
STOCHASTICITY IS NOT THE HARD PART: REDUCTION AND COMPLEXITY IN INSTRUCTIONAL SEQUENCING OVER PREREQUISITE DAGs

A PREPRINT

Zonglin Han¹ Yichen Chen^{1†} Jiawen Jiang^{2†} Tongan Shi^{3†} Kristian A. Stevens¹

¹Department of Computer Science, University of California, Davis

²International Digital Economy College, Minjiang University

³School of Computer Science and Artificial Intelligence, Liaoning Normal University

[†]Equal contribution (alphabetical order)

ABSTRACT

When a student must learn concepts connected by prerequisite dependencies, when does the order of instruction matter, and what does it cost to find the best one? We study instructional sequencing as a stochastic shortest-path problem in which attempting a concept succeeds with a state-dependent probability and failure leaves the learner state unchanged. We first prove that this stochasticity can be eliminated *exactly*: the problem collapses to a deterministic shortest-path problem on the lattice of prerequisite order ideals, preserving optimal values and actions. The collapse removes stochastic complexity but not combinatorial complexity: optimal sequencing remains NP-hard—via reduction from feedback arc set in tournaments—even with no prerequisite edges, unit costs, uniform binary nonnegative transfer, and success probabilities at least $1/2$. Hardness is not uniform: when realizable transfer preferences remain jointly acyclic with the prerequisites, any topological order of the residual joint graph is optimal, and fixed prerequisite width yields polynomial-time exact dynamic programming. A computable diagnostic, $m\Delta$, bounds the value of sequencing before optimization. On 70,893 interactions from an introductory CS course, the diagnostic certifies a doubly easy regime—little value to optimize and little space to search—while constructed transfer instances realize the challenging regime, where myopic sequencing suffers large regret yet exact A^* with a consistent heuristic expands only linearly many states on that family.

1 Introduction

When a student must learn a set of concepts connected by prerequisite dependencies, what is the computational complexity of finding the learning order that minimizes expected instructional cost? Platforms such as Khan Academy serve more than 180 million learners over prerequisite-structured curricula [1]. Individualized instruction can substantially improve learning outcomes [2–4], while prerequisite structure can be recovered from curricula and interaction data [5, 6]. Yet the computational question remains unresolved: *when does the order of instruction actually matter, and what does it cost to find the best one?*

Existing work largely asks how to generate a learning sequence or policy, using bandits [7], reinforcement learning and partially observable models [8, 9], prerequisite-aware MDPs [10, 11], or, more recently, large language models. Four questions logically precede policy design: whether stochastic learning dynamics genuinely require stochastic planning; when learner-state dependence creates value in sequencing; which transfer structures induce combinatorial conflict; and what governs the difficulty of exact search. We answer all four. The stochastic dynamics can be eliminated *exactly*; the resulting ordering problem remains NP-hard even under severe restrictions; it becomes polynomial when realizable transfer preferences remain jointly acyclic with the prerequisites; and a simple a priori diagnostic bounds the value of sequencing—correctly predicting that the course instances we evaluate lie in a doubly easy regime.

We model a curriculum as a prerequisite learning graph $G = (V, E)$, a DAG with a learner state given by a downward-closed set of mastered concepts. At state s , the feasible actions are the unmastered concepts whose prerequisites are already contained in s , a graph-theoretic reading of the zone of proximal development [12]. Attempting concept v incurs cost T_v , succeeds with learner-oracle probability $p(v, s)$, and otherwise leaves the decision state unchanged. Reaching a prerequisite-closed target set at minimum expected total cost defines an instructional sequencing problem over the prerequisite graph, a stochastic shortest-path problem [13]. Because failures preserve the decision state and the success probabilities are time-homogeneous, however, the stochastic self-loops can be eliminated *exactly*: each action becomes a deterministic transition of cost $T_v/p(v, s)$ on the lattice of prerequisite order ideals, with optimal values and minimizing actions preserved (a complementary deterministic-valuation theory on the same lattice appears in concurrent work [14]). This collapse removes stochastic complexity but not combinatorial complexity. Optimal sequencing remains NP-hard even with no prerequisite edges, unit attempt costs, uniform binary nonnegative transfer, monotone success probabilities, and success probabilities at least $1/2$.

The hardness is not uniform. For an explicit pairwise-transfer oracle, if the realizable soft ordering preferences and hard prerequisites form an acyclic residual joint graph, then any topological ordering of that graph is globally optimal and can be found in polynomial time; for uniform affine transfer, optimal sequencing is further fixed-parameter tractable in the distance to acyclicity. More generally, if the prerequisite poset has width w , exact dynamic programming is polynomial for fixed w . Between these structural regimes, a computable quantity $m\Delta$ upper-bounds how much any valid sequence can improve over another, prior to sequence optimization. On 70,893 interactions from an introductory computer science course (ECS32A), this diagnostic certifies a doubly easy regime: the learned evaluator leaves less than 0.2% model-based improvement available to sequencing, while the target prerequisite closures have width at most four and at most 47 reachable ideals. A larger public curriculum, Junyi Academy, exhibits the opposite topology: sink-target closures reach width 18, with 68 closures exceeding 10^7 reachable ideals, so structurally challenging topologies do occur in real curricula. A first-attempt diagnostic proxy further suggests substantially greater state dependence, though the proxy is exploratory rather than a certified bound. Constructed transfer instances realize the challenging regime in full: myopic sequencing incurs large regret exactly where the order-ideal lattice grows exponentially, while exact A* with a consistent heuristic expands only linearly many states on the constructed family.

To our knowledge, this is the first computational characterization of failure-invariant instructional SSPs on prerequisite ideal lattices. We establish an exact stochastic-to-deterministic reduction; NP-hardness alongside residual-acyclic, width-XP, and affine-FPT regimes; the $m\Delta$ opportunity bound; and exact A* with a consistent heuristic. Experiments contrast doubly easy ECS32A instances with high-regret constructed instances (Sections 3–6).

