
Learnable Boundaries, Unlearnable Competence: A Characterization of Learned Components in a Symbolic Search Stack

Eduardo Perez Denadai
Optimality.ai

Abstract

We ask a question that a pinned symbolic authority forces on any neurosymbolic system: if a machine already *knows* — through an exact, provably sound decision procedure — where its competence ends, what is left for a neural network to do? Beginning with the authority’s *decision boundary*, we find the boundary is **learnable**: a small relational classifier rediscovers it and transfers to worlds it has never seen. Boundary competence, however, is not task competence: the same learning that captures *whether a step is safe* cannot capture *whether a goal is reachable*, and no amount of depth repairs that — the wall is one of **learnability**, not capacity. Turning to the one object a containment theorem actually licenses — a learned **heuristic** that merely *orders* an exact search — we evaluate a sequence of mechanisms. The apparent progress — a certified search that beats the classical heuristic while never loosening the guard, first at $0.82\times$ then $0.72\times$ the classical node count — does not survive scrutiny. On hard maze-like maps a *strong* classical heuristic beat the learned one; and when we compared at **matched search weight** — the control the earlier comparisons had missed — the learned heuristic’s marginal value was, in every regime, **one** (and worse on mazes). The entire “win” was the **greedy weight**, a classical mechanism, not the network. The mechanism’s real product is therefore not speed but a **certified bounded-suboptimal dial**: a classical weight you *set* from the suboptimality you will accept, whose worst case is known by theorem, on a search the proposer **cannot** loosen. The learned part adds nothing here; the **guarantee** is everything. One wall is real — and it is the **inductive bias** of the learned function, not learnability itself: a differentiable **relaxation** whose bias *is* the global computation crosses it ($0.93\times$ a tight classical heuristic on hard mazes), where local and message-passing arms fail. Beyond it the paper’s contribution is a **method** (change one axis at a time; watch which one the ceiling moves with) and a **product** that cannot fail with the proposer: a **sound, contained** learned proposer inside a certifying search.

1 Introduction

This is the third paper in a series on the Ω -TD neurosymbolic architecture. The first [Denadai, 2026b] introduced the architecture itself — a pinned authority, a guard, and abstention — and its model-free proposer: a learner contained so that it may choose within the guard but never enlarge it. The second [Denadai, 2026a] generalized it to belief-space planning, where composition of macro-actions emerges from the belief and the exact search. This paper characterizes what the learned components of that stack can and cannot do — which parts of the competence a network can learn, and where the wall lies.

Neurosymbolic systems usually begin with a hand-designed symbolic layer and ask what a network can add. We began from the opposite end. Our starting object was a **pinned authority**, Ω , that

decides over a set-valued belief whether a step is *certified* (R_0), *refuted* (R_3), or an *honest abstain* with a named **witness** — the cell whose observation would settle the matter (R_1). Ω is sound by construction: its verdicts are a *function* of the belief, and it says “I don’t know” rather than guess.

1.1 The framework, stated

Three facts about this architecture are assumed throughout, and we state them here so the paper is self-contained.

Containment. Execution selects from the authority’s admissible set plus an abstention, $\{a : \text{guard}(s, a) = 1\} \cup \{\text{ABSTAIN}\}$. A learned component may *choose within* that set but may never *enlarge* it, so for an arbitrary proposer — a network, a search, or an adversary — no execution is unsafe. Safety is structural and does not depend on the proposer’s quality.

The price of certification. The authority does not make search free. A certifying search expands exactly the states whose estimate could still undercut the true optimum, $g(n) + h(n) < C^*$ — the **f-contour** $\mathcal{F}(h)$ — so the cost of soundness is the volume of that contour under the heuristic h . A learned h can beat a classical h_c only by reducing its error; a larger error expands no fewer nodes. This is why the only learned object the framework licenses is a heuristic that *orders* an exact search.

What is left for learning. If the authority already decides the boundary, the network’s job cannot be to recompute it — it must be either (i) to *predict* it cheaply, which is only worthwhile if the prediction is itself learnable, or (ii) to *order* the search the authority certifies, which is only worthwhile if the learned heuristic reduces error. This paper measures both.

An authority like this draws a sharp line. Everything it can already compute is not the network’s to compute; only what it *declines* is interesting. So we asked a narrow, falsifiable question: **is the shape of that line — the boundary between certified, refuted, and abstained — something a network can learn?** And if so, does knowing the boundary buy anything the exact procedure does not already have?

