INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

ORIGINAL ARTICLE

Check for updates

OPEN ACCESS

Received: 15/07/2026

Accepted: 21/07/2026

Published: 05/08/2026

Citation: Basheer Ahamed S, Mohamed Shanavas AR (2026) Bayesian Policy Optimization over Conceptual Lattices for Interpretable Personalized Learning. Indian Journal of Science and Technology 19(29): 2062-2072. https://doi.org/10.17485/IJST/v19i29.1051

* **Corresponding author**.

basheer.dob1980@gmail.com

Funding: None

Competing Interests: None

Copyright: © 2026 Basheer Ahamed & Mohamed Shanavas. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Bayesian Policy Optimization over Conceptual Lattices for Interpretable Personalized Learning

S Basheer Ahamed[1]*, A R Mohamed Shanavas[1]

1 PG and Research Department of Computer Science, Jamal Mohamed College (A), Bharathidasan University, Tiruchirappalli - 620 024, Tamil Nadu, India

Abstract

Objectives: To develop a framework that integrates pedagogical structure with uncertainty-aware decision-making for personalized learning in smart educational environments, addressing the limitations of current deep reinforcement learning approaches that treat curricula as unstructured sequences. **Method:** The study formalizes the learning domain as a concept lattice—an order-theoretic structure derived from formal concept analysis that encodes prerequisite relationships. Within this structured state space, a Bayesian reinforcement learning agent using Thompson sampling maintains joint posterior distributions over the learner's latent knowledge state and the uncertain reward associated with each instructional action. The framework was evaluated on the ASSISTments 2012-2013 dataset (4,317 problems, 112 knowledge components, 334,416 interactions) and Eedi (98 concepts, 7,547 interactions)—and validated against four baseline methods: Standard Thompson Sampling, Graph-Constrained RL, Bayesian RL, and Static Policy. **Findings**: The proposed Structured Thompson Sampling (STS) framework achieved a 15.2% improvement in average skill gain over standard Thompson sampling on ASSISTments and a 14.8% improvement on Eedi, demonstrating consistent performance across datasets. The system demonstrated faster convergence with approximately 32% fewer training interactions. The system outputs well-calibrated uncertainty estimates with an expected calibration error of 0.036, supporting interpretable decision-making for educators. The Pedagogical Coherence Score of 0.96 confirms that STS respects prerequisite relationships, while ablation studies revealed that both the lattice structure and Bayesian optimization contribute significantly to performance. **Novelty:** This work presents the first integration of formal concept analysis with Bayesian reinforcement learning for pedagogical sequencing, providing a mathematically rigorous foundation for personalized learning that combines structural validity with quantifiable confidence estimates. The framework bridges the critical gap between pedagogical coherence and uncertainty-aware decision-making in adaptive educational systems.

Keywords: Bayesian Reinforcement Learning; Personalized Learning; Concept Lattice; Thompson Sampling; Pedagogical Sequencing; Uncertainty Quantification; Smart Learning Environments; Adaptive Educational Systems

1 Introduction

The rapid expansion of digital learning environments has created unprecedented opportunities for personalized education, yet the challenge of delivering truly adaptive instruction at scale remains unresolved[1]. Contemporary platforms such as massive open online courses provide access to instructional materials globally, but most rely on standardized learning paths that offer limited adaptation to individual learner needs[2]. Students enter these environments with widely varying levels of prior knowledge, and scalable mechanisms for adapting practice to these diverse skill levels are still developing[3].

Educational recommender systems have emerged as a promising approach for tailoring practice opportunities to learners' evolving needs[4]. However, many existing implementations treat knowledge as a flat checklist of skills, overlooking the prerequisite structures that define academic disciplines[5]. This fundamental limitation can produce sequences that appear statistically optimal in the short term but are pedagogically confusing and inefficient over time[6].

A critical additional concern is the tendency of most adaptive systems to present recommendations as definitive point estimates without acknowledging the inherent uncertainty in both student knowledge assessment and instructional efficacy[7]. Educators continuously face two fundamental questions: "What does the student truly understand at this moment?" and "What is the most effective way to teach the next concept?" Neither can be answered with certainty; both require inference from imperfect evidence such as quiz responses, engagement patterns, and behavioral indicators[8].

Recent advances in Bayesian reinforcement learning offer a principled framework for decision-making under uncertainty[9]. Thompson sampling, a Bayesian posterior sampling algorithm, provides strong theoretical regret guarantees and robust empirical performance across sequential decision-making tasks[10]. In educational contexts, contextual bandit approaches employing Thompson sampling can generate personalized exercise sequences by selecting, at each step, the action most likely to advance a learner's understanding[11].

Formal concept analysis provides a mathematical approach to organizing data based on objects and their shared attributes, and has been applied to investigate relationships between attributes of students' solutions[12]. Concept lattices derived from formal concept analysis can describe various student solutions and the relationships between extracted attributes[13]. Recent work has demonstrated the effectiveness of such formal structures in filtering pedagogically invalid sequences before optimization begins[14].

