is optimal	$P(h) = 2^{-K(h)}$
Free Energy Principle	What to minimize	$F = \mathbb{E}[\log q] - \mathbb{E}[\log p]$
CRR	When and how	$\mathcal{S} \to \delta \to R$ with threshold Δ, temperature Ω

2 Preliminaries

2.1 Algorithmic Information Theory

Definition 2.1 (Kolmogorov Complexity)**.** The *Kolmogorov complexity* of a string x with respect to universal machine U is:

$$K_U(x) = \min\{|p| : U(p) = x\} \tag{1}$$

where $|p|$ denotes the length of program p in bits.

Definition 2.2 (Conditional Kolmogorov Complexity)**.** The complexity of x given y is:

$$K(x|y) = \min\{|p| : U(p, y) = x\} \tag{2}$$

Theorem 2.3 (Invariance Theorem)**.** *For any two universal machines U_1, U_2, there exists a constant c such that for all x:*

$$|K_{U_1}(x) - K_{U_2}(x)| \leq c \tag{3}$$

2.2 Solomonoff Induction

Definition 2.4 (Solomonoff Prior)**.** The *Solomonoff prior* over hypotheses (programs) is:

$$P_{\text{Sol}}(h) = 2^{-K(h)} \tag{4}$$

This assigns higher probability to simpler hypotheses.

Definition 2.5 (Solomonoff Posterior)**.** Given data y, the posterior over hypotheses is:

$$P_{\text{Sol}}(h|y) \propto P(y|h) \cdot 2^{-K(h)} = 2^{-K(y|h)} \cdot 2^{-K(h)} \approx 2^{-K(h,y)} \tag{5}$$

where the approximation uses the chain rule $K(h, y) = K(h) + K(y|h) + O(\log n)$.

Theorem 2.6 (Solomonoff Convergence)**.** *Let μ be any computable measure. The expected total squared prediction error is bounded:*

$$\mathbb{E}_\mu \left[\sum_{n=1}^{\infty} (P_{Sol}(x_{n+1}|x_{1:n}) - \mu(x_{n+1}|x_{1:n}))^2 \right] \leq K(\mu) \ln 2 \tag{6}$$

Remark 2.7 (Incomputability)**.** Solomonoff induction is incomputable because computing $K(x)$ requires solving the halting problem.

2.3 CRR Framework

Definition 2.8 (CRR System)**.** A *CRR system* is a tuple $(\mathcal{M}, Y, \Pi, \Delta, \Omega, \mathcal{S}, R)$ where:

- $\mathcal{M} = \{m, m', \ldots\}$ is a set of generative models
- $Y = \mathbb{R}^d$ is the observation space
- $\Pi : \mathcal{M} \to \mathrm{PD}(d)$ assigns precision matrices
- $\Delta : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$ is the **switching threshold** (pairwise, model-dependent)
- $\Omega > 0$ is the **temperature/rigidity parameter** (global)
- $\mathcal{S} : \mathcal{M} \times \mathbb{N} \to \mathbb{R}_+$ is the surprisal accumulator
- R is the regeneration operator

Remark 2.9 (Two Distinct Parameters)**.** We explicitly separate:

- $\Delta_{m,m'}$: The threshold for switching from model m to m', related to description-length overhead and prior odds
- Ω: A temperature-like parameter controlling how sharply regeneration weights historical states

These serve different functions and should not be conflated.

Definition 2.10 (CRR Operators)**.** The three CRR operators are:
Surprisal (Cost) Accumulation:

$$\mathcal{S}_m(n) = \frac{1}{2} \sum_{i=1}^{n} (y_i - g_m(\mu_i))^\top \Pi_m (y_i - g_m(\mu_i)) \tag{7}$$

This is a *cost*: higher values indicate worse model fit.
Rupture Condition:

$$\text{Switch } m \to m' \quad \text{when} \quad \mathcal{S}_m(n) - \mathcal{S}_{m'}(n) > \Delta_{m,m'} \tag{8}$$

Regeneration:

$$R[\phi](t) = \frac{1}{Z} \int_0^t \phi(\tau) \cdot \exp\left(-\frac{\mathcal{S}(\tau)}{\Omega}\right) \cdot \Theta(t - \tau)\, d\tau \tag{9}$$

Note the **negative sign**: states with *lower* surprisal (better fit) receive *higher* weight.