2 Related Work

2.1 Adaptive Learning and Instructional Sequencing

Our work characterizes the instructional sequencing problem that adaptive-learning systems attempt to solve. Instructional sequencing has been addressed with expert rules [3, 15], bandits [7], reinforcement learning [9], POMDP planning [8], and MDP-based learning-path recommendation over prerequisite structures [10, 11, 16]. A parallel literature sequences training tasks for artificial agents [17, 18]. These methods primarily address how to construct sequences under rich practical constraints, and prior complexity analyses in this literature typically arise from content-selection, resource, or multi-objective formulations [19]. Our analysis instead isolates a minimal planning core—state-conditioned mastery dynamics on a prerequisite lattice—and derives hardness, tractable subclasses, and instance diagnostics for that core, separating statistical and model-learning challenges from the intrinsic optimization structure.

2.2 Knowledge Tracing

Knowledge tracing estimates the learner state our planner conditions on; we contribute no new KT model and remain oracle-agnostic. The lineage runs from Bayesian Knowledge Tracing [20] through recurrent [21] and attention-based architectures [22]. In our framework a student model enters only through the induced interface $p(v, s)$: the success probability of a feasible concept given a hypothesized mastery state. Models whose native conditioning is an interaction history require an explicit projection onto mastery states, and this projection is part of the modeling choice rather than a free equivalence. For sequencing, what matters is the state variation an oracle induces in collapsed costs—which governs the available planning opportunity—together with its fidelity to the true learner, which governs whether optimizing against it helps.

2.3 Structure and Search on Ideal Lattices

The order-ideal state space itself has a long history in mathematical psychology: knowledge space theory models a learner as the subset of items mastered, and prerequisite partial orders induce the quasi-ordinal knowledge spaces represented here as order-ideal lattices [23, 24], with efficient algorithms for assessment and learning-sequence enumeration [25]. Our contribution is not the state space but the planning problem on it: expected-cost-optimal sequencing under a state-conditioned stochastic oracle, its exact deterministic collapse, and its complexity.

Our stochastic model is a stochastic shortest-path problem [13], and our solver is A* with a consistent heuristic [26]. LAO* extends heuristic search to stochastic shortest-path problems whose solution structures contain loops [27, 28]; our collapse result shows that such machinery is unnecessary in this subclass, making graph-search A* the natural exact solver, and we retain a LAO*-derived implementation as a solver baseline. Ye et al. [29] apply SSP to spaced-repetition scheduling of individual items; their formulation optimizes review timing for one memory trace, whereas ours orders many concepts under prerequisite and transfer coupling. Our hardness proof reduces from feedback arc set in tournaments, whose NP-hardness was established by Alon [30] and Charbit et al. [31]; our width-based XP bound rests on Dilworth’s decomposition theorem [32]; and our fixed-parameter result for uniform affine transfer follows from the FPT classification of MinCSP($<, \leq$), equivalently Directed Subset Feedback Arc Set [33].

Closest to our setting, Pasechnyuk-Vilensky [14] independently studies order sensitivity for deterministic edge-additive interventions on the ideal lattice of a prerequisite poset, characterizing path independence through vanishing diamond curvature and bounding path-value differences by the number of local diamond swaps; the planning consequence is dynamic programming on a depth-truncated Hasse diagram, validated on general sequential-intervention experiments without educational learner-model evaluation. Our setting is complementary: we begin with a stochastic instructional shortest-path problem in which failed learning attempts leave the learner state unchanged, prove an exact reduction to deterministic edge costs $T_v/p(v, s)$, establish worst-case NP-hardness and structural tractable regimes, and develop exact A* search with a consistent heuristic on the implicit lattice. The two order-sensitivity analyses quantify different objects: diamond curvature is a local, second-order mixed difference, whereas our Δ_v measures the global feasible-state range of a concept’s collapsed cost. Local curvature is controlled by global variation but not conversely, and neither resulting global bound uniformly dominates the other; we formalize this relationship in Section 4.

3 Model and Exact Reduction

We formalize the instructional sequencing task described above as the *prerequisite sequencing problem*, defined by the following components. The section introduces the primitive instructional SSP, establishes its exact deterministic collapse, records the resulting order-ideal structure, and delineates the boundary of the reduction.

Definition 1 (Prerequisite structure and mastery states). *Let $G = (V, E_G)$ be a finite prerequisite DAG given by its transitive reduction. For $u, v \in V$, write $u \prec_G v$ when G contains a directed path from u to v , and let $\text{anc}_G(v) = \{u \in V : u \prec_G v\}$ denote the ancestor (full prerequisite) set of v . A mastery state is an order ideal of this partial order:*

$$\mathcal{S} = \{s \subseteq V : v \in s \implies \text{anc}_G(v) \subseteq s\}.$$

The initial state $s_0 \in \mathcal{S}$ is supplied by an upstream assessment procedure. Instruction is confined to a fixed instructional domain: a prerequisite-closed set Q with $s_0 \subseteq Q \subseteq V$ —in our experiments, the prerequisite closure of a designated target set—whose concepts are the only admissible actions, and the goal is $s^ = Q$. Because transfer can originate outside a target’s closure, this exclusion is a modeling assumption rather than a lossless normalization: transfer-inducing detours outside Q are not represented. We relabel Q as V hereafter, so $s^* = V$ and the complete-mastery case is $Q = V$.*

Throughout, s is the complete decision-relevant state: every quantity that affects transition probabilities or costs is a function of s alone. Mastery is monotone—once a concept crosses the mastery threshold, it remains mastered—matching the absorbing learned-state semantics used in classical knowledge tracing models [20]. This assumption will imply that every successful transition moves upward in the order-ideal lattice.