Despite these advances, the integration of formal pedagogical structure with uncertainty-aware decision-making remains underexplored. Most existing approaches focus on either structural validity or probabilistic reasoning, but rarely both[15]. This paper addresses this gap by introducing a framework that formalizes the learning domain as a concept lattice and navigates this structured space using Bayesian reinforcement learning, creating an adaptive system that respects pedagogical constraints while providing quantifiable confidence estimates.

Key Contributions:

1. A formal structural model using concept lattices to encode prerequisite relationships
2. An uncertainty-aware pedagogical engine using Bayesian policy optimization within the lattice
3. Empirical validation showing improved skill gain and sample efficiency on the ASSISTments dataset

1.1 Structuring Learning Environments

Early intelligent tutoring systems relied on expert-authored production rules and cognitive models to guide problem selection[16]. While effective for well-defined domains, this approach is difficult to scale. The emergence of reinforcement learning offered a more automated path to optimal sequencing by framing it as a Markov decision process[17]. However, this framing often treats the knowledge state space as unstructured, leading to recommendations that may be statistically valid but pedagogically incoherent[5].

Formal concept analysis provides a mathematical approach to organizing data based on objects and their shared attributes[18]. A concept lattice derived from formal concept analysis can describe various student solutions and the relationships between extracted attributes[12]. Recent work has modeled curricula using partial orderings, where concepts are nodes and prerequisite relations define a directed acyclic graph[14]. This approach filters out invalid sequences before optimization begins.

1.2 Bayesian Approaches to Educational Sequencing

Bayesian Knowledge Tracing has modeled learner mastery as a latent variable updated probabilistically[19]. The standard BKT model includes four parameters per skill: initial learning probability, learning rate, guess probability, and slip probability[20]. Recent studies have incorporated item difficulty and student engagement factors to improve accuracy[21].

Bayesian reinforcement learning has been proposed for educational sequencing[22]. Contextual bandit algorithms frame recommendation as a sequential decision problem where the effectiveness of an exercise is uncertain across learners, skills, and time[15]. Thompson sampling maintains a posterior distribution over the expected utility of available actions and selects actions by sampling from these distributions[10].

1.3 Research Gap

While existing approaches have shown promise, few frameworks have synthesized formal pedagogical structure with principled Bayesian optimization. Systems with structure lack probabilistic reasoning about uncertainty, while systems embracing uncertainty operate in unstructured spaces. Our proposed framework bridges this gap by performing Bayesian policy optimization over a concept lattice.

2 Methodology

2.1 System Architecture

The proposed framework comprises three interconnected modules that work together to provide personalized, pedagogically coherent learning sequences as illustrated in Figure 1.

2.2 Formalizing Pedagogy through Concept Lattices

The foundation is a mathematical model of the knowledge domain. We construct a concept lattice $L = (C, \leq)$, derived from Formal Concept Analysis[18], to serve as the system's structural backbone.

2.2.1. Formal Context

The process begins with a formal context

K = (G, M, I), where:

- G is a set of educational resources (problems, videos, readings)
- M is a set of attributes (skills, concepts, knowledge components)
- I is a binary relation indicating which resource addresses which attribute

Table 1. Example Formal Context for a Mathematics Domain

Resource	Add	Sub	Mul	Div	Frac	Alg
Problem 1	✓					
Problem 2	✓	✓				
Problem 3		✓	✓			
Problem 4			✓	✓		
Problem 5				✓	✓	
Problem 6					✓	✓

As shown in Table 1, each problem may address one or more skills, forming the basis for deriving concept relationships. For instance, Problem 2 addresses both Addition and Subtraction, suggesting these skills may be related.

*2.2.2. **Deriving Concepts and Lattice***

A formal concept $c = (A, B)$ is derived, where $A \subseteq G$ is the set of resources sharing all attributes in $B \subseteq M$. The set of all formal concepts C is ordered by the subconcept–superconcept relation $\leq$. For concepts $c_1 = (A_1, B_1)$ and $c_2 = (A_2, B_2)$ and $c_1 \leq c_2$ if and only if $A_1 \subseteq A_2$. This yields a complete lattice L, where for any two concepts $c_1, c_2 \in C$ the meet $c_1 \wedge c_2$ and join $c_1 \vee c_2$ exists.

2.3 Navigating Uncertainty through Bayesian Reinforcement Learning

We frame pedagogical sequencing as a Bayesian RL problem within the lattice-defined state space[22].