Remark 2.11 (Sign Convention)**.** The negative sign in the regeneration exponent is essential for consistency:

- $\mathcal{S}$ is a cost (higher = worse fit, analogous to $-\log P$ or K)
- $\exp(-\mathcal{S}/\Omega)$ gives higher weight to lower-cost states
- This aligns with Boltzmann weighting $\exp(-E/kT)$ where low energy is favored

3 Correspondence Theorems

3.1 Surprisal and Kolmogorov Complexity

Theorem 3.1 (Surprisal-Complexity Correspondence)**.** *Let $y_{1:n}$ be an observation sequence and m a computable generative model. Define the* algorithmic surprisal*:*

$$\mathcal{S}_m^{alg}(n) = \sum_{i=1}^{n} K(y_i | y_{<i}, m) \tag{10}$$

Then for CRR surprisal under Gaussian observations:

$$\mathcal{S}_m(n) = \mathcal{S}_m^{alg}(n) + O(n) \tag{11}$$

where the $O(n)$ term accounts for encoding overhead.

Proof. **Step 1:** By Levin's coding theorem, for any computable distribution P:

$$-\log P(x) = K(x) + O(K(P)) \tag{12}$$

Step 2: CRR surprisal under Gaussian observations equals negative log-likelihood:

$$\mathcal{S}_m(n) = -\log p(y_{1:n}|m) + \frac{n}{2}\log\det(2\pi\Sigma_m) \tag{13}$$

Step 3: Combining:

$$\mathcal{S}_m(n) = \sum_{i=1}^{n}[-\log p(y_i|y_{<i}, m)] + O(n) \tag{14}$$

$$= \sum_{i=1}^{n}[K(y_i|y_{<i}, m) + O(1)] + O(n) = \mathcal{S}_m^{\text{alg}}(n) + O(n) \tag{15}$$

□

Remark 3.2 (Temporal Survival as Complexity Proxy)**.** Models with low surprisal accumulation over time *behave as if* they have low Kolmogorov complexity, without computing K directly. This provides a computable proxy for the incomputable complexity measure.

3.2 Rupture and Minimum Description Length

Theorem 3.3 (Rupture as MDL Model Switching)**.** *The CRR rupture condition implements MDL model switching. Specifically, the switching threshold corresponds to:*

$$\Delta_{m,m'} = K(m') - K(m) + K(switch) \tag{16}$$

which is the description-length overhead for adopting model m' over m.

Proof. The MDL principle selects the model minimizing total description length:

$$m^* = \arg\min_m [K(m) + K(y_{1:n}|m)] \tag{17}$$

Model m' is preferred over m when:

$$K(m') + K(y_{1:n}|m') < K(m) + K(y_{1:n}|m) \tag{18}$$

Rearranging:

$$K(y_{1:n}|m) - K(y_{1:n}|m') > K(m') - K(m) \tag{19}$$

Substituting CRR variables via Theorem 3.1:

$$\mathcal{S}_m(n) - \mathcal{S}_{m'}(n) > \underbrace{K(m') - K(m)}_{\Delta_{m,m'}} \tag{20}$$

□

Remark 3.4 (Threshold is Model-Dependent)**.** The switching threshold $\Delta_{m,m'}$ depends on the complexity difference between models. This is *not* the same as the temperature parameter Ω.

3.3 Regeneration and Evidence Weighting

Theorem 3.5 (Regeneration as MDL Evidence Weighting)**.** *CRR regeneration implements a soft MDL evidence-weighting mechanism. For a hypothesis h evaluated on data y:*

$$w(h; y) \propto \exp\left(-\frac{\mathcal{S}_h(y)}{\Omega}\right) \approx \exp\left(-\frac{K(y|h)}{\Omega \ln 2}\right) = 2^{-K(y|h)/\Omega'} \tag{21}$$

where $\Omega' = \Omega \ln 2$.

*This weights hypotheses by how well they compress the data (the **likelihood/evidence term**), not by the prior complexity of the hypothesis itself.*

Proof. From Definition 2.10, regeneration weights by $\exp(-\mathcal{S}/\Omega)$.

By Theorem 3.1:

$$\mathcal{S}_h(y) \approx K(y|h) + O(n) \tag{22}$$

Therefore:

$$w(h; y) \propto \exp\left(-\frac{K(y|h) + O(n)}{\Omega}\right) \propto 2^{-K(y|h)/(\Omega \ln 2)} \tag{23}$$

□

Remark 3.6 (Distinction from Solomonoff Prior)**.** This is **not** the Solomonoff prior $2^{-K(h)}$ over hypotheses. The correspondence is:

Solomonoff	**CRR Regeneration**
Prior: $P(h) \propto 2^{-K(h)}$	Not directly represented
Likelihood: $P(y|h) \propto 2^{-K(y|h)}$	$w(h; y) \propto \exp(-\mathcal{S}_h(y)/\Omega)$
Posterior: $P(h|y) \propto 2^{-K(h,y)}$	Would require adding prior term

CRR regeneration captures the **evidence/likelihood** component of Bayesian inference, weighting hypotheses by how well they explain the data. To recover the full Solomonoff posterior, one would need to additionally weight by a prior term proportional to $2^{-K(h)}$ or equivalently $\exp(-K(h)/\Omega')$.