At a nonterminal state s , the feasible actions are the unmastered concepts on the prerequisite frontier,

$$\mathcal{A}(s) = \{v \in V \setminus s : \text{anc}_G(v) \subseteq s\}.$$

This frontier gives a graph-theoretic interpretation of the zone of proximal development [12]: instruction is restricted to concepts for which the modeled prerequisites have already been mastered. The frontier is part of the hard-prerequisite semantics of the model rather than an algorithmic pruning rule; soft cross-concept effects among feasible actions are represented through $p(v, s)$.

Definition 2 (Primitive instructional SSP). *The primitive instructional SSP is*

$$\mathcal{M} = (\mathcal{S}, s_0, s^*, \mathcal{A}, P, T).$$

For every $s \neq s^*$ and $v \in \mathcal{A}(s)$, one instructional attempt incurs a finite cost $T_v > 0$ and succeeds with probability $p(v, s) \in (0, 1]$. The transition kernel is

$$P(s' \mid s, v) = \begin{cases} p(v, s), & s' = s \cup \{v\}, \\ 1 - p(v, s), & s' = s, \\ 0, & \text{otherwise.} \end{cases}$$

The success probabilities and attempt costs are time-homogeneous. A failed attempt means that the learner has not crossed the mastery threshold, so the complete decision state remains s ; a successful attempt adds exactly v . The goal s^* is absorbing with zero cost. Instruction is serial, so each attempt targets one concept.

For a policy π , let $\tau = \inf\{t : s_t = s^*\}$. The objective is the risk-neutral additive expected total cost

$$J^\pi(s) = \mathbb{E}_s^\pi \left[\sum_{t=0}^{\tau-1} T_{v_t} \right], \quad J^*(s) = \inf_{\pi} J^\pi(s).$$

Because V is finite and all feasible success probabilities are positive, $p_{\min} = \min_{s, v \in \mathcal{A}(s)} p(v, s) > 0$. From any state s , every policy needs at most $|V \setminus s|$ successes, each requiring at most $1/p_{\min}$ attempts in expectation, so $\mathbb{E}_s^\pi[\tau] \leq |V \setminus s|/p_{\min}$ and the expected cost is at most $T_{\max}|V \setminus s|/p_{\min}$, where $T_{\max} = \max_v T_v$; every policy is proper.

Theorem 1 (Exact stochastic collapse). *Consider the finite primitive instructional SSP above, in which each feasible action has a positive time-homogeneous success probability and finite time-homogeneous attempt cost, failure leaves the complete Markov state unchanged, success has the unique progress successor $s \cup \{v\}$, and performance is measured by risk-neutral additive expected total cost. Define a deterministic graph on the same state set $\mathcal{S}$, with one edge and cost*

$$s \longrightarrow s \cup \{v\}, \quad c(s, v) = \frac{T_v}{p(v, s)}, \quad v \in \mathcal{A}(s).$$

Then the primitive SSP and this deterministic shortest-path problem have identical optimal values and identical sets of Bellman-minimizing actions at every state. Every optimal deterministic stationary collapsed policy lifts to an optimal deterministic stationary primitive policy by retrying its selected action after failure until success, and every optimal deterministic stationary primitive policy induces an optimal collapsed policy through the same state-action map.

Proof sketch. For a nonterminal state s , the primitive Bellman equation is

$$J^*(s) = \min_{v \in \mathcal{A}(s)} [T_v + p(v, s)J^*(s \cup \{v\}) + (1 - p(v, s))J^*(s)].$$

For each feasible action define

$$\tilde{Q}_v(s) = \frac{T_v}{p(v, s)} + J^*(s \cup \{v\}).$$

If $Q_v^{\text{prim}}(s)$ denotes the corresponding primitive Bellman expression, then

$$Q_v^{\text{prim}}(s) - J^*(s) = p(v, s)(\tilde{Q}_v(s) - J^*(s)).$$

Since $p(v, s) > 0$, the two differences have the same sign and vanish simultaneously. Hence the primitive and collapsed models have the same minimizing action set, although suboptimal actions need not have the same ranking. It follows that

$$J^*(s) = \min_{v \in \mathcal{A}(s)} \left[\frac{T_v}{p(v, s)} + J^*(s \cup \{v\}) \right].$$

Under a stationary deterministic choice of v , attempts until success are geometric with mean $1/p(v, s)$, so retrying after each failure realizes exactly the collapsed edge cost and successor, establishing policy lift and recovery. The complete correspondence proof is given in the supplementary material.

Corollary 1 (Acyclic order-ideal lattice). *Under monotone mastery, the deterministic graph of Theorem 1 has vertex set $\mathcal{S}$ (the order ideals of G), directed edges $s \rightarrow s \cup \{v\}$ for each $v \in \mathcal{A}(s)$, and edge weight $T_v/p(v, s)$. It is acyclic and layered by mastery cardinality.*

Every collapsed edge satisfies $|s \cup \{v\}| = |s| + 1$, so no directed cycle is possible. We refer to this deterministic shortest-path problem as the *collapsed Ariadne sequencing problem* and to its state graph as the *collapsed lattice*.

Remark. The self-loop elimination underlying Theorem 1 is standard in time-homogeneous SSPs; the contribution is its rigorous instantiation on the ideal lattice (properness, policy correspondence) and the consequence that LAO*-style loop handling is unnecessary. The reduction generally fails when failures update beliefs, costs depend on attempt history, mastery can be forgotten, actions have multiple outcomes, or the objective is risk-sensitive. It shows the problem is deterministic; it does not show it is easy.

4 Computational Complexity of the Collapsed Problem

We now map the computational complexity of the collapsed Ariadne sequencing problem. The landscape ranges from order-insensitivity through polynomial and parameterized tractability to NP-hardness, with an instance-level diagnostic bounding the value of sequencing before sequence optimization.