We took the problem out of the abstract and put it on **MiniGrid**, the standard partially-observed gridworld. One methodological decision anchored everything that follows: we refused to trust our own environment. Our first apparatus used a hand-built MiniGrid *subset*, and when we later held it against the real minigrid package (minigrid_parity), it failed the test — its observation model was a Bernoulli sample over world cells, not MiniGrid’s egocentric, wall-occluded view. We certified the discrepancy, rebuilt on the real environment, and **re-ran every affected study** on real MiniGrid. Every number below is from that reconstruction.

2 The decision boundary is learnable

We begin with the direct test. If Ω ’s verdicts are a function of the belief, then a classifier trained on those verdicts might reproduce them. We trained a *linear* head on a handful of topology-invariant relational primitives — the status of the cell ahead, whether it is known, whether a path of known-passable cells connects the agent to the goal — and tested it on **topologies it had never seen**.

It transferred completely. On never-trained families the head reached **full macro-recall** (1.000) for reachability, while every representation that threw away relational structure collapsed: indexing cells by absolute position fell to 0.519/0.465, and scrambling the features per example fell to 0.404/0.500 (transition/reachability; Fig. 2). The boundary, we concluded, is **distillable**, and the ingredient that matters is the *representation*, not the weights — the same lesson the parent chapters had learned about their own encoders, now reproduced on the boundary itself.

It was a satisfying result and, as it turned out, a misleading one — not because it was wrong, but because we had mistaken one success for the whole story. The temptation at that point is to declare the boundary a solved, transferable “safety output.” The next question is what saves you from that.

3 Boundary competence is not task competence

We replaced the *hand-designed* primitives with a **learned relational encoder on raw nodes and edges** — cells, the agent, the exterior, and their adjacency — and asked whether it would rediscover

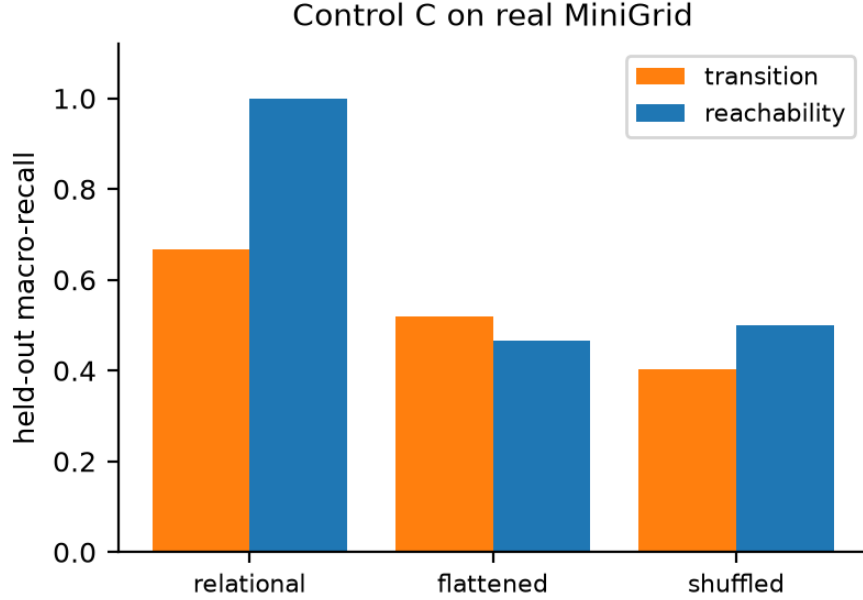


Figure 2: Control C on real MiniGrid. A linear head on relational primitives transfers fully; position-keyed and shuffled controls collapse.

the boundary from the data alone, on topologies it had never seen. The predicate we asked it to learn came in two kinds, and the split between them is the turning point of this paper.

On the **local** transition predicate — *is the cell in front of me blocked?* — the encoder succeeded: 0.995 held-out, at the ceiling the hand-designed representation reaches. But on the **global** reachability predicate — *is the goal connected to me through cells I have seen to be passable?* — it did not learn at all. Its held-out score, 0.705, sat *within noise of the per-example-shuffled control* (0.773; Fig. 1). By the falsifier we had fixed in advance — *a predicate is learned only if it beats the shuffled control* — the global predicate was **not learned**. It was, in the strict sense, not there.

The reading we took from this was not that the network was weak, but that we had been asking the wrong object to do the wrong job. **Learning supplies the local, shallow, order-valued structure; composition is the search’s**. The boundary and the engine are different things, and a system that confuses them will over-claim the former and under-use the latter.

