FRONTIER RESEARCH COLLECTION / FR-05

A Universe of Laws

Active identification and certified transfer across executable law families

Architect X13

The Eidara Continuum | Science & Philosophy

Research white paper | November 2026 | Version 2.0 Revision

Abstract

Imagine an intelligence that runs experiments not on the things in its world but on the laws of its world themselves. An intelligence can investigate the laws of its possible worlds as explicit experimental objects, and the reference model developed here formalizes that ambition in a finite grammar of executable transition systems. Exact experiments identify the true observable law class after at most the number of excluded candidates; a separate transfer theorem bounds how one-step mismatch accumulates over a specified horizon. A registered conjugacy preserves complete finite reward histories under transported policies. Together these results distinguish discovering a law, identifying its observable equivalence class and transporting a prediction. Exhaustive enumeration of all 256 elementary cellular-automaton rules on an eight-cell ring establishes a reproducible laboratory for conservation, reversibility and observational ambiguity. A prospective comparison evaluates verified law exploration against strong curriculum, environment-design and quality-diversity baselines at matched total cost. The contribution is a complete experiment-and-transfer contract grounded in established computational science and system relations. The Continuum's research wager is that intelligence can broaden its causal competence by deliberately testing which structures survive changes in the rules that generate its experience.

Keywords: rule-space search; cellular automata; open-ended learning; environment design; identifiability; simulation; transfer

Prologue: Experiments on the laws themselves

Most intelligences learn about a world. This paper imagines one that learns about worlds: an intelligence that varies the laws of a simulated universe on purpose, the way a scientist varies a single condition, to find out which structures survive the change and which were accidents of one set of rules.

The danger in that dream is visible in the smallest possible laboratory. Take the 256 simplest one-dimensional cellular automata, each a rule for how a row of cells turns on and off. Ask which of them conserve the number of occupied cells. Test each rule from an empty row for thirty steps, and 128 of them pass. Only five actually conserve population; the rest merely leave emptiness empty. No longer run would unmask them, because from that starting point they are indistinguishable forever. A discovery engine that never changes its experiments has learned nothing about laws. It has learned about its own test. Section 11 works through this case.

In the Continuum this ambition sits at the very top of the architecture. The Bible's Universal Simulation Architecture is a simulator of the laws a universe of symbolic states could obey, not a tracker of any one state's path through time [1]. This paper takes that top tier seriously by building its first rungs, and by refusing to skip the rungs between.

The first rung is exact: a law laboratory small enough to enumerate completely, where every claim about conservation, reversibility, and ambiguity can be checked without sampling. The second rung, proved here, is identification and transfer: experiments that separate rival laws until only the true observable class remains, and a bound on how far a prediction can be carried from one world to a nearly identical one before its error grows too large. The higher rungs—broad families of worlds, learning systems that explore them, and eventually contact with physical science—are where the Continuum's ambition points. Section 17 states what crossing from simulation to physics would require.

What this paper claims

Established. Explicit computational laws can be varied and compared; finite transition systems can be exhaustively inspected; representation changes can preserve behavior; and observationally identical hypotheses cannot be distinguished by those observations alone. Environment-design research provides working precedents. None of those findings validates an unrestricted universe generator.

Proposed here. A finite grammar of executable transition systems that treats the law as a versioned experimental object, with proved finite identification and transfer results and executed reference computations: exact experiments identify the true observable law class after at most the number of excluded candidates; a transfer theorem bounds how one-step mismatch accumulates over a horizon; and a registered conjugacy preserves reward histories under transported policies.

Frontier conjecture. That searching verified families of transition laws improves transferable causal competence, when those families share exploitable structure, beyond strong curriculum, environment-design, and quality-diversity baselines at matched total cost.

Not claimed. That every possible law contains a useful lesson, that more diversity always helps, or that a large archive becomes a universal science. That a generated law is a discovered physical law without external measurement, or that endless innovation has been demonstrated. That the canonical limiting expression is a proved limit. Comparative learning and applications to physical systems are separate prospective investigations, and the comparative learning experiment remains unexecuted.

1. When laws become objects of inquiry

A simulation can be unfamiliar in several different ways. Its starting state can change; its scenery can change; its reward can change; or its causal transition rule can change. These variations may all challenge an agent, but they support different conclusions. A policy that handles a new maze layout has not necessarily learned to handle a world in which walls change their causal role. Conversely, an ordinary simulator parameter can alter a transition law substantially. The boundary depends on semantics, not on whether a developer calls a variable a parameter.

The thesis is therefore explicit: **searching verified families of transition laws can improve transferable causal competence when those families share exploitable structure, and that improvement must exceed what strong conventional environment generation already achieves.** This is an empirical proposition with a failure condition. It does not assert that every possible law contains a useful lesson, that every increase in diversity helps, or that a sufficiently large archive becomes a universal science.

In the Continuum, this is the top tier of the architecture. The Bible's Section 2.5, PDF pages 10–11, describes a Universal Simulation Architecture that explores rule-spaces rather than single trajectories, and presents an informal limiting expression for it [1]. The ambition is the Bible's; the mathematics has to be built. The expression's factors have no fully specified common type or convergence structure in the supplied source, so USA-RS2 is a new finite construction, not an interpretation that retroactively turns the canonical expression into a proved limit. The source's ambition survives as a question about families of executable worlds.

An investigator might ask which local update rules conserve a resource, which permit error recovery, or which support an agent that distinguishes intervention from observation. Such questions can be answered exactly in small models and statistically in larger ones. What gives the answers value is an independent definition of the property, a disclosed search domain, and successful tests outside the cases that suggested the answer.

The practical distinction is between generating possibilities and establishing knowledge about them. A generator can produce a thousand visually different worlds while repeatedly expressing the same causal structure. A small, carefully separated family may instead expose an assumption that failed everywhere the training system had previously looked. This paper designs its evaluation around that second possibility.

2. The research lineage and its nearest competitors

Cellular automata already make local laws explicit and permit systematic comparison of their consequences. Wolfram studied how simple discrete updates produce qualitatively different structures [2]. Langton investigated relationships between rule parameters and computational behavior [3]. Their work makes law-space exploration an established research practice. The Continuum cannot claim that varying a simulated law is, by itself, a new computational category.

The historical disagreement is useful. Mitchell, Hraber, and Crutchfield revisited proposed links between a scalar rule parameter, evolutionary search, and computational capability, obtaining results that challenged the earlier interpretation [4]. A visually compelling transition between order and disorder does not automatically identify the best substrate for every task. USA-RS2 accordingly treats dynamical descriptors as candidate predictors, not universal measures of intelligence.