Fix an initial mastery state s_0 , let $R = V \setminus s_0$, and write $m = |R|$. Let $\Sigma_G(s_0)$ denote the prerequisite-valid complete orderings of R . For $\sigma = (v_1, \dots, v_m)$, define $s_i = s_0 \cup \{v_1, \dots, v_i\}$ and

$$C(\sigma) = \sum_{i=1}^m c(v_i, s_{i-1}), \quad c(v, s) = \frac{T_v}{p(v, s)}.$$

Paths from s_0 to s^* in the collapsed lattice correspond bijectively to orderings in $\Sigma_G(s_0)$, so

$$C^* = \min_{\sigma \in \Sigma_G(s_0)} C(\sigma).$$

When can sequencing matter? For each $v \in R$, let $\mathcal{F}_v = \{s \in \mathcal{S} : s_0 \subseteq s, v \in \mathcal{A}(s)\}$ be its reachable feasible states, and define

$$\Delta_v = \max_{s, s' \in \mathcal{F}_v} |c(v, s) - c(v, s')|, \quad \Delta = \max_{v \in R} \Delta_v.$$

When every $\Delta_v = 0$, each concept has the same collapsed cost whenever it is legally attempted; all valid complete orderings therefore have equal total cost. The following result quantifies the general case.

Proposition 1 (Sequencing-value bound). *For any $\sigma, \sigma' \in \Sigma_G(s_0)$,*

$$|C(\sigma) - C(\sigma')| \leq \sum_{v \in R} \Delta_v \leq m\Delta.$$

Consequently, $C(\sigma) - C^ \leq m\Delta$ for every valid complete ordering σ .*

Proof sketch. Both orderings master each $v \in R$ exactly once, possibly at different feasible states. The two contributions of v differ by at most Δ_v ; summing concept-wise proves the claim. Full proofs appear in the supplementary material. Under monotonicity, the same $2|R|$ oracle calls also yield a multiplicative bound: $C(\sigma)/C^* \leq \max_v p_v^{\max}/p_v^{\min}$, so large relative regret requires large oracle dynamic range.

Concurrent work bounds order insensitivity via local diamond curvature ε and swap distance [14]; $\varepsilon \leq 2\Delta$ but neither global bound dominates the other, since vanishing curvature can coexist with large Δ_v . Under a monotone oracle, Δ_v is exact in two oracle calls— $p_{\min}$ at $s_0 \cup \text{anc}_G(v)$, $p_{\max}$ at $V \setminus (\{v\} \cup \text{desc}_G(v))$ —giving certified $\sum_v \Delta_v$ in $O(|R|)$ calls with no lattice enumeration (supplementary material). For non-monotone oracles the diagnostic may require enumeration or sampling; only certified values yield the formal guarantee.

A polynomial transfer subclass. To isolate the role of transfer structure, consider an explicit nonnegative pairwise-transfer oracle

$$p(v, s) = \phi_v \left(b_v^0 + \sum_{u \in s \cap R} w_{uv} \right), \quad w_{uv} \geq 0,$$

where b_v^0 is transfer already supplied by s_0 and each $\phi_v : \mathbb{R}_{\geq 0} \rightarrow (0, 1]$ is nondecreasing. Let $E_H^R = \{(u, v) \in R^2 : w_{uv} > 0\}$. A soft edge $u \rightarrow v$ is hard-contradictory when $v \prec_G u$, because every prerequisite-valid sequence must place v before u . Remove these impossible preferences and define

$$E_H^+ = E_H^R \setminus \{(u, v) : v \prec_G u\}, \quad D_R^+ = (R, E_G[R] \cup E_H^+).$$

For each v , its concept-wise attainable transfer and cost bounds are

$$\bar{A}_v = b_v^0 + \sum_{u: (u, v) \in E_H^+} w_{uv}, \quad \underline{c}_v = \frac{T_v}{\phi_v(\bar{A}_v)}.$$

Theorem 2 (Residual-acyclic optimality). *If D_R^+ is acyclic, every topological ordering of D_R^+ is globally optimal, with*

$$C^* = \sum_{v \in R} c_v.$$

Proof sketch. Any topological ordering of D_R^+ respects the hard prerequisites and places every noncontradictory transfer source u before its recipient v . Each concept therefore receives $\bar{A}_v$ and attains its individual lower bound c_v ; their sum is a lower bound for every valid sequence and is attained simultaneously. The full proof and the transfer-edge classification appear in the supplementary material.

With prerequisite reachability available, constructing and checking D_R^+ is linear in $|R| + |E_G[R]| + |E_H^R|$.

Worst-case hardness. Residual acyclicity is a sufficient certificate, not a universal property. To expose the remaining combinatorial difficulty, consider *Restricted Ariadne Sequencing*: the input contains n concepts, rational attempt costs, a prerequisite DAG, a binary matrix $B = (b_{uv})$, and a rational threshold K ; success probabilities are fixed by

$$p(v, s) = \left(2 - \frac{1}{n} \sum_{u \in s} b_{uv} \right)^{-1},$$

and the question is whether a prerequisite-valid complete ordering σ satisfies $C(\sigma) \leq K$.

Theorem 3 (Hardness of optimal sequencing). *Restricted Ariadne Sequencing is NP-complete even when the prerequisite graph is empty, all attempt costs are one, transfer is nonnegative with uniform binary magnitude, success probabilities are monotone in s , and every success probability is at least $1/2$. Consequently, its optimization version is NP-hard under the same restrictions.*

Proof sketch. Reduce from Minimum Feedback Arc Set in Tournaments [30, 31]. Given a tournament, set $b_{uv} = 1$ exactly for its arcs and remove all prerequisite edges. Then

$$c(v, s) = 2 - \frac{1}{n} \sum_{u \in s} b_{uv}, \quad C(\sigma) = 2n - \frac{\text{fwd}(\sigma)}{n}.$$

Minimizing instructional cost is therefore equivalent to maximizing forward tournament arcs, or minimizing backward arcs. A complete ordering is a polynomial certificate; details appear in the supplementary material. The restrictions show that hardness comes from ordering conflicts induced by state-conditioned transfer, rather than from prerequisite legality, heterogeneous costs, extreme probabilities, negative effects, or stochastic self-loops.