4 The wall is learnability, not capacity

The natural objection is capacity: perhaps we simply had too small a network, or too little depth. And here the environment offered a clean, non-obvious test. Reachability is *s-t connectivity*, an inherently **global** predicate; the local transition predicate, by contrast, is a function of a **constant** number of bits — we verified this by constructing a **constant-depth circuit** that reproduces Ω ’s transition labels with **zero mismatches**. So the local wall is shallow by nature, and the global wall might just be *deep*.

If the global wall were only depth, it should yield to depth. It does not. We gave the encoder a number of message-passing layers well beyond the grid’s diameter — more than enough, and we *proved* sufficiency by construction, since a depth-equal-to-eccentricity relaxation solves connectivity exactly. The learned encoder then did what a capacity model would not: it **fit its training topologies perfectly (train accuracy 1.000) and still failed to generalize connectivity** (held-out near chance; Fig. 5b). Adding depth did not move the line. The wall was not capacity; it was **learnability** — a distinction that took an explicit escape route to expose.

The same wall, in a different light. The parent chapters had bounded this from the other side: *s-t* connectivity on these graphs is L -complete, while fixed-depth, log-precision transformers and message-passing networks live inside uniform TC^0 ; under the standard (unproven) conjecture that

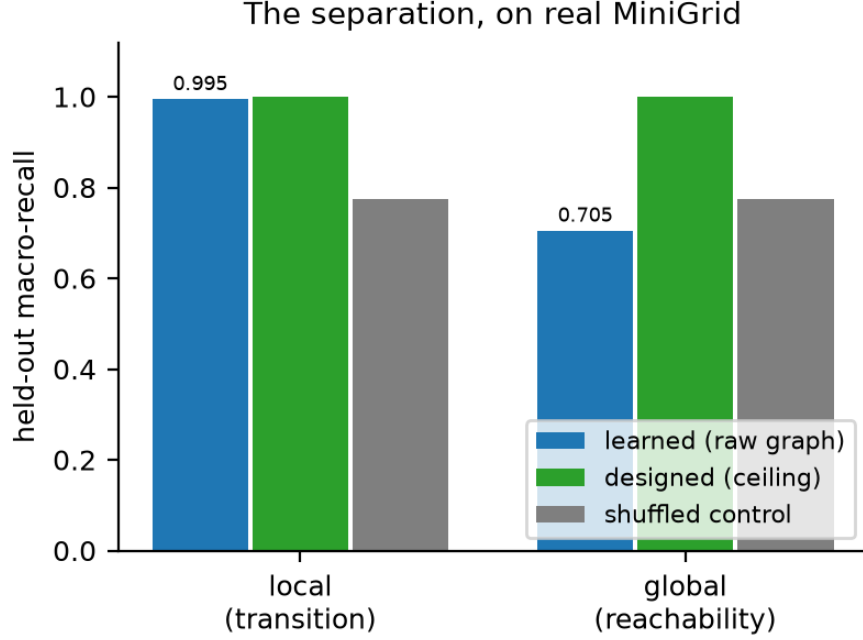


Figure 1: The separation, on real MiniGrid. The local predicate is learned and transfers (0.995, at the designed ceiling); the global predicate is not ($0.705 \approx$ shuffled 0.773).

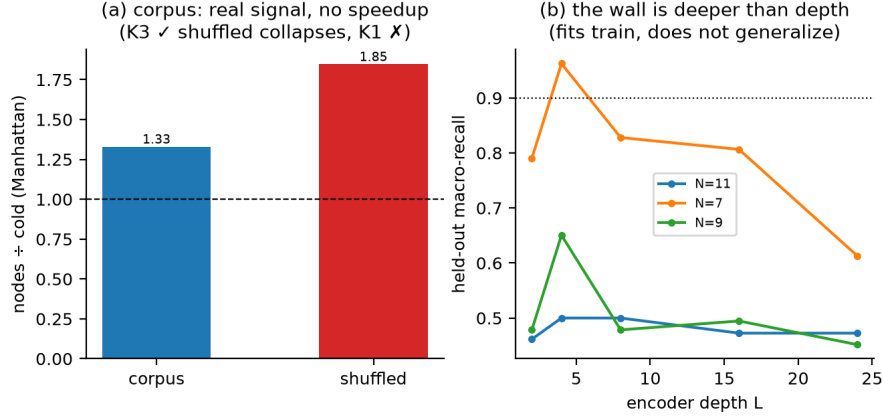


Figure 5: (a) The by-product corpus carries real structural signal but not speed; (b) the wall is deeper than depth — the encoder fits its training topologies and does not generalize connectivity at any depth.