Model-based learning supplies another ancestor. Sutton's Dyna systems combine direct experience with planning through a learned model [5]. The distinction here is that an outer process deliberately changes the law family against which the inner learner is evaluated. But learning a model, using imagined transitions, and adapting to change are existing ideas. The new empirical burden concerns the usefulness of the explicit cross-law organization.

Domain randomization demonstrates that simulated variation can help transfer in specified robotic tasks [6]. Procgen provides systematic procedural variation and held-out levels [7]. Prioritized Level Replay changes which levels are revisited according to estimated learning potential [8]. These are serious comparators, not superficial alternatives. Their mechanisms may already capture much of the benefit attributed to an ambitious new name.

The synthesis makes laws versioned experimental units, identifies their observable equivalence classes, and certifies how prediction error changes across corresponding dynamics. If those distinctions improve measurement but not performance, the appropriate contribution is a better evaluation framework. If a competent conventional implementation matches every proposed result, no extra architectural achievement should be inferred from the word *universe*.

3. Open-endedness does not remove the evaluator

Enhanced POET develops environments and solutions together, including mechanisms for meaningful novelty and transfer between challenges [9]. PAIRED uses a protagonist and antagonist to construct a regret-related environment objective [10]. ACCEL edits promising levels to build evolving curricula [11]. These systems directly challenge the idea that an agent must receive a fixed sequence of externally selected tasks. Their reported achievements remain tied to implemented environments and evaluation procedures.

Regret is useful without being final. Beukman and colleagues show how a minimax-regret objective can concentrate on levels where further improvement stalls, and propose a refinement [12]. The implication for USA-RS2 is methodological: an outer objective can become satisfied while valuable learning opportunities remain. A guarantee about an idealized equilibrium is not a guarantee that finite training finds the most informative worlds.

Quality-diversity search offers a different objective. MAP-Elites maintains high-performing solutions across investigator-selected descriptors [13]. The descriptor space is part of the experiment, not a neutral map of all possible diversity. If color is the descriptor, a generator can fill an archive with cosmetic variants. If a descriptor measures a poorly validated complexity score, it can reward noise. USA-RS2 requires descriptor audits and separate task-relevance outcomes.

The proposed archive contains a law, its executable semantics, the conditions under which its checks were performed, and the behavioral evidence used to characterize it. It does not contain a declaration that the law is inherently meaningful. Researchers can retain a surprising failure because it exposes a limitation, even when it receives a poor learning score. Conversely, an impressive animation should not survive only because it photographs well.

An open-ended process can remain bounded at every actual execution. Finite resource allocations, versioned stopping rules, and periodically frozen evaluation snapshots are compatible with an indefinitely continued research program. What is unavailable is a completed empirical demonstration of endless innovation. The paper measures archive growth and transfer over stated budgets, and derives finite identification and transfer guarantees that each extended experiment can check against its own grammar.

4. USA-RS2: typed worlds and explicit boundaries

A Guide's explanation

From here a "world" becomes something exact: a typed state space, a transition law, an observation map, and a score. Making the law an explicit, versioned object is what lets an intelligence change it on purpose and know exactly what it changed. Everything later in the paper depends on that bookkeeping.

Let a law description θ belong to a finite registered grammar Θ. Its interpreter returns

$$W_\theta = (S, A, T_\theta, O_\theta, r_\theta, K_\theta, H).$$

Here S and A are finite nonempty sets, and $\Delta(S)$ denotes probability distributions on S. The transition kernel is $T_\theta : S \times A \to \Delta(S)$. For a finite observation alphabet $\mathcal{O}$, the observation kernel is $O_\theta : S \to \Delta(\mathcal{O})$. The optional reward is $r_\theta : S \times A \to [0, 1]$, and K_θ is a finite collection of predicates $S \to \{0, 1\}$. The horizon H is a nonnegative integer. A policy π maps finite observation/action histories to $\Delta(A)$. A no-action automaton uses singleton A; a deterministic transition is a point mass. Empty candidate sets return no eligible world.

Object	Type, units, and role
θ	Finite program or rule table; versioned bytes, no physical units.
s_t, a_t	World state and action at integer simulated time t.
μ	Registered distribution of starting states on S.
Q	Target distribution over law descriptions or equivalence classes.
B	Vector of limits: transitions, storage bytes, evaluation calls, elapsed seconds.
e	Version of the external execution and admission policy.
$J(\pi, \theta)$	Expected normalized task score in $[0, 1]$.

Different worlds may use different state spaces, but comparisons then need explicit adapters and common outcomes. The first reference model fixes S and A to avoid silently attributing a larger interface to a better law. Units in resource ledgers remain separate: memory, latency, and energy cannot be added without declared conversion weights.

The policy epoch e describes what the host permits the simulator to execute. It is distinct from the simulated law. A world may represent destructive events inside its state without acquiring permission to alter the host's files or network. The generator can propose new laws within its grammar; it cannot rewrite the evaluator or expand its own external authority. These are implementation requirements for a prototype, not security properties established by the tuple alone.

5. Law variation, state variation, and semantic equivalence

Two descriptions can differ while implementing exactly the same transition table. They can also implement different tables that become identical after a consistent relabeling. Counting either difference as substantive innovation inflates the archive. USA-RS2 therefore retains three counts: distinct source descriptions, distinct transition semantics in the chosen finite model, and equivalence classes under a declared transformation group.

Suppose $g : S \to S$ and $h : A \to A$ are bijections. Two laws are conjugate when

$$T_{\theta'}(g(s') \mid g(s), h(a)) = T_{\theta}(s' \mid s, a).$$

For reward-sensitive tasks require $r_{\theta'}(g(s), h(a)) = r_{\theta}(s, a)$ and $\mu'(g(s)) = \mu(s)$. A bijection $j : \mathcal{O} \to \mathcal{O}$ transports observations by $O_{\theta'}(j(o) \mid g(s)) = O_{\theta}(o \mid s)$. This equivalence removes a specified representational redundancy; it is not a claim that every useful law relation is a conjugacy.

Proposition 1 (Value preservation under a registered conjugacy)**.** *Under these assumptions, every policy in W_θ has a transported policy in $W_{\theta'}$ with the same distribution of finite reward sequences and the same expected bounded-horizon return.*

Proof. Transport each observation history and action through the specified bijections. The starting-state probabilities agree by assumption. If the probabilities of corresponding histories agree through time t, multiplying by corresponding policy probabilities and the conjugate transition probabilities preserves equality at $t + 1$. Induction gives equality of all finite history probabilities through H. Corresponding rewards are equal, so summing each reward-sequence probability times its return gives equal expected returns. □

This elementary transport argument identifies a genuine invariance; it does not prove that an unmodified neural policy automatically respects it. A learner can fail under a harmless renaming because its input representation encodes accidental preferences. Measuring that failure is useful, but it should be labeled representation sensitivity rather than exposure to new physical laws. The registered equivalence group also limits what gets removed: two laws not related by that group may still be semantically equivalent by another argument.