Width-parameterized dynamic programming. Worst-case hardness does not preclude exact optimization on narrow prerequisite orders.

Proposition 2 (Width-based XP algorithm). *Let the prerequisite poset on R have width w . Then*

$$|\mathcal{S}_{\text{reach}}| \leq \left(\frac{m}{w} + 1 \right)^w,$$

and exact dynamic programming runs in $O(m(m/w + 1)^w)$.

Proof sketch. By Dilworth’s theorem, the poset partitions into w chains of lengths $n_1, \dots, n_w$ with $\sum_i n_i = m$ [32]. Every order ideal is uniquely encoded by one prefix length per chain, so its count is at most $\prod_i (n_i + 1)$, which is bounded by $(m/w + 1)^w$ via AM–GM. A reverse-cardinality Bellman pass scans at most m actions per state. This is XP, not FPT, in w , because the exponent depends on the parameter. In our data, ECS32A closures have $w \leq 4$ and at most 47 reachable ideals, whereas Junyi sink closures reach width 18 and 68 exceed our 10^7 -ideal enumeration guard, motivating implicit heuristic search. Full derivations appear in the supplementary material.

Distance to residual acyclicity. Residual acyclicity is the zero-distance endpoint of a broader parameterized regime. For a valid ordering σ , let $B_H^R(\sigma)$ be the residual soft edges directed backward by σ , and define

$$k = \min_{\sigma \in \Sigma_G(s_0)} |B_H^R(\sigma) \cap E_H^+|.$$

Equivalently, k is the minimum number of residual soft edges whose deletion makes D_R^+ acyclic; hard prerequisite edges cannot be deleted.

Theorem 4 (Affine fixed-parameter tractability). *Suppose the collapsed costs have the uniform affine form*

$$c(v, s) = a_v - \lambda |\{u \in s \cap R : (u, v) \in E_H^+\}|, \quad \lambda > 0,$$

with all feasible costs positive. Exact optimal sequencing is fixed-parameter tractable in k : an optimal sequence can be computed in $f(k) \cdot (|R| + |E_G| + |E_H|)^{O(1)}$ time, where the polynomial is in the encoding of the residual graph, not the lattice.

Proof sketch. Each satisfied residual soft edge reduces total cost by exactly λ , so

$$C(\sigma) = C_0 + \lambda |B_H^R(\sigma) \cap E_H^+|.$$

Cost minimization is therefore exactly minimum residual frustration. This is Directed Subset Feedback Arc Set, equivalently MinCSP($<, \leq$), with soft precedence constraints breakable and hard prerequisites fixed; known algorithms are FPT in the violated-constraint budget [33]. The supplementary material gives the reduction and sequence recovery. The affine assumption is essential to our reduction: for general nonlinear monotone oracles, fixed-parameter tractability in this distance remains open.

5 Exact Search on the Implicit Lattice

The landscape tells us where exact planning is feasible; we now ask how to perform it when the lattice is too large to enumerate.

A consistent sum heuristic. For each concept $v \in R$, let $\bar{p}_v$ be any certified upper bound satisfying

$$\bar{p}_v \geq \max_{s \in \mathcal{F}_v} p(v, s).$$

The tightest separable choice is $p_v^* = \max_{s \in \mathcal{F}_v} p(v, s)$; under a monotone oracle this is analytic: $p_v^* = p(v, V \setminus (\{v\} \cup \text{desc}_G(v)))$ (Section 4). When no oracle structure is available, the trivial bound $\bar{p}_v = 1$ remains valid. Define

$$h_{\text{sum}}(s) = \sum_{v \in V \setminus s} \frac{T_v}{\bar{p}_v}.$$

Proposition 3 (Consistency of the sum heuristic). *The heuristic h_{sum} is consistent on the collapsed Ariadne lattice and satisfies $h_{\text{sum}}(s^*) = 0$. It is therefore admissible.*

Proof sketch. For every collapsed edge $s \rightarrow s \cup \{v\}$,

$$h_{\text{sum}}(s) - h_{\text{sum}}(s \cup \{v\}) = \frac{T_v}{\bar{p}_v} \leq \frac{T_v}{p(v, s)} = c(v, s).$$

This is exactly the consistency inequality $h_{\text{sum}}(s) \leq c(v, s) + h_{\text{sum}}(s \cup \{v\})$; goal-zero consistency implies admissibility. A direct concept-wise admissibility proof is given in the supplementary material. Consistency also ensures that graph-search A* never reopens a closed state: when a state is removed from the priority queue, its g -value is final.

A* on the implicit lattice. The order-ideal graph is represented implicitly. Starting from s_0 , expanding a state s enumerates only its frontier actions $\mathcal{A}(s)$ and generates successors $s \cup \{v\}$ on demand. Standard graph-search A* orders open states by

$$f(s) = g(s) + h_{\text{sum}}(s)$$

and terminates when the goal is first removed from the queue. Because the heuristic is consistent and all collapsed costs are positive, each state is expanded at most once and the returned path is optimal. In the worst case, A* may still expand every reachable order ideal, reducing to exhaustive dynamic programming with priority-queue overhead; its advantage is that informative lower bounds can avoid generating large portions of a wide lattice.

Solver comparison. We compare full reverse-topological dynamic programming, Dijkstra’s algorithm ($h = 0$), A* with h_{sum} , and a LAO*-derived solution-graph baseline. All four return identical optimal values to within 10^{-12} . A* provides the strongest expansion and wall-clock performance on the wide instances, whereas the additional revision machinery of the LAO*-derived solver offers no benefit after collapse—consistent with Theorem 1, which leaves no stochastic loops for that machinery to exploit. Detailed counts appear in Section 6 and the supplementary material.