Corollary 3.7 (Full Posterior Recovery)**.** *To recover Solomonoff-like posterior weighting, CRR would need:*

$$w_{full}(h; y) \propto \underbrace{\exp\left(-\frac{K(h)}{\Omega'}\right)}_{\textit{prior term}} \cdot \underbrace{\exp\left(-\frac{K(y|h)}{\Omega'}\right)}_{\textit{CRR regeneration}} = \exp\left(-\frac{K(h, y)}{\Omega'}\right) \tag{24}$$

where the prior term could be implemented via the model prior $p(m)$ in the CRR system.

3.4 Parameter Identification

Proposition 3.8 (Temperature-Complexity Scaling)**.** *For CRR regeneration weights to match Solomonoff likelihood weighting exactly:*

$$\exp\left(-\frac{\mathcal{S}}{\Omega}\right) = 2^{-K(y|h)} \quad \Rightarrow \quad \Omega = \frac{1}{\ln 2} \approx 1.443 \text{ (in nats)} \tag{25}$$

Remark 3.9 (The $\Omega = 1/\pi$ Conjecture)**.** CRR documentation conjectures a universal value $\Omega = 1/\pi \approx 0.318$. This is **not** the same as the Solomonoff-matching value $1/\ln 2 \approx 1.443$.

If $\Omega = 1/\pi$, then:

$$w(h; y) \propto 2^{-K(y|h)\cdot\pi/\ln 2} \approx 2^{-4.53\cdot K(y|h)} \tag{26}$$

This represents *more aggressive* complexity penalization than standard Solomonoff weighting—a "colder" selection temperature.

Whether $\Omega = 1/\pi$ has deeper significance remains an open question, but it is mathematically distinct from Solomonoff correspondence.

4 What Each Framework Provides

4.1 What Solomonoff Provides

1. **Optimal prior:** $P(h) = 2^{-K(h)}$ over hypotheses (programs)
2. **Optimal posterior:** $P(h|y) \propto 2^{-K(h)} \cdot 2^{-K(y|h)}$
3. **Convergence guarantee:** Total prediction error bounded by $K(\mu)\ln 2$

4.2 What Solomonoff Does Not Provide

1. **When to act:** No decision rule for committing to a model
2. **Computability:** $K(h)$ is uncomputable
3. **Model change mechanism:** How one model replaces another
4. **Parameter flexibility:** Prior is fixed by choice of universal machine

4.3 What CRR Adds

1. **Decision rule:** Switch when $\mathcal{S}_m - \mathcal{S}_{m'} > \Delta_{m,m'}$
2. **Computable proxy:** Temporal surprisal accumulation approximates algorithmic complexity
3. **Regeneration mechanism:** $R[\phi] = Z^{-1}\int \phi \cdot \exp(-\mathcal{S}/\Omega)\, d\tau$
4. **Two tunable parameters:**
   - $\Delta_{m,m'}$: Switching threshold (model-dependent)
   - Ω: Selection temperature (global)

5 Compatibility Analysis

Theorem 5.1 (Compatibility)**.** *CRR and Solomonoff induction are mathematically compatible. CRR operationalizes aspects of Solomonoff for finite agents:*

Solomonoff	***CRR***
$K(y\|h)$ *(data complexity)*	$\mathcal{S}_h(n)$ *(accumulated surprisal)*
$K(m') - K(m)$ *(model complexity diff)*	$\Delta_{m,m'}$ *(switching threshold)*
Likelihood weighting $2^{-K(y\|h)}$	*Regeneration* $\exp(-\mathcal{S}/\Omega)$
Continuous mixture	*Discrete rupture at threshold*

Remark 5.2 (What CRR Does Not Capture)**.** CRR regeneration captures evidence weighting but not the Solomonoff prior $2^{-K(h)}$ directly. The prior enters CRR through:

- The model set $\mathcal{M}$ (which models are considered)
- The prior $p(m)$ (implicit in $\Delta_{m,m'} = \log(p(m)/p(m'))$ interpretation)

Remark 5.3 (One-Sentence Summary)**.** CRR describes how bounded agents can approximate Bayesian evidence weighting through temporal dynamics and threshold-triggered model switching, using surprisal accumulation as a computable proxy for algorithmic complexity.