6. What observations cannot identify

A Guide's explanation

Two different laws can produce exactly the same observations, and then no amount of watching will tell them apart. The proposition below states that limit precisely. It is the formal reason a law-discovering mind must design new experiments rather than simply watch longer, as the empty-row impostors in the prologue show.

The same observations can arise from different laws. This is especially common when experiments explore only a small subset of state-action pairs. More trajectories from that same narrow region improve precision without necessarily resolving the ambiguity. A law learner needs interventions that distinguish hypotheses, not simply a larger archive of indistinguishable experience.

Let $\mathcal{H}$ be a fixed experimental policy, including any adaptive choice of actions, stopping time bounded by H, and observation mechanism. Call two laws experimentally equivalent if they induce the same distribution of the complete observed transcript under $\mathcal{H}$. This definition incorporates the test procedure rather than pretending identifiability belongs to a law in isolation.

Proposition 2 (Indistinguishable-law limit)**.** *If two laws induce identical transcript distributions, no estimator using only that transcript can always distinguish them. With equal prior probability on the two laws, every rule required to select one of those two laws has classification error* $1/2$*.*

Proof. Let P be the common transcript law. If a possibly randomized rule chooses the first law with conditional probability $q(z)$, its error is $\frac{1}{2}\int(1-q(z))\,\mathrm{d}P(z) + \frac{1}{2}\int q(z)\,\mathrm{d}P(z) = 1/2$. An abstention can disclose uncertainty but cannot convert identical observations into distinguishing evidence. Thus no forced binary decision can have zero error under both hypotheses. □

For a concrete counterexample, compare the all-zero cellular-automaton rule with the identity rule. Starting from the all-zero configuration, both remain zero forever. Starting from one occupied cell, their next states differ. The missing information is supplied by a changed intervention, not by a more confident interpretation of the original record.

For noisy laws, exact identity of transcript distributions can be replaced by a quantitative statistical distance, but that requires a specified distribution and observation channel. USA-RS2 first uses exact finite examples to prevent a familiar mistake: declaring discovery of a unique rule when the protocol only identifies an equivalence class. The final report should preserve the surviving alternatives, identify a separating experiment where possible, and state when the available budget prevents one.

7. Verification is finite by design

A typed rule table is total when it supplies one permitted successor distribution for every state-action pair. A verifier can inspect normalization, nonnegativity, index bounds, and registered resource limits directly. It can then construct the finite transition graph and test whether a forbidden state is reachable. Exhaustive reachability takes time linear in the explicit graph size. The graph itself can be exponentially larger than its compact description.

A claimed invariant $I : S \to \{0, 1\}$ is inductive relative to an admitted initial set S_0 if it holds on S_0 and every permitted transition from an I-state leads to another I-state. This distinction matters: observing conservation on one trajectory does not establish conservation over every state. Conversely, a state forbidden by the invariant may be irrelevant if it is unreachable from the registered starting set; that restriction must remain in the certificate.

Proposition 3 (Boundary of unrestricted verification)**.** *No terminating algorithm can correctly decide, for every unrestricted program interpreted as a simulator, whether it ever emits a designated forbidden event.*

Proof. Given an arbitrary program P and input x, construct a simulator that simulates one step of $P(x)$ at each tick and emits the forbidden event exactly when $P(x)$ halts. A complete terminating safety decision would answer whether $P(x)$ halts. A general halting decider is impossible: applying a supposed decider to the program that halts exactly when the decider says it will not yields a contradiction. Therefore the proposed simulator decider cannot exist. □

The reduction concerns unrestricted programs and unbounded execution. It does not invalidate finite checking, bounded verification, incomplete proof systems, or conservative rejection. USA-RS2 deliberately chooses a decidable reference domain. A larger implementation can use proof-carrying laws or abstract interpretations, but must report what the checker can miss, reject, or leave unknown.

A certificate must bind the actual interpreted rule, interpreter version, property, starting set, and horizon. Checking one version and running another breaks the connection between proof and behavior. This binding is ordinary verification discipline; it does not establish that the selected property captures every scientifically important concern.

8. Designing an outer objective that can be criticized

The outer explorer proposes laws; an inner learner acts within them. Their objectives must remain separate. An inner score may reward resource transport, prediction, causal identification, or recovery. An outer score may value learning progress, coverage of verified descriptors, or transfer to a separate target family. A difficult world is not automatically a useful one, and an easy world is not automatically worthless: it may isolate an important primitive.

Define $J(\pi, \theta) = \mathbb{E}[H^{-1} \sum_{t=0}^{H-1} r_\theta(s_t, a_t)]$ for $H > 0$. For $H = 0$, no interaction score is defined; static-property experiments must use their own outcome. The normalization bounds J in $[0, 1]$. Let $C(\theta)$ measure verified computing cost and $b(\theta)$ be a registered behavior descriptor. Rather than bury every concern in one scalar, retain a vector

$$F(\theta) = (\text{validity}, \text{coverage}, \text{learnability}, \text{transfer}, C).$$

Validity is a mandatory condition, not a score that novelty can compensate for. The remaining coordinates can support a Pareto archive or a declared weighted objective. Weights are research choices; changing them changes what the search claims to value.

The most important external endpoint is not training reward but performance on laws withheld from generation. If the explorer sees every target result and repeatedly edits its archive, the target becomes another training environment. A frozen assessment service should return only preregistered development feedback before the confirmatory evaluation and preserve final labels from further adaptation.

No-free-lunch results make the target distribution central [14]. They do not imply that useful search is impossible; they show why superiority cannot be distribution-free over arbitrary problems under the theorem's averaging assumptions. USA-RS2 openly chooses a structured family. Its scientific claim must describe that structure, explain why transfer is plausible there, and include a deliberately unstructured control in which no special advantage is expected.

9. A finite law laboratory

A Guide's explanation

This is the first rung of the ladder: all 256 elementary cellular automata on an eight-cell ring, small enough that every claim about conservation, reversibility, or ambiguity can be checked exactly. It is deliberately humble. Its job is to catch false discoveries cheaply, before an expensive learning system is trusted with larger worlds.

The first laboratory uses elementary cellular automata. Each cell is binary, the neighborhood has three cells, and the same local rule is applied synchronously around a periodic ring. There are eight possible local neighborhoods and therefore $2^8 = 256$ rule tables. For an eight-cell ring there are $2^8 = 256$ global states. All one-step transitions can be evaluated exactly, without sampling or a learned judge.