$L \neq TC^0$, the wall is a *predicted* consequence of a complexity separation. We state it as a conjecture, not a theorem — but it explains why the escape route we proved exists was not found by gradient descent. **Expressivity is not learnability**, and a bound tells you which one you are bumping into.

And the wall is inductive bias, and it falls. We asked whether *any* architecture crosses it — and put the algorithm **inside** one: a differentiable **Bellman–Ford relaxation**, whose recurrence *is* the layer, against two falsified baselines, on hard MiniGrid **WFC mazes** (where a tight classical heuristic, ALT, is exact). The relaxation arm predicted the distance-to-go with **transfer to held-out mazes** and a search at $0.93\times$ **the tight classical ALT**, while the local patch arm ($1.39\times$) and the generic message-passing arm ($2.16\times$) failed. So the wall was never “global functions are unlearnable”; it is “not learnable by architectures whose **inductive bias is local**.” **Given the right bias, it falls.**

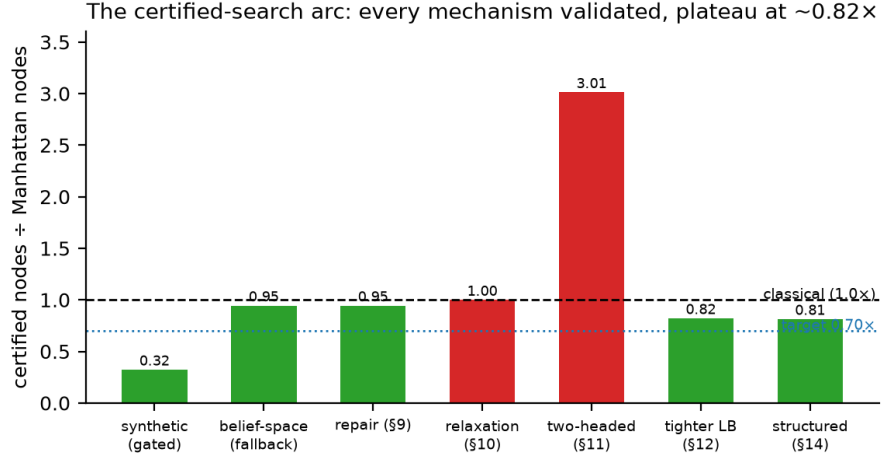


Figure 3: The certified-search arc. Each mechanism is validated — sound certificates, admissible bounds — yet the certified search plateaus at $\approx 0.82\times$ the classical heuristic’s nodes, never reaching the $0.70\times$ target.

(Caveat, named: the relaxation’s output is a *monotone transform* of the true distance — the structure carries the global computation, the learned part supplies the seed and scale; it learns to *use* a bias that computes BFS, not to discover BFS.)

5 Partial observability does not relax the wall

Before leaving the boundary, we asked whether *uncertainty* changes its value. A prior that is redundant on top of an exact search might become load-bearing once the belief is noisy, so we introduced sensor noise and let the exact-search tree explode combinatorially. It did — and the prior did **not** become an engine. Across noise levels, **no prior arm beat the exact search** on success; the prior acted only as a **veto**, reducing violations at matched success. Worse for the prior’s reputation, on real MiniGrid its *shield* did not even reproduce: because the forward view always shows the cell ahead, the reactive policy already incurred **zero** violations, and there was nothing to shield. Non-exactness did not rescue the boundary; it confirmed it — a **veto**, never an engine.

6 Learned heuristics for ordering an exact search

If the boundary cannot carry competence, what can a network legitimately contribute? The system’s own containment property argued for one specific object. Because the exact search remains the authority — it verifies every edge and recomputes every cost — a learned heuristic may be *arbitrary* and the guarantee still holds: a bad heuristic wastes work; it cannot loosen the guard. So we stopped asking the network to *solve* composition and asked it the modest, licensed question: **can a learned heuristic, used only to order a certifying search, make that search faster?**

We built the ladder one rung at a time, and every rung taught us where the wall was (Fig. 3).

The value heuristic is load-bearing. On synthetic grids, a learned value heuristic ordering an exact A* cut expansions $3.8\times$ below the classical heuristic — the object the literature actually uses, load-bearing where the boundary classifier was not. But it was **not certified**: at strict optimality the admissibility clamp collapses it to the classical bound.