The structural analysis motivates two empirical regimes; we now test where real and constructed instances land.

Table 1: Sequence quality under the frozen public evaluator (repetitions averaged within each target; ten targets weighted equally). Ariadne Exact shares the evaluator’s planning oracle by design, so its zero regret is a protocol identity rather than an empirical claim about real-student optimality. Mean costs for all conditions appear in the supplementary material.

Condition	Normalized regret
Ariadne Exact	0
Ariadne Greedy	5.934×10^{-7}
Frequency Greedy	1.047×10^{-5}
DKT-set Greedy	4.935×10^{-5}
DKT-set Exact	1.828×10^{-4}
BKT-set Greedy	1.995×10^{-4}
BKT-set Exact	2.481×10^{-4}
Random Frontier	2.633×10^{-4}
Frequency Exact	2.658×10^{-4}
Linear Syllabus	3.385×10^{-4}

6 Experiments

Setup. We evaluate Ariadne on 70,893 mapped interactions from 294 students in an introductory computer science course (ECS32A), organized by a 61-concept prerequisite DAG. The evaluation uses ten prespecified target closures, an empty initial mastery state, and a uniform nominal attempt cost $T_v = 60$, so expected cost is proportional to expected attempts. A deterministic CPU *FrozenMonotonicOracle* serves as both Ariadne’s planning oracle and the public evaluator for every condition; on held-out sessions it is only moderately predictive (full-feature AUC 0.775; zero-history ablation AUC 0.611; no estimate over prerequisite-closed planner states is available), so our conclusions concern the planning landscape induced by this frozen model rather than true learner counterfactuals. The closed-loop design makes the exact Ariadne reference zero-regret by construction; the informative comparisons are the deviations of the remaining methods from that evaluator-defined optimum. Normalized regret of a sequence σ is $(C_{\text{eval}}(\sigma) - J_{\text{eval}}^*)/J_{\text{eval}}^*$, where C_{eval} is its cost and J_{eval}^* the exact optimum under the frozen evaluator for that target; repetitions are averaged within each target (100 seeds for Random Frontier, 20 per LLM condition, one run otherwise) before the ten targets are weighted equally. Table 1 reports the ten primary conditions: Ariadne, Frequency, BKT-set, and DKT-set planners under Exact and Greedy solvers, plus Random Frontier and Linear Syllabus orderings; Exact rows are computed by graph-search A*, and the LAO*-derived implementation returns identical values (Section 5). Four tool-free LLM conditions evaluated under the identical protocol land in the same easy band—best mean regret 6.6×10^{-5} (GPT-5.6 SOL, full context), with full context outperforming zero context for both models (Holm-adjusted exact permutation $p = 0.016$)—and are analyzed, with validity audits, in the supplementary material, together with full data processing, oracle validation, and provenance.

Q1: Does the diagnostic predict sequencing opportunity? For the *FrozenMonotonicOracle*, the target-level opportunity bound satisfies $\sum_v \Delta_v/J^* < 0.2\%$; the BKT- and DKT-derived set oracles expose wider envelopes of approximately 1.9%–4.2% and 2.1%–4.7%, respectively. Table 1 confirms that the largest target-equal mean regret is only 3.385×10^{-4} (0.034%), about one sixth of the Frozen-oracle envelope. The table also reveals oracle mismatch. BKT-set and DKT-set Exact match reverse-topological dynamic programming under their own objectives and improve or tie their corresponding Greedy policies internally, yet score worse under the Frozen evaluator. Exact optimization can therefore amplify learner-model error; this reversal is not a solver-correctness failure.

Q2: Why are the ECS32A instances easy? The objective leaves little value to optimize, and the prerequisite topology leaves little state space to search. Across the ten closures, width is at most four and the number of reachable ideals is at most 47, making exhaustive dynamic programming effectively instantaneous. This is a double degeneracy: weak feasible-state cost variation makes nearly every valid order competitive, while narrow prerequisite structure makes exact enumeration cheap.

A wider public topology. The public Junyi Academy corpus contains 835 exercises and exhibits the opposite structural regime (topology figure in the supplementary material). Its sink-target closures have median width eight and maximum width 18; their median ideal count is 96,608, and 68 closures exceed the 10^7 -state enumeration guard. A first-attempt bucketing proxy gives median proxy ratio $\sum_v \Delta_v/J_{\text{proxy}}^* = 5.6\%$ and a maximum of 12.5%, suggesting substantially greater state dependence than in ECS32A. This proxy is not directly comparable to the trained Frozen-