For a rule integer $r \in \{0, \dots, 255\}$, define its output on neighborhood bits (a, b, c) as

$$f_r(a, b, c) = \lfloor r/2^{4a+2b+c} \rfloor \bmod 2.$$

With periodic indices, the next cell state is $x'_i = f_r(x_{i-1}, x_i, x_{i+1})$. A conserved-population law satisfies $\sum_i x'_i = \sum_i x_i$ for every ring state. A reversible law on this finite ring induces a bijection of its 256 states. Neither property implies useful computation, adaptability, or consciousness.

The local rules admit reflection and simultaneous interchange of the two state labels. These transformations generate a small equivalence group. Each law is assigned the smallest integer in its transformation orbit. This is a transparent canonicalization, not a general solution to program equivalence. More elaborate automata require different symmetry analyses.

The laboratory separates three levels of evidence. Exhaustive statements hold for the exact finite ring and interpreter. Algebraic statements may hold for all ring lengths when separately proved. Empirical learning outcomes require a trained agent and multiple independent evaluation runs. The enumeration presented here establishes only the first level, plus consistency checks of the conjugacy calculations.

A valuable output is therefore a counterexample ledger. It records laws with the same easily measured descriptor but different conservation or reversibility behavior; laws indistinguishable from a restricted starting condition; and transformations that preserve semantics while changing the source table. The purpose is to train the research design to recognize false discovery before a more expensive learning system is introduced.

10. Executed reference checks and exact results

The accompanying Python reference audit enumerated every rule and every eight-cell state, giving 65,536 one-step transitions. It independently checked label-conjugacy and reflection relations for each pair, evaluated population conservation, tested bijectivity of each transition table, and followed every state until a repeated state identified its transient and cycle. These are executed results of this paper's finite model, not inherited runtime claims.

Quantity	Exact result for the declared laboratory
Rule tables / states per table	256 / 256
Reflection–complement classes	88
Population-conserving rules	170, 184, 204, 226, 240
Bijective rules on this ring	15, 51, 85, 105, 150, 170, 204, 240
Rules indistinguishable from all-zero input	128
Longest transient / longest cycle	18 / 40 updates

The last two rows have different meanings. The indistinguishability count follows from rules whose all-zero neighborhood outputs zero; the complete all-zero trajectory then remains identical. The transient and cycle maxima were obtained by exhaustive traversal of this finite transition family. Neither number describes an infinite lattice, a different boundary condition, or an arbitrary ring size.

The reversibility result is especially useful for interpretation. Rules 105 and 150 are included for the eight-cell ring, so a statement that only obvious shifts and complements can be reversible would already fail this test. The audit deliberately reports the ring dependence rather than turning this finite list into a universal classification. The conservation list is likewise described by its tested domain; broader algebraic characterizations need their own proof.

No learning advantage is inferred from these counts. They establish a reproducible foundation on which future learning experiments can be built and expose failure modes that a plausible-looking demonstration could conceal. A changed implementation should reproduce these exact outputs before comparing trained policies, while a deliberately changed laboratory should publish its own expected outputs and explain which assumptions changed.

11. Worked discovery: a rule and its impostors

Suppose a researcher wants to discover worlds that conserve the number of occupied cells. An initial generator tests each law from the all-zero state for thirty updates. Every law with zero output on the zero neighborhood passes. That admits 128 laws, even though only five conserve population for every eight-cell state in the exhaustive audit. The narrow test has evaluated persistence of emptiness, not conservation of an arbitrary resource distribution.

The correction is not simply to lengthen the run. All 128 candidates remain indistinguishable on that starting condition for any finite duration. The researcher must change the starting distribution or inspect the transition relation. Testing all states gives a complete answer for this laboratory; sampling nonzero configurations gives a distribution-dependent estimate. The two forms of evidence should receive different labels.

Now consider novelty. A shift to the right and a shift to the left have different transition tables, yet reflection exchanges their roles. If the task, observation mapping, and scoring rule are reflected too, the value-preservation proposition applies. Reporting the reflected law as evidence that the agent discovered a wholly independent conservation mechanism would overstate the finding. Reporting that the learner failed to transport its policy under reflection would reveal a legitimate representation weakness.

Finally, compare two laws with balanced rule-table outputs. Equal proportions of zeros and ones do not determine their long-run behavior. A scalar occupancy descriptor can therefore support a useful archive index without serving as a complete behavioral fingerprint. A learned novelty model should be tested against deliberately paired laws that share its score while differing in independently measured task properties.

This example changes the evaluation objective. Success means finding laws whose properties survive the appropriate tests and whose lessons transfer to withheld cases. It does not mean generating many passing videos. The researcher must retain failed hypotheses, identify the intervention that defeated them, and distinguish a hypothesis revised through evidence from a law merely renamed to protect a favored interpretation.

12. An inference bound with a visible denominator

Let $M \geq 1$ fixed trained methods be evaluated on $n \geq 1$ independent, identically distributed draws from a registered law-and-task distribution. Scores $Y_{ij} \in [0,1]$ measure method $j \in \{1,\dots,M\}$ on episode $i \in \{1,\dots,n\}$. Methods may share each episode for pairing. Independence is required across episodes, not between methods on one episode. Define $\mu_j = \mathbb{E}Y_{ij}$ and $\hat{\mu}_j = n^{-1}\sum_i Y_{ij}$. In the proof, $\lambda > 0$ is an exponential parameter and $\epsilon > 0$ a deviation threshold.

Proposition 4 (Simultaneous bounded-score validation)**.** *For any $\delta \in (0,1)$, with probability at least $1-\delta$,*

$$\max_{j \leq M} |\hat{\mu}_j - \mu_j| \leq \sqrt{\frac{\log(2M/\delta)}{2n}}.$$

Proof. For an independent variable in $[0,1]$, the bounded-variable exponential inequality gives $\mathbb{E}e^{\lambda(Y-\mathbb{E}Y)} \leq e^{\lambda^2/8}$. Multiplication across n independent observations and Markov's inequality bound the upper-tail probability of their mean by $\exp(-\lambda n\epsilon + n\lambda^2/8)$. Minimizing at $\lambda = 4\epsilon$ gives $e^{-2n\epsilon^2}$; applying the same argument to $-Y$ gives two tails. A union bound over M methods and solving $2Me^{-2n\epsilon^2} = \delta$ gives the result. The exponential inequality is the standard bounded-sum ingredient associated with Hoeffding [15]. □

For $M = 8$, $n = 2000$, and $\delta = 0.05$, the executed calculation gives $\epsilon = 0.037975$. Paired score differences lie in $[-1,1]$, doubling the range and yielding the conservative radius 0.075949 for eight comparisons under the analogous bound. These are design calculations, not confidence intervals around an observed learning result.