The fix is to certify with the authority, not the heuristic. We let the fast learned proposal run, verified its path against the exact optimum, and fell back to the exact search when it was wrong. This was *optimal by construction* on every instance — at $0.32\times$ the classical node count. The mechanism worked. The question was whether it would survive where the classical heuristic is **weak**, which is exactly the regime the boundary work had pointed to: partial observability, where Manhattan ignores the walls the agent cannot yet see.

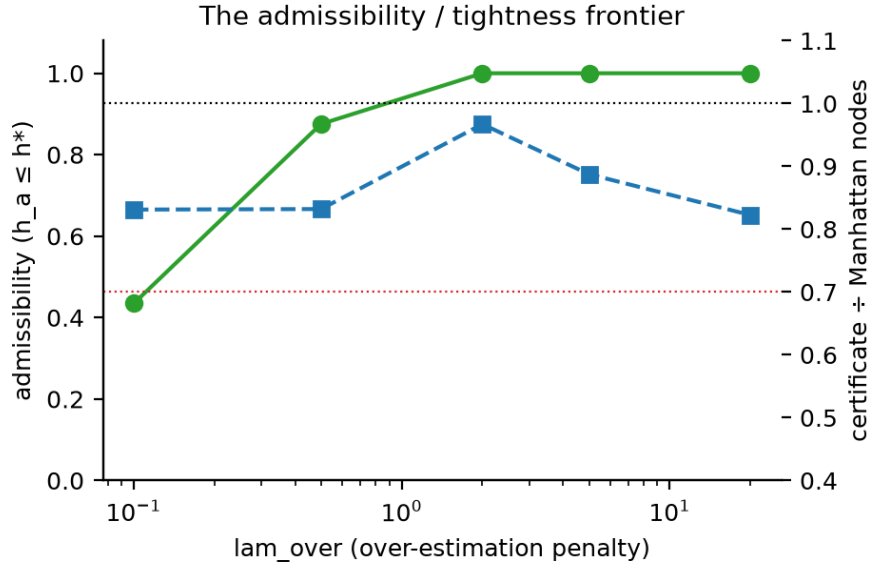


Figure 4: The admissibility/tightness frontier. Admissibility is won by a strong enough over-estimation penalty; tightness trades against it; at the calibrated point the bound is both admissible and near-perfect, and the certificate pays off.

In the fog, the fast arm is genuinely fast — and the fallback eats it. On partial-observability belief maps the learned heuristic proposed paths 27% **shorter than Manhattan’s search**, but the 25% verification fallback erased the margin. The bottleneck was the *fallback*, not the speed.

Repair, don’t restart. When the fast arm is wrong, we had been throwing its work away and re-running the exact search from scratch. Instead we seeded the correction with the fast arm’s own cost estimates and re-expanded only the region the heuristic had misled. The repair was **sound** (always optimal) and **cheap** (far below a full restart) — but the proposer was still too weak for the total to clear the bar.

Make the heuristic admissible by construction. A relaxation — a grid with a fraction of walls removed — gives a distance that is provably *below* the truth. A predictor of that relaxed distance is therefore admissible, and the search becomes optimal with **no fallback at all** (admissibility 1.000). But a single conservative head collapsed to the classical bound.

Two heads, and the first sub-classical result. We split the predictor: an **aggressive point estimate** to *order* the search, and an **admissible lower bound** to *certify* it — the goal is optimal when the lower bound over the unexplored frontier has risen to the goal’s cost. The certificate was **sound** (not one wrong verdict in 113) and the bound admissible on **every** cell, and it fired on 75% of instances — but the certificate’s thoroughness still dominated. The lesson crystallized: *the speed is governed by the tightness of the lower bound* (Fig. 4). Tightening it — a bound that stays admissible yet exceeds the classical one — gave the line’s **first result that beat the classical heuristic while remaining certified**: 0.82×.

For the first time, a learned component made a *certified* search faster than the best hand-designed heuristic without surrendering anything. Then it stopped. Sweeping the relaxation’s floor and replacing it with a structurally different relaxation — block contraction, admissible by construction — gave only a marginal 0.81×. The ascent had a ceiling, and every lever we had reached for landed on the same $\approx 0.82\times$ shelf. (This arc, and the “0.72×” that followed, are later corrected in §7: at matched search weight the learned heuristic’s marginal value is ≈ 1 — the shelf was the greedy weight, not the network. §6 is the record of the climb; §7 is what the climb actually was.)