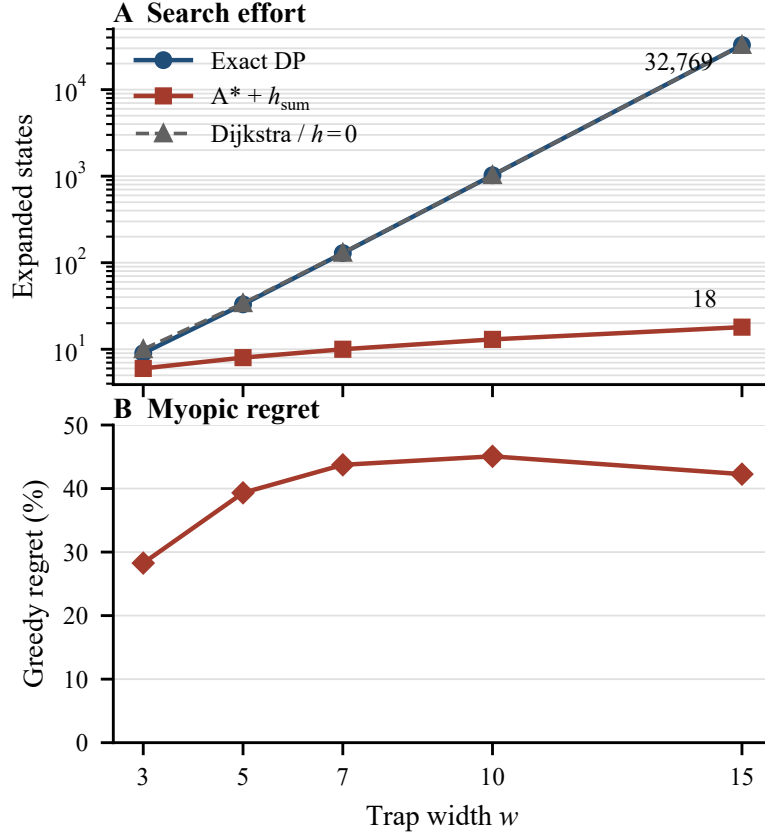


Figure 1: Sibling-transfer trap family. Exact DP expands all $2^w + 1$ ideals and Dijkstra nearly as many, whereas A^* with h_{sum} expands $w + 3$; greedy regret remains 28.3%–45.1%. These scalings are family-specific.

oracle diagnostic: students who have already mastered prerequisites may be systematically stronger, inflating apparent state dependence. The proxy is computed only on the observable restriction of each closure (241 closures yield values; median concept coverage 0.78), so it is an exploratory diagnostic rather than a certified full-closure bound. We therefore use Junyi only as topology and diagnostic evidence, not as planner evaluation.

Q3: When does the heuristic reduce exact search? On the narrow ECS32A closures, A^* and Dijkstra expand the same states. This is expected: the intermediate analytic certificate $\bar{p}_v = p(v, V \setminus \{v\})$ used in our experiments leaves a 2.5%–3.2% residual gap at s_0 , which exceeds the total cost spread $\sum_v \Delta_v / J^* < 0.2\%$ by an order of magnitude, so f -values cannot separate optimal from suboptimal states regardless of lattice size. On wider synthetic lattices where $\sum_v \Delta_v$ is large relative to the certificate slack, the distinction appears. Relative to Dijkstra, h_{sum} reduces expansions by about $2\times$ at width two, $4\times$ at width three, and 10 – $25\times$ at width five. Replacing the trivial certificate $\bar{p}_v = 1$ with this analytic bound further tightens search without changing the optimal value; complete counts are reported in the supplementary material.

Q4: Can myopic sequencing be substantially suboptimal? For each width $w \in \{3, 5, 7, 10, 15\}$, the sibling-transfer trap has a root r , an antichain of w siblings $a_1, \dots, a_w$, and a sink z , with prerequisite edges $r \rightarrow a_i \rightarrow z$. Success probabilities are $p(v, s) = \sigma(\varphi_v + 1.5 \sum_u w_{uv} \mathbf{1}[u \in s])$ with $T_v = 60$, $\varphi_r = 1$, $\varphi_z = 0$, $\varphi_{a_i} = -0.8 + 0.12i$; transfer weights form a non-prerequisite chain $w_{a_i, a_{i+1}} = 1$. Greedy favors later siblings (higher baseline p) but the forward order $a_1, \dots, a_w$ exploits sequential transfer. The poset width is exactly w , giving $2^w + 2$ reachable ideals. Figure 1 shows that DP expands $2^w + 1$ states while A^* expands only $w + 3$; greedy regret ranges from 28.3% to 45.1%. The $O(w)$ A^* profile depends on h_{sum} being near-tight on this family and is not a general complexity claim. By contrast, the NP-hardness family of Theorem 3 has $p \geq 1/2$, forcing $c(v, s) \in [1, 2]$ and a multiplicative ratio at most two; hardness there concerns exact optimization, not the magnitude of achievable gains.

All splits, seeds, model checkpoints, LLM prompts and responses, diagnostic artifacts, and aggregation scripts are frozen and hash-addressed; code accompanies the submission.

7 Conclusion

Under failure-invariant, time-homogeneous learner dynamics, instructional sequencing collapses exactly to deterministic shortest paths on the order-ideal lattice, yet remains NP-hard in general; residual acyclicity, bounded width, and—for uniform affine transfer—small distance to acyclicity identify tractable regimes. The $m\Delta$ diagnostic bounds the value of sequencing before optimization, and a consistent heuristic supports exact A* search on the implicit lattice.

Our empirical conclusions are model-relative: the closed-loop ECS32A evaluation makes exact-planner zero regret a protocol identity, not evidence of real-student optimality, and the Junyi analysis provides topology evidence without planner evaluation. The residual-acyclic and FPT results assume explicit nonnegative pairwise transfer (affine costs for the latter) and do not extend automatically to black-box oracles.

Extensions include belief-space planning for failure-dependent updates, nonmonotone state models for forgetting, validated oracles for Junyi-scale evaluation, and robust learner modeling to address misspecification. Planning is a lever; the learner model is its fulcrum.