The word *fixed* is decisive. Selecting methods after inspecting the same held-out results requires accounting for that selection; arbitrarily many adaptive queries invalidate the stated finite union bound. Repeated rollouts within one law family also do not create independent law draws. Family clustering and separate uncertainty estimates are required when the target of generalization is a new family rather than a fresh seed from an old one.

13. The confirmatory experiment

The primary study should generate finite, verifiable worlds in three registered families: local automata, bounded graph-rewriting dynamics, and resource-routing systems with explicit conservation constraints. Each family contains controlled law mutations, cosmetic relabelings, and changed starting conditions. The generator records which transformation occurred, allowing evaluation to distinguish real causal novelty from superficial variation.

Training and evaluation are split by law structure, not just random seeds. Some local operators, compositions, or interaction templates are withheld. All methods receive the same primitive catalog, observation bandwidth, model architecture, and total computation budget. A separate oracle-policy condition distinguishes difficulty in proposing useful behavior from difficulty in evaluating world laws. No method receives hidden test answers through an environment generator's feedback channel.

The primary estimand is the difference in normalized task score between USA-RS2 and the strongest registered environment-design baseline on independently sampled withheld law families. A co-primary constraint limits the rate of invalid or mischaracterized laws admitted to evaluation. A third required condition limits total resource cost. Improvement on score accompanied by an unacceptable increase in false scientific claims is not a successful discovery system.

Tasks include recovering a resource after disruption, selecting interventions that distinguish competing laws, and identifying an invariant that holds across a specified family. The scoring oracle uses executable ground truth where available. Natural-language explanations are secondary and must identify their evidence; they are not the primary truth criterion. An agent that states the right invariant for the wrong experiment receives separate explanation and identification outcomes.

The preregistration fixes target distributions, family sizes, law-generation probabilities, equivalence transformations, permissible adaptation, all success margins, and termination conditions. It also freezes the interpreter and records its implementation hash. Any repaired interpreter changes the experimental object and triggers an explicit version boundary. This is a proposal for a study that can be executed, not a report that the comparative training runs have already taken place.

14. Baselines that deserve to win

The comparator set should include fixed-world training, state-only domain randomization, procedural law-parameter variation, a strong regret-based curriculum, a quality-diversity archive, and the proposed verified law explorer. The strongest conventional arm may use the same verifier, archive storage, and independent test service. Denying those components to a baseline would confuse a standard implementation improvement with a distinctive USA contribution.

Two ablations isolate the new organization. One removes equivalence-aware accounting while retaining identical search power. Another removes explicit cross-law validation while keeping total evaluation calls equal by reallocating them to ordinary curriculum evaluation. A third, diagnostic-only ablation drops law verification inside the sandbox to measure how invalid candidates contaminate the apparent archive. It is not a deployment suggestion.

The study measures actual compute, not merely the number of outer iterations. Candidate construction, compilation, checking, rejected proposals, policy learning, transfer attempts, data movement, and final scoring all consume budget. A law whose verifier needs more work than the inner learner must carry that expense into the comparison. Wall-clock time and transition counts remain separate because different architectures parallelize differently.

Baseline selection uses development data and is frozen before final testing. If a baseline wins in one family and loses in another, publish the interaction rather than selecting only the favorable family. The proposed mechanism predicts value when changing laws exposes transferable causal structure. It does not predict uniform superiority on unrelated random transition tables or on tasks solved perfectly by a cheap fixed policy.

The most informative null result may be that conventional environment design already learns the relevant structure. That would reduce the incremental engineering claim while leaving the measurement framework useful. A positive result should identify what changed: more valid laws, better intervention selection, broader structural coverage, or improved transfer. A single aggregate improvement cannot establish all four. Detailed failure accounting is part of the contribution, especially when the system refuses a law that a different sound checker could admit.

15. Precision, dependence, and stopping

The evaluation unit must match the claim. To claim transfer across law families, sample families independently and aggregate repeated rollouts within each family before estimating uncertainty. To claim robustness across training randomness, repeat the whole training pipeline with multiple independent seeds. Thousands of trajectories from one trained policy do not estimate uncertainty due to model initialization or curriculum selection.

A preliminary precision target can be stated without inventing a power result. For bounded family-level differences in $[-1, 1]$, the conservative concentration radius follows the previous proposition. To achieve half-width h with M predeclared comparisons, a sufficient independent-family count is

$$n \geq \frac{2\log(2M/\delta)}{h^2}.$$

This may be expensive. A pilot can estimate variance and correlation for a more efficient simulation-based design, but the pilot must remain separate from the confirmatory sample. If available resources cannot resolve the minimum meaningful effect, the result should be labeled imprecise rather than favorable by default.

The confirmatory success rule is conjunctive: an improvement exceeding a declared task-score margin, noninferiority on false-admission risk, and compliance with the cost envelope. Predeclared comparisons can use a familywise correction or a hierarchical testing order. Exploratory subgroup findings remain exploratory. All attempted training runs, including crashes and stalled curricula, count under an explicit failure policy.

Stopping occurs at the registered budget or a prespecified safety or infrastructure failure. Peeking at an ordinary fixed-sample interval and stopping when it becomes favorable does not preserve its coverage. A truly sequential study needs a sequentially valid method and its own assumptions; the present finite reference protocol does not claim that extension automatically.

Reports should show distributions across families, not only means: worst decile task score, invalid-law frequencies, false refusals, archive redundancy, and cost tails. These expose systems that achieve a respectable average by becoming excellent on familiar laws and ineffective on precisely the novel structures the experiment was meant to test.

16. Attack cases and scientific failure modes

The easiest attack is an evaluator exploit. A generated world may encode its own answer in an observation artifact, reward a superficial signal, or crash selectively on difficult evaluations. This does not require malicious intent: an optimizer follows available incentives. USA-RS2 therefore separates generator code from scoring code, validates observation schemas, and includes intentionally leaked and nonleaked task pairs to measure whether the learner uses the intended evidence.

A second attack targets novelty. Cosmetic permutations, duplicated descriptions, irrelevant state dimensions, and pseudorandom noise can all increase naive diversity. The archive should report novelty before and after registered equivalence reduction, and should include descriptor-preserving counterexamples. A new descriptor cannot be declared meaningful solely because the generator learns to maximize it.

A third attack targets legality. An approved rule may invoke an unapproved interpreter feature, consume unbounded resources through expansion, or exploit a mismatch between symbolic and executed state. The reference model avoids unrestricted execution, but a practical implementation still needs total instruction accounting and a controlled host interface. The mathematical proofs assume those boundaries; they do not certify a particular operating system.