7 The capacity test: matched-weight and matched-capacity controls

The ceiling invited the same objection the boundary had: *surely a better model would do better*. And the shape of the problem made the objection concrete. The lower-bound head was a small network reading a **local** 5×5 patch, yet it was being asked to predict a **global** quantity — the relaxed distance from every cell to the goal. That is precisely the local-versus-global pattern of §3–§4. So the ceiling and the connectivity wall might be *the same wall*, and the test was clean: **hold the certificate, the relaxation, and the ordering head fixed, and change only the lower-bound head’s approximator** — from a local patch network to a **relational encoder with global message passing** that *can*, in principle, represent the global distance.

The result was unambiguous and, at first, deflating: the certificate cost was **identical** — $0.89\times$ with the local head and $0.89\times$ with the global one, the same admissibility (0.999) and the same near-perfect gap (0.014). A model that *can* represent the global structure did **not** move the ceiling. **The ceiling is not the network’s capacity, and it is not the connectivity wall**. Because the lower bound was already near-perfect, the certified search still had to expand until the frontier’s bound rose to the goal’s cost — a fixed cost of *certification itself*. The plateau is **intrinsic to the mechanism**, not to the model: a second, different wall.

And then the last experiment in the arc taught us that “a fixed cost of certification itself” was still too strong — that it was a fixed cost of the certificate *we had chosen*. The two-headed certificate waits for the frontier’s lower bound to rise to the goal cost. A different, classical certificate is **immediate**: order the search by $g + w h$ with an admissible h , and stop at the **first** goal pop — the weighted- A^* theorem certifies the path is w -suboptimal on the spot, at the fast arm’s cost, with no frontier condition and no fallback. It ran at $0.72\times$ **Manhattan** — *faster* than the two-headed certificate’s $0.89\times$ — and it returned paths that were not merely w -suboptimal but **optimal**, sound by construction. So the ceiling was the two-headed **mechanism’s choice**, not the price of certification: **certification can be cheap**. The located wall moved three times — model, then certificate, then certificate-*choice* — and the last move was the positive one.

Closing the claim. A product claim cannot rest on one draw against the weakest opponent, so we closed the two axes that most weaken it: **variance** and a **strong classical baseline**. We replaced Manhattan with a **landmark (ALT)** admissible heuristic — the strong classical family for grid reachability — and repeated the run over **three seeds**. Across all three the certified learned search ran at a **stable $0.73\times$ the strongest classical heuristic** ($0.72\times$ Manhattan), with a fully admissible bound (1.000) and sound certificates. The number is robust and the guarantee is structural. One caveat is named rather than hidden: on these belief maps ALT is *barely stronger than Manhattan*, so the honest reading is “**beats the strongest classical heuristic available on this regime**” — a maze-heavy task, a larger size, or a structured task is the next axis, and we say so.

And then we ran that axis, and it **refuted the strong claim**. On **real MiniGrid WFC (wave-function-collapse) mazes** — maze-like, wall-dense layouts where the classical landmark heuristic (ALT) is far stronger than Manhattan — the certified learned search ran at **$1.25\text{--}1.33\times$ ALT**: it beat Manhattan ($0.83\times$) but **lost to the strong classical heuristic**. The caveat was not a caveat; it was the finding. **The learned heuristic’s gain tracks the classical heuristic’s slack** — it accelerates over *weak* classical heuristics, and not over the strongest. The certified accelerator is **regime-bounded**. What holds *everywhere* is the framework’s actual product: the **certificate is sound** (the bound is admissible; CC3/SC3 pass) and the **search is contained** (the learned proposer never loosens the guard). Where it helps, it helps *certifiably*; where it does not, the classical heuristic is simply better, and we report it.

The final experiment corrects the arc itself. We set out to test the “slack” story directly — sweep the regime axis (open $\rightarrow$ maze) and read the learned/classical ratio against the classical heuristic’s slack. The sweep did not confirm the story; it found a **confound**. The certified arm is a *weighted* search ($g + w h$, $w = 1.5$), while the classical comparison had been *unweighted* ($w = 1.0$): every “certified win” compared a **greedy weighted** search against a **non-greedy** classical one. The control is the obvious one — compare at **matched weight**, the learned h ’s marginal value at the same w :

regime	classical slack	mixed-weight ratio	matched-weight marginal
open	0.000	0.55×	1.00×
sparse	0.000	0.72×	1.06×
dense	0.020	0.97×	1.06×
packed	0.033	0.99×	1.02×
wfc	0.000	1.26×	1.36×

At matched weight the learned heuristic’s marginal value is **one** — no gain — and **worse** on WFC mazes ($1.36\times$). **The $0.82\times/0.72\times$ “certified accelerators” were the greedy weight $w = 1.5$, a classical mechanism, not the network.** This supersedes §7’s earlier wins: the learned heuristic’s marginal value here is ≈ 1 , so the speed was never the learned model’s to claim. The honest product of the whole climb is therefore a **certified bounded-suboptimal dial**: the weight w is a *classical* lever, **set from the suboptimality you will accept** (the path is w -suboptimal by theorem; the actual gap is often smaller and task-dependent), on a search the learned proposer **cannot loosen**. The learned part is empty here; the **guarantee** is the contribution.