6 Comparison Tables

Table 1: Parameter Roles

Parameter	**Role**	**Solomonoff Analogue**
$\Delta_{m,m'}$	Switching threshold; when to change models	$K(m') - K(m)$ (complexity overhead)
Ω	Temperature; how sharply to weight by fit	$1/\ln 2$ for exact correspondence

Table 2: What Each Framework Provides

Feature	**Solomonoff**	**CRR**
Prior over hypotheses $2^{-K(h)}$	Yes	Indirect (via $p(m)$)
Likelihood weighting $2^{-K(y\|h)}$	Yes	Yes ($\exp(-\mathcal{S}/\Omega)$)
Optimal prediction guarantee	Yes	No
Decision rule (when to switch)	No	Yes ($\mathcal{S}_m - \mathcal{S}_{m'} > \Delta$)
Computable	No	Yes (for finite $\mathcal{M}$)
Tunable parameters	No	Yes (Δ, Ω)

7 Limitations and Open Questions

7.1 Limitations of CRR

1. **No universal prior:** CRR regeneration captures evidence weighting but not the Solomonoff prior directly
2. **Model class restriction:** Requires specifying $\mathcal{M}$ in advance

3. **Parameter dependence:** Results depend on Δ and Ω, which must be determined
4. **No convergence guarantee:** Unlike Solomonoff, no proof that CRR converges to truth

7.2 Limitations of Solomonoff

1. **Incomputable:** Cannot be implemented
2. **No decision rule:** Provides probabilities but not when to act
3. **No temporal structure:** Does not address discrete change

7.3 Open Questions

1. Can CRR be extended to include a proper Solomonoff-like prior term?
2. Is there a principled derivation of Ω (independent of Solomonoff matching)?
3. How should $\Delta_{m,m'}$ be set in practice?
4. Does CRR with appropriate parameters converge to Solomonoff as $|\mathcal{M}| \to \infty$?

8 Conclusion

This analysis establishes:

1. **Partial correspondence:** CRR surprisal corresponds to $K(y|h)$; regeneration implements likelihood/evidence weighting $2^{-K(y|h)}$, not the full Solomonoff posterior
2. **Parameter separation:** The switching threshold Δ and temperature Ω serve distinct roles and should not be conflated
3. **Sign consistency:** Regeneration must use $\exp(-\mathcal{S}/\Omega)$ (negative exponent) for mathematical consistency
4. **Complementarity:** Solomonoff specifies *what* is optimal; CRR specifies *when* to switch and *how* to weight evidence

The relationship: *CRR implements the evidence-weighting component of Bayesian inference with explicit decision rules, using surprisal accumulation as a computable proxy for algorithmic complexity.*

References

[1] Solomonoff, R.J. (1964). A formal theory of inductive inference. *Information and Control*, 7(1), 1-22.

[2] Kolmogorov, A.N. (1965). Three approaches to the quantitative definition of information. *Problems of Information Transmission*, 1(1), 1-7.

[3] Levin, L.A. (1974). Laws of information conservation. *Problems of Information Transmission*, 10(3), 206-210.

[4] Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, 14(5), 465-471.

[5] Wald, A. (1947). *Sequential Analysis.* John Wiley & Sons.

[6] Hutter, M. (2005). *Universal Artificial Intelligence.* Springer.

[7] Grünwald, P.D. (2007). *The Minimum Description Length Principle.* MIT Press.

A Technical Notes

A.1 Levin's Coding Theorem

Theorem A.1 (Levin)**.** *For any computable probability distribution P and string x:*

$$-\log P(x) = K(x) + O(K(P)) \tag{27}$$

A.2 Chain Rule for Kolmogorov Complexity

Theorem A.2 (Chain Rule)**.**

$$K(x, y) = K(x) + K(y|x^*) + O(\log K(x, y)) \tag{28}$$

where x^* is the shortest program for x.

A.3 Why the Negative Sign Matters

In the original CRR formulation, if coherence C is defined as accumulated prediction error (a cost), then weighting by $\exp(+C/\Omega)$ would favor *worse*-fitting states.

The correct form is either:

- Define C as negative surprisal (higher = better fit), weight by $\exp(+C/\Omega)$
- Define $\mathcal{S}$ as surprisal/cost (higher = worse fit), weight by $\exp(-\mathcal{S}/\Omega)$

This document uses the second convention for clarity.