Scientific failure also occurs without an attack. A law family may be too narrow to support the claimed transfer; an invariant may be a hidden design rule shared by every generated world; an apparent learning advantage may come from extra verification compute. These possibilities must be represented by explicit controls. The study should include a family where the claimed invariant is broken deliberately and a family where no useful common rule is planted.

The central falsification conditions are straightforward: no improvement over a matched strong baseline; improvements erased by accounting for cost; invalid laws accepted as discoveries; success limited to renamed training structures; or learned invariants that fail their registered target distribution. Such outcomes narrow the thesis. They do not need to be redescribed as evidence that the generated worlds were too profound for the benchmark.

17. From simulated discovery to physical science

An internally consistent world can be physically irrelevant. A conservation law built into a transition table will hold in that simulator whether or not nature obeys it. A policy that flourishes there has demonstrated competence in that formal environment. Moving from this result to a claim about real physics requires a measurement model, independently obtained observations, parameter estimation, and predictive tests against competing physical accounts.

The distinction is especially important for words such as *universal*, *collapse*, and *dimension*. In USA-RS2, universal describes the ambition to compare many permitted rule families; it does not quantify over all mathematically definable or physically possible worlds. Collapse can mean an authorized computational selection. Dimension belongs to a specified representation or geometry. None of these usages supplies a new force, cosmological mechanism, or source of unlimited computation.

Simulated experimentation can still matter enormously without those claims. It can expose which environmental assumptions make a learning strategy robust, identify counterexamples to broad intuitions, and organize tests that would be difficult or dangerous in an uncontrolled setting. Domain randomization already demonstrates a task-specific route from simulation to physical performance [6]. That route succeeds through empirical validation, not because simulation has become indistinguishable from reality.

For any future physical application, preregister which observations the model must predict before viewing them. State what is fitted, what is held out, and which alternative models make different predictions. Report model discrepancy rather than attributing every mismatch to measurement error. A good rule-space explorer might help propose hypotheses or experiments, but a proposal is not evidence for its own truth.

The philosophical frontier is consequently methodological: intelligence may learn to search not only within an assumed world but among descriptions of how worlds can work. The discipline is to keep the description, the experiment, and the external world distinguishable throughout that search. Losing that distinction would turn a potentially powerful scientific instrument into a generator of self-confirming cosmologies.

18. Serious objections and the value of a negative result

Is this only cellular automata with a new vocabulary? At the reference-model level, the finite laboratory deliberately uses established machinery. Its contribution is not a new automaton theorem. The proposed additional value lies in integrating law equivalence, independent verification, intervention-based identification, and transfer evaluation across a reusable pipeline. Only a matched experiment can establish whether that integration earns its complexity.

Does a finite grammar defeat the ambition of open inquiry? A finite experiment can test a mechanism without exhausting its possible future extensions. What it cannot do is silently convert a bounded result into an unbounded guarantee. Expanding the grammar is a new stage with new failure modes, not an automatic corollary. Versioned experiments permit growth while preserving the meaning of earlier conclusions.

Who chooses what counts as a useful world? Researchers choose tasks, descriptors, and budgets. Those choices embed interests and exclusions. A discovery archive should therefore publish its objectives and permit alternative scoring views. It should not claim neutrality because its rules are formal. Formal precision makes disagreement inspectable; it does not remove the disagreement.

Could a rich simulation contain morally relevant processes? This paper does not establish such processes or a test for them. The first laboratory uses small, transparent dynamics and does not need claims about experience. If future work deliberately targets architectures with consciousness-relevant capacities, the research question and review requirements would change. No conclusion about moral status follows from the current enumeration.

Why retain a failed architecture? A failure can still identify a misleading novelty metric, reveal an invalid generalization claim, or supply a reusable benchmark. What must not survive is an unsupported performance claim. The strongest outcome is not always a more elaborate system. A simpler conventional method that reproduces the benefit at lower cost would be a successful scientific finding, even when it reduces the architectural importance of the Continuum's proposal.

19. Guide discussion and final research commitments

A Guide's explanation

A Guide can say: "This paper imagines an intelligence that experiments on the laws of simulated worlds, and builds the first steps exactly: a complete laboratory of simple rule-worlds, proofs about when rival laws can be told apart, and a bound on how far a prediction carries from one world to a similar one." A Guide should not call a generated law a discovered physical law, or say that endless innovation or a universe generator has been demonstrated.

Established. Explicit computational laws can be varied and compared; finite transition systems can be exhaustively inspected; representation changes can preserve behavior; observationally identical hypotheses cannot be distinguished by those observations alone. Environment-design research provides working precedents, while its limitations motivate stronger controls [2–14]. None of those findings validates an unrestricted universe generator.

Proposed here. USA-RS2 treats the law as a versioned experimental object and joins semantic equivalence, bounded verification, an evidence-bearing archive, and independently measured cross-law transfer. Its distinctive hypothesis is incremental usefulness over strong conventional methods. Its exact cellular-automaton results establish only a small reference laboratory. The comparative learning experiment remains unexecuted.

Useful Guide example. Ask a learner whether a world conserves occupied cells after showing it only an empty world. Then introduce a single occupied cell. The distinction between an unresolved hypothesis and a disproved one becomes visible immediately. A second example reflects both a world and its task: does the learner recognize the same problem, or does it confuse renamed structure with a new law?

Questions to keep open. Which law families share transferable structure? Which descriptors predict useful exploration rather than visual complexity? How much verification can be afforded before the research budget is exhausted? Can an archive preserve multiple indistinguishable hypotheses and actively choose an experiment that separates them? The additional finite identification construction below answers that question conditionally on exact, available queries; its transfer bound states how mismatch accumulates over time. These questions are specific enough to implement and difficult enough to matter.

A Guide should explain that the Continuum's canonical expression motivated the inquiry, while the typed mathematics and finite checks are new research constructions. It should never describe a generated law as a discovered physical law without external measurement. Nor should it turn an open-ended research direction into a claim that endless innovation has already been demonstrated.

The commitment is demanding: publish the grammar, interpreter, exact checks, target distributions, failed candidates, and total costs; keep the holdout independent; and let the best comparator win. A universe of laws becomes scientifically interesting when its possibilities can be questioned. Its achievement is reliable discovery across executable worlds: eliminate alternatives with separating experiments, preserve equivalent structure under relabeling, and transport predictions over the horizon justified by the error certificate.