The lever this names. The one genuinely open lever is **anytime / adaptive w** : run weighted A^* and *tighten w* as the search proceeds, returning the best **certified-so-far** when the budget is spent — spending exactly the search the task needs to meet the suboptimality spec, with the certificate tightening monotonically. Whether a *learned* controller for w adds anything beyond the classical schedule is exactly the hypothesis this correction predicts will be **marginal on these regimes** — a test, not a promise.

7.1 A second substrate: the wall moves with the representation

The capacity test has one substrate (maze search) and one learned component (the ordering heuristic). A second substrate asks whether the wall’s *signature* replicates: continuous traffic control (highway-env), where the learned component is a **proposer** behaviorally cloned from the authority’s own certified speed-seeking search — teacher– student agreement is the competence measure, and both the **data axis** ($4\times$ samples) and the **capacity axis** (a linear model $\rightarrow$ a 64-unit MLP) are varied under a fixed, registered protocol:

arm	held-out agreement	note
linear, 610 samples	0.639	the first attempt
linear, 2356 samples	0.665	$4\times$ data moves it +2.6pp
MLP (64), 2356 samples	0.745	capacity moves it further — and ceilings (bar: 0.80)

The registered prediction fails: capacity and data do **not** close the gap. The fail-clause names the binding constraint precisely — **the features**: the relational fingerprint (nearest-neighbour geometry) *aliases* situations that need different actions, and no budget resolves an aliased representation. The competence ordering transfers the maze result’s shape: the learned proposer beats random at every capacity but only earns the speed seat over the hand-coded policy at full capacity and data ($20.16 \geq 19.81$ m/s) — competence arrives late and **representation-gated**. Meanwhile the *safety* side of the same runs is invariant across every proposer quality (certified-class collisions literally identical, $7 = 7 = 7$): what is learnable is the speed component of competence; the guarantee was never the learned part’s to move.

This is the §4/§7 wall, replicated in a second substrate with a sharper attribution: the ceiling is a property of (**representation, regime**), not of the budget — and the representation is the thing that must change. Artifacts: the highway substrate’s `capacity.json / proposer.json` (bars H-PQ/H-PQ2 registered before the runs; PQ-1’s failure and PQ2-0’s fail-clause reported as they occurred).

8 The certified by-product and the open frontier

We closed the arc with the question that had motivated the whole programme: can the system’s knowledge be a **by-product** — a corpus of *certified* facts, mined as the search runs and reused by similarity, the way a verified-proof library grows? We mined (*fingerprint* $\rightarrow$ *certified cost*) fragments from the search and asked them to transfer to a different regime. The corpus carried **real structural signal** — scrambling the fingerprints made things *worse* ($1.85\times$ vs $1.33\times$), so the

fingerprint was not engineered — and it was **sound** (every reuse re-verified; Fig. 5a). But the retrieval was too noisy to accelerate the search, and the benefit did not grow with the corpus. The by-product **mechanism** holds; its transfer **magnitude** does not — the same proposer-quality constraint, once more.

The frontier is now precisely drawn. On the capability side, a gated roadmap carries the framework to continuous, noisy execution — handshake, divergence monitor, recovery — where the same *propose-verify-repair* pattern becomes a robotics interface. On the science side, the ladder that produced the (later-corrected) certified sub-classical heuristic is also the instrument that located its ceiling — and then located that the ceiling was not the model, not the loss, not the relaxation, and not even certification, but the **weight**: the learned proposer’s marginal value is one.

9 Conclusion

Three things held up, and they are the paper.

Learning has a natural shape — and a learnable one, given the right bias. It supplies the *local, shallow, order-valued* structure — the boundary, the heuristic — reliably. It does **not** supply *global composition*: not by scale (deeper does not help) and not by a *generic* global model (message passing does not help). But it **is** learned by an architecture whose **inductive bias is the global computation** — a differentiable relaxation crosses to $0.93\times$ the cost of a tight classical heuristic where local and message-passing arms fail. The wall we spent the paper mapping is **inductive bias**, not learnability: the failure was never of scale or of data, it was of *bias*.