References

- [1] Sal Khan. Khan academy’s Sal Khan on AI in education. Interview on *Decoder*, The Verge, 2025. <https://www.theverge.com/decoder-podcast-with-nilay-patel/766082/>, accessed July 2026.
- [2] Benjamin S. Bloom. The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6):4–16, 1984.
- [3] Kurt VanLehn. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4):197–221, 2011.
- [4] James A. Kulik and J. D. Fletcher. Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research*, 86(1):42–78, 2016.
- [5] Pavlos Stavrinos and Konstantin M. Zuev. Course-prerequisite networks for analyzing and understanding academic curricula. *Applied Network Science*, 8(1):1–37, 2023.
- [6] Weiming Lu, Yangfan Zhou, Jiale Yu, and Chenhao Jia. Concept extraction and prerequisite relation learning from educational data. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 9678–9685, 2019.
- [7] Benjamin Clement, Didier Roy, Pierre-Yves Oudeyer, and Manuel Lopes. Multi-armed bandits for intelligent tutoring systems. *Journal of Educational Data Mining*, 7(2):20–48, 2015.
- [8] Anna N. Rafferty, Emma Brunskill, Thomas L. Griffiths, and Patrick Shafto. Faster teaching via POMDP planning. *Cognitive Science*, 40(6):1290–1332, 2016.
- [9] Min Chi, Kurt VanLehn, Diane Litman, and Pamela Jordan. Empirically evaluating the application of reinforcement learning to the induction of effective and adaptive pedagogical strategies. *User Modeling and User-Adapted Interaction*, 21(1–2):137–180, 2011.
- [10] Qi Liu, Shiwei Tong, Chuanren Liu, Hongke Zhao, Enhong Chen, Haiping Ma, and Shijin Wang. Exploiting cognitive structure for adaptive learning. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 627–635, 2019.
- [11] Xinghe Cheng, Zihan Zhang, Jiapu Wang, Liangda Fang, Chaobo He, Quanlong Guan, Shirui Pan, and Weiqi Luo. GraphRAG-induced dual knowledge structure graphs for personalized learning path recommendation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pages 14610–14620, 2026.
- [12] Lev S. Vygotsky. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, 1978.
- [13] Dimitri P. Bertsekas and John N. Tsitsiklis. An analysis of stochastic shortest path problems. *Mathematics of Operations Research*, 16(3):580–595, 1991.
- [14] Dmitry Pasechnyuk-Vilensky. Order-sensitive sequential interventions on ideal lattices, 2026. arXiv:2604.26472.
- [15] Shayan Doroudi, Vincent Aleven, and Emma Brunskill. Where’s the reward? *International Journal of Artificial Intelligence in Education*, 29(4):568–620, 2019.

- [16] Qingyao Li, Wei Xia, Li'ang Yin, Jian Shen, Renting Rui, Weinan Zhang, Xianyu Chen, Ruiming Tang, and Yong Yu. Graph enhanced hierarchical reinforcement learning for goal-oriented learning path recommendation. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, pages 1318–1327, 2023.
- [17] Alex Graves, Marc G. Bellemare, Jacob Menick, Rémi Munos, and Koray Kavukcuoglu. Automated curriculum learning for neural networks. In *Proceedings of the 34th International Conference on Machine Learning*, pages 1311–1320, 2017.
- [18] Sanmit Narvekar, Bei Peng, Matteo Leonetti, Jivko Sinapov, Matthew E. Taylor, and Peter Stone. Curriculum learning for reinforcement learning domains: A framework and survey. *Journal of Machine Learning Research*, 21:1–50, 2020.
- [19] Sean A. Mochocki, Mark G. Reith, Brett J. Borghetti, Gilbert L. Peterson, John D. Jasper, and Laurence D. Merkle. Personalized learning path problem variations: Computational complexity and AI approaches. *IEEE Transactions on Artificial Intelligence*, 6(3):574–588, 2025.
- [20] Albert T. Corbett and John R. Anderson. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction*, 4(4):253–278, 1995.
- [21] Chris Piech, Jonathan Bassen, Jonathan Huang, Surya Ganguli, Mehran Sahami, Leonidas J. Guibas, and Jascha Sohl-Dickstein. Deep knowledge tracing. In *Advances in Neural Information Processing Systems*, volume 28, 2015.
- [22] Aritra Ghosh, Neil Heffernan, and Andrew S. Lan. Context-aware attentive knowledge tracing. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 2330–2339, 2020.
- [23] Jean-Paul Doignon and Jean-Claude Falmagne. Spaces for the assessment of knowledge. *International Journal of Man-Machine Studies*, 23(2):175–196, 1985.
- [24] Jean-Claude Falmagne and Jean-Paul Doignon. *Learning Spaces: Interdisciplinary Applied Mathematics*. Springer, Berlin, Heidelberg, 2011.
- [25] David Eppstein. Learning sequences: An efficient data structure for learning spaces. In Jean-Claude Falmagne, Dietrich Albert, Christopher Doble, David Eppstein, and Xiangen Hu, editors, *Knowledge Spaces: Applications in Education*, pages 287–304. Springer, Berlin, Heidelberg, 2013.
- [26] Peter E. Hart, Nils J. Nilsson, and Bertram Raphael. A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions on Systems Science and Cybernetics*, 4(2):100–107, 1968.
- [27] Eric A. Hansen and Shlomo Zilberstein. LAO*: A heuristic search algorithm that finds solutions with loops. *Artificial Intelligence*, 129(1–2):35–62, 2001.
- [28] Nils J. Nilsson. *Principles of Artificial Intelligence*. Morgan Kaufmann, 1980.
- [29] Junyao Ye, Jingyong Su, and Yilong Cao. A stochastic shortest path algorithm for optimizing spaced repetition scheduling. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 4381–4390, 2022.
- [30] Noga Alon. Ranking tournaments. *SIAM Journal on Discrete Mathematics*, 20(1):137–142, 2006.
- [31] Pierre Charbit, Stéphan Thomassé, and Anders Yeo. The minimum feedback arc set problem is NP-hard for tournaments. *Combinatorics, Probability and Computing*, 16(1):1–4, 2007.
- [32] Robert P. Dilworth. A decomposition theorem for partially ordered sets. *Annals of Mathematics*, 51(1):161–166, 1950.
- [33] George Osipov, Marcin Pilipczuk, and Magnus Wahlström. Parameterized complexity of MinCSP over the point algebra. In *Proceedings of the 32nd Annual European Symposium on Algorithms (ESA 2024)*, volume 308 of *LIPIcs*, pages 93:1–93:15, 2024.