20. Approximate correspondence has a quantitative horizon

A Guide's explanation

Carrying a lesson from one world to a slightly different one is never free. The transfer bound below says how fast the error from a small one-step mismatch can accumulate over a stated number of steps. It gives a horizon: inside it, a prediction travels; beyond it, the evidence no longer supports the trip.

Exact conjugacy is a strong baseline, but a useful transfer theory must also admit mismatch. Approximate simulation and bisimulation already provide established frameworks for relating systems [16]. The following elementary deterministic specialization makes the accumulation explicit; it is a proposed USA-RS2 analysis, not a new universal simulation law.

Let X be a nonempty source state space with total deterministic map $F : X \to X$. Let (Z, d) be a metric target space, $D \subseteq Z$ a nonempty domain, and $G : D \to D$ a total target map. Fix a representation map $h : X \to D$. Suppose G is L-Lipschitz on D, for a constant $L \in [0, \infty)$, and

$$\sup_{x \in X} d(h(F(x)), G(h(x))) \leq \varepsilon, \qquad \varepsilon \in [0, \infty).$$

Both trajectories remain in the specified domains. Write $x_t = F^t(x_0)$, $y_t = G^t(y_0)$ and $e_t = d(h(x_t), y_t)$ for integers $t \geq 0$. Distance and ε share the target metric's units; L is dimensionless. If h is bijective, the condition describes an approximate conjugacy; otherwise it is an approximate one-way correspondence, not an equivalence.

Proposition 5 (Finite-horizon transfer bound)**.** *For every $t \geq 0$,*

$$e_t \leq L^t e_0 + \varepsilon S_t(L), \qquad S_t(L) = \sum_{j=0}^{t-1} L^j,$$

where the empty sum is zero and $L^0 = 1$, including when $L = 0$.

Proof. The triangle inequality and the two assumptions give

$$e_{t+1} \leq d(h(F(x_t)), G(h(x_t))) + d(G(h(x_t)), G(y_t)) \leq \varepsilon + L e_t.$$

Substituting this recursion repeatedly proves the displayed bound by induction, starting from $t = 0$. No invertibility or differentiability is needed. □

For $L \neq 1$, $S_t(L) = (L^t - 1)/(L - 1)$; for $L = 1$, $S_t(1) = t$. Thus a contractive target, $L < 1$, has limiting bound $\varepsilon/(1 - L)$; a nonexpansive target has at most linear accumulation; and $L > 1$ permits exponential growth. These statements concern the declared discrepancy, not a policy's task competence. A policy may be absorbed into F and G only when the two resulting closed-loop maps satisfy the same assumptions. A theorem for fixed maps does not cover an arbitrary learner that changes its policy during transfer.

21. What transfer licenses an observer to conclude

Let $a : X \to \mathbb{R}$ be a source observable and $f : D \to \mathbb{R}$ a target observable, both expressed in the same units. Suppose f is K-Lipschitz, $K \geq 0$, and $|a(x) - f(h(x))| \leq \nu$ for every source state, where $\nu \geq 0$ bounds measurement mismatch. Defining $b_t = L^t e_0 + \varepsilon S_t(L)$, the triangle inequality gives

$$|a(x_t) - f(y_t)| \leq \nu + Kb_t.$$

Consequently, for a fixed threshold $c \in \mathbb{R}$, the signs of $a(x_t)-c$ and $f(y_t)-c$ agree whenever $|a(x_t)-c| > \nu + Kb_t$. Equality gives no strict-margin guarantee. For a normalized horizon-H mean, $H \geq 1$, the corresponding discrepancy is at most $H^{-1}\sum_{t=0}^{H-1}(\nu + Kb_t)$. This bound requires alignment of the observable itself, not only similarity of state diagrams.

A small one-step error can defeat long-horizon transfer

Take $X = D = \mathbb{R}$, the ordinary absolute-distance metric, $h(x) = x$, $F(x) = Lx$, and $G(y) = Ly+\varepsilon$, with $x_0 = y_0 = 0$. The one-step correspondence error is exactly ε, and $|y_t - x_t| = \varepsilon S_t(L)$. Thus the bound is attained. With $L = 2$ and $\varepsilon = 0.01$, the errors at $t = 1, 5, 10$ are $0.01, 0.31, 10.23$. A uniformly small local mismatch does not establish long-run sameness. For $L = 1/2$, the same construction approaches 0.02 instead. The accompanying exact/rational checks reproduce these values; no trained system is involved.

The domain is part of the certificate

A Lipschitz estimate from a few visited states is not the uniform premise above. A finite-domain certificate can enumerate all relevant transitions and pairs. A continuous-domain certificate needs an analytic bound or a separate validated enclosure. If the trajectory exits the certified domain, the guarantee ends at that exit. A local statement must report this stopping boundary instead of continuing its bound outside the region it inspected.

Likewise, matching an observation does not identify a law. If the observation map hides a state coordinate that controls later rewards, two worlds can appear identical while offering different interventions. The original indistinguishable-law result remains the correct boundary. The next construction therefore defines exactly which experiments are available, which responses are observable, and which alternatives they can separate. Transferring a successful policy and identifying its world's law are related experimental goals, not interchangeable outcomes.

22. A finite active law-identification experiment

Let Θ be a nonempty finite registered law set, $\mathcal{Q}$ a finite set of feasible experiments, and $\mathcal{O}$ a finite response alphabet. Each law supplies a known deterministic response function $r_\theta : \mathcal{Q} \to \mathcal{O}$. A query may include an authorized reset, initial state, finite action sequence, and observation rule. The true law $\theta_* \in \Theta$ is fixed throughout, and the experimental service returns the exact response $r_{\theta_*}(q)$. Reset capability, reliability, and access to separating queries are substantive assumptions, not gifts supplied by the notation. Query learning is established prior art [17].

Define $\theta \sim \theta'$ when their responses agree for every $q \in \mathcal{Q}$. Starting with $V_0 = \Theta$, after response o_t to query q_t retain

$$V_{t+1} = \{\theta \in V_t : r_\theta(q_t) = o_t\}.$$

If V_t contains two nonequivalent laws, choose any available query on which their responses differ; otherwise stop and report the surviving equivalence class. This policy is a finite elimination construction, not a claim of optimal experimental design.