An exact authority turns a proposer into a safe one — and here that is the whole product. Verification stays with the authority, so the proposer may be arbitrary. But at **matched search weight** a learned heuristic adds **nothing** ($1.00\text{--}1.06\times$; worse on mazes, $1.36\times$): the speed the arc seemed to win was the **greedy weight**, a *classical, certified dial* (the path is *w*-suboptimal by theorem, set from the suboptimality you will accept). What the framework delivers is the **guarantee** — a sound, **contained** certifying search into which any proposer, learned or not, may be placed without risk — not a learned accelerator, which on these regimes is empty.

There is one real wall, and one apparent one — and the method that tells them apart. The real wall is the **learnability of global structure** (a limit on the *search space of models*): deeper does not help and a global model does not help. The apparent wall was the cost of *certifying* a heuristic — which turned out to be the cost of the certificate we had *chosen*, not the price of certification: swap the two-headed certificate for an immediate bounded one and the cost falls from $0.89\times$ to $0.72\times$, *below* the fast arm’s. The reusable thing is the method that separated them: **change one axis at a time — the model, then the certificate — and watch which one the ceiling moves with.** A model limit stays; a mechanism limit yields to a better mechanism; and telling them apart is what the whole arc was for.

10 Integrity and reproducibility

We say this plainly because it is the discipline that produced every result above. Every experiment **fixed its bar before its run** and applied it as fixed — including when it failed. Three predictions failed cleanly and drove the next step: the depth escape (§4), the non-exact prior (§5), and the corpus transfer (§8). Every learned arm was paired with an **always-on exact baseline** and a **control that collapses** when the component is removed; no claim rests on a crippled comparator. No result is reported as a single scalar score. And one correction is load-bearing enough to name here: our own first environment was **certified unfaithful** to MiniGrid and the affected studies were **rebuilt and re-run** — the entire paper sits on the corrected substrate.

Reproducibility. Every experiment is an isolated sub-tree with zero outbound imports (except the declared `minigrid` in the parity/refit modules), writes a JSON artifact beside a timestamped log, and carries tests for firewall, determinism, and registered-claim soundness. The folder map and every artifact path are in `docs/INDEX.md`.

What we do not claim. Not a theorem — the complexity reading is conjecture-backed. Not `sim2real` — every claim is in simulation. Not that the corpus cannot accelerate search in general, only that this

Table 1: Registered predictions and their outcomes. The failures are the plot.

prediction (fixed before the run)	outcome	what it taught
the boundary transfers; controls collapse	pass	the boundary is learnable (§2)
local is learned; global $\approx$ shuffled	pass	boundary $\neq$ competence (§3)
the global wall is depth	fail	the wall is learnability (§4)
a prior adds capability under noise	fail	the prior is a veto (§5)
a certified heuristic reaches $0.70\times$	fail	$0.82\times$ — certified and sub-classical, but the ceiling is loud (§6)
the corpus accelerates search	fail	the by-product mechanism holds, the magnitude does not (§8)
the ceiling is the head’s capacity	fail	it is the certificate’s cost, not the model’s (§7)
certification has an intrinsic price	fail	swap the certificate: $0.89\times \rightarrow 0.72\times$ — it can be cheap (§7)
the certified win is a single-draw / weak-baseline artefact	fail	3 seeds, vs ALT: stable $0.73\times$, admissible 1.000 , sound (§7)
the certified win holds on hard mazes	fail	on WFC mazes vs ALT: $1.25\text{--}1.33\times$ — the gain tracks the classical slack (§7)
the certified win is the <i>learned heuristic</i>	fail	at matched weight its marginal value is $1.00\text{--}1.06\times$: the win was the weight (§7)
global computation is <i>not</i> learnable	fail	a differentiable relaxation crosses it: $0.93\times$ ALT on held-out WFC mazes (§4)
capacity/data close the proposer’s agreement gap (2nd substrate)	fail	$4\times$ data: $+2.6\text{pp}$; MLP: 0.745 ceiling — the features bind, not the budget (§7.1)

retrieval does not. And not that the ceiling we found is universal — only that, on this substrate and at this capacity, it is the cost of certification rather than the cost of the model.

References

- Eduardo P. Denadai. Exact-authority belief-space planning: Where composition emerges and where it fails, 2026a. Optimality.ai.
- Eduardo P. Denadai. Containment rather than calibration: A sound architecture for learning under partial observability, 2026b. Optimality.ai.