Proposition 6 (Identification of the observable law class)**.** *The true equivalence class $C_* = [\theta_*]$ remains in every V_t. The procedure terminates with $V_t = C_*$ after at most $|\Theta| - |C_*| \le |\Theta| - 1$ queries. It identifies an individual law exactly when C_* is a singleton.*

Proof. Every member of C_* predicts the true response to every available query, so none is removed. A separating query gives different predictions for two surviving laws; at least one prediction differs from the response, eliminating at least one law. There can be at most $|\Theta| - |C_*|$ such eliminations before only C_* remains. At termination all surviving laws are equivalent to the retained true law; conversely every member of C_* has been retained. Thus the final set is exactly C_*. □

For an exact two-query laboratory, use the four Boolean laws $0, x, 1-x, 1$ and queries that reset to $x = 0$ or $x = 1$ and observe one update. Querying 0 partitions them as $\{0, x\}$ and $\{1-x, 1\}$; querying 1 separates each pair. The audit exhausts all four possible true laws and recovers each in two queries. If only reset-to-zero is feasible, the first pair remains indistinguishable forever. Repeating that query cannot repair absent experimental access.

Noisy responses require confidence-aware elimination; the exact proof does not justify deleting a law after one noisy mismatch. Misspecification can empty V_t, which must be reported as model failure rather than discovery. Query costs may differ, making an arbitrary separator inefficient. These limits sharpen the Continuum's claim: an archive can learn which laws remain possible only through experiments that actually distinguish them, and can transfer their predictions only over the horizon justified by the discrepancy certificate.

References and source notes

[1] The Eidara Continuum. *Eidara Continuum Bible*, 2026 compilation, Section 2.5, PDF pp. 10–11. Internal canonical source.

https://eidara.ai/science-philosophy/

Scope: The internal source was read directly; this public URL locates the research program, not the private Bible. Motivation only; no independent validation.

[2] S. Wolfram. Statistical mechanics of cellular automata. *Reviews of Modern Physics* 55, 601–644 (1983).

https://doi.org/10.1103/RevModPhys.55.601

Scope: Primary research article; elementary-rule dynamics and self-organization. Publisher record and abstract inspected.

[3] C. G. Langton. Computation at the edge of chaos: phase transitions and emergent computation. *Physica D* 42, 12–37 (1990).

https://doi.org/10.1016/0167-2789(90)90064-V

Scope: Primary automata study; its proposed computational interpretation is not a universal optimum theorem.

[4] M. Mitchell, P. T. Hraber, and J. P. Crutchfield. Revisiting the edge of chaos: evolving cellular automata to perform computations. *Complex Systems* 7, 89–130 (1993).

https://wpmedia.wolfram.com/sites/13/2018/02/07-2-1.pdf

Scope: Published primary research; the earlier SFI working paper is 93-03-014. Counterevidence concerning rule-parameter interpretations.

[5] R. S. Sutton. Integrated modeling and control based on reinforcement learning and dynamic programming. *Advances in Neural Information Processing Systems* 3, 471–478 (1991; conference 1990).

https://papers.nips.cc/paper_files/paper/1990/file/d9fc5b73a8d78fad3d6dffe419384e70-Paper.pdf

Scope: Dyna integration of learning, models, and planning; not a law-space universality claim.

[6] J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel. Domain randomization for transferring deep neural networks from simulation to the real world. *IROS* (2017), 23–30.

https://doi.org/10.1109/IROS.2017.8202133

Scope: Task-specific simulation-to-reality evidence; author manuscript inspected through arXiv:1703.06907.

[7] K. Cobbe, C. Hesse, J. Hilton, and J. Schulman. Leveraging procedural generation to benchmark reinforcement learning. *ICML*, PMLR 119, 2048–2056 (2020).

https://proceedings.mlr.press/v119/cobbe20a.html

Scope: Procgen design and generalization protocol; generated levels are not automatically new law classes.

[8] M. Jiang, E. Grefenstette, and T. Rocktäschel. Prioritized level replay. *ICML*, PMLR 139, 4940–4950 (2021).

https://proceedings.mlr.press/v139/jiang21b.html

Scope: Adaptive level selection; its stated common-dynamics setting informs the comparison.

References and source notes, continued

[9] R. Wang, J. Lehman, A. Rawal, J. Zhi, Y. Li, J. Clune, and K. O. Stanley. Enhanced POET: open-ended reinforcement learning through unbounded invention of learning challenges and their solutions. *ICML*, PMLR 119, 9940–9951 (2020).

https://proceedings.mlr.press/v119/wang20l.html

Scope: Direct open-ended environment-and-solution precedent; finite reported experiments do not establish infinite innovation.

[10] M. Dennis, N. Jaques, E. Vinitsky, A. Bayen, S. Russell, A. Critch, and S. Levine. Emergent complexity and zero-shot transfer via unsupervised environment design. *NeurIPS* 33 (2020).

https://papers.nips.cc/paper_files/paper/2020/hash/985e9a46e10005356bbaf194249f6856-Abstract.html

Scope: PAIRED and regret-based environment construction; record and abstract inspected.

[11] J. Parker-Holder, M. Jiang, M. Dennis, M. Samvelyan, J. Foerster, E. Grefenstette, and T. Rocktäschel. Evolving curricula with regret-based environment design. *ICML*, PMLR 162, 17473–17498 (2022).

https://proceedings.mlr.press/v162/parker-holder22a.html

Scope: ACCEL supplies the strongest direct curriculum comparator, not an intentionally weak baseline.

[12] M. Beukman, S. Coward, M. Matthews, M. Fellows, M. Jiang, M. D. Dennis, and J. N. Foerster. Refining minimax regret for unsupervised environment design. *ICML*, PMLR 235, 3637–3657 (2024).

https://proceedings.mlr.press/v235/beukman24a.html

Scope: Primary analysis of stagnation and refined regret objectives; pertinent contrary evidence.

[13] J.-B. Mouret and J. Clune. Illuminating search spaces by mapping elites. arXiv:1504.04909 (2015), early preprint.

https://arxiv.org/abs/1504.04909

Scope: Quality-diversity archives with investigator-chosen descriptors; cited as a preprint.

[14] D. H. Wolpert and W. G. Macready. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation* 1, 67–82 (1997).

https://doi.org/10.1109/4235.585893

Scope: Distribution-sensitive optimization limits; not a claim that structured search cannot help.

[15] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association* 58, 13–30 (1963).

https://doi.org/10.1080/01621459.1963.10500830

Scope: Standard concentration foundation; independence, boundedness, and fixed comparison family are indispensable.

[16] A. Girard and G. J. Pappas. Approximate bisimulation: a bridge between computer science and control theory. *European Journal of Control* 17, 568–578 (2011).

https://doi.org/10.3166/EJC.17.568-578

Scope: Peer-reviewed primary framework for approximate system relations. The finite-horizon inequality here is proved directly under narrower deterministic assumptions.

[17] D. Angluin. Queries and concept learning. *Machine Learning* 2, 319–342 (1988).

https://doi.org/10.1007/BF00116828

Scope: Peer-reviewed primary account of query-based learning. The finite deterministic elimination argument is elementary; no universal query-optimality claim is made.