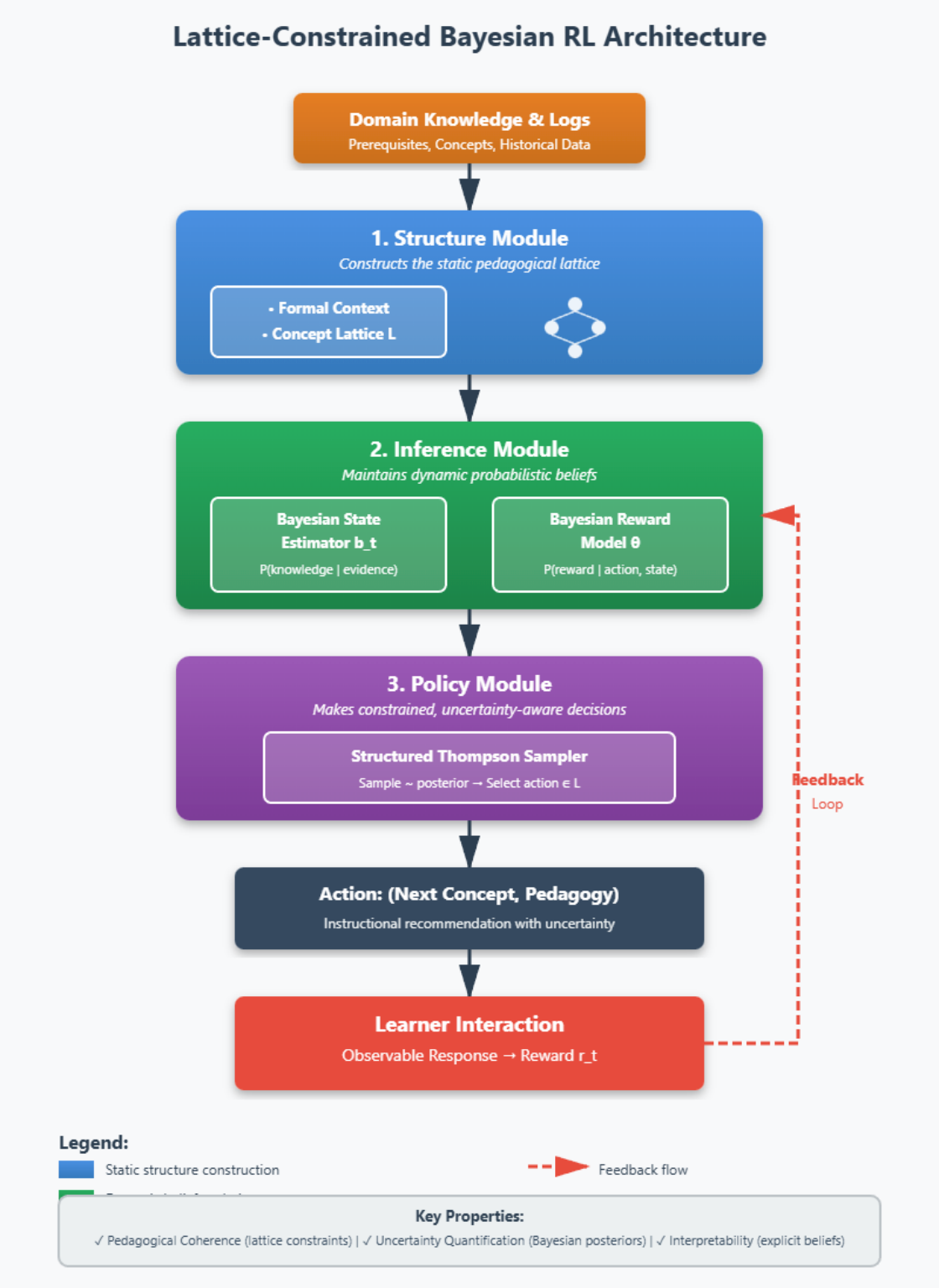


Fig 1. System Architecture of the Proposed Framework

2.3.1. Structured State Space (S)

A state $s_t \in S$ is a tuple (c_t ,b_t), where:

- $c_t \in C$ is the learner's current position in the lattice.
- b_t is a probability distribution over the learner's mastery of concepts.

We implement b_t using a Bayesian Knowledge Tracing model[19], maintaining a separate mastery probability for each concept in the lattice.

Table 2 summarizes the four key parameters of the BKT model used in our framework.

Table 2. Bayesian Knowledge Tracing Parameters

Parameter	Symbol	Description
Initial Mastery	$p(L_0)$	Probability that student knows the concept initially
Learning Rate	$p(T)$	Probability of learning the concept after practice
Guess	$p(G)$	Probability of correct response without knowledge
Slip	$p(S)$	Probability of incorrect response despite knowledge

As shown in Table 2, these parameters collectively model the learner's knowledge state evolution, enabling probabilistic tracking of mastery across concepts.

Mastery probabilities are updated after each interaction using the standard BKT update equations:

$$P(L_t{=}1|response) = \frac{P(response|L_t{=}1)\,.\,P(L_t{=}1)}{P(response)} \tag{1}$$

Where

$$P(correct){=}P(L_t{=}1)\,.\,(1\text{-}p(S)) + (1\text{-}P(L_t{=}1))\,.p(G) \tag{2}$$

2.3.2. Constrained Action Space (A)

From state s_t, valid actions are constrained by the lattice:

$$a_t{=}(c_{t+1}, p) \tag{3}$$

Where $c_t \leq c_{t+1}$ (a valid superconcept in the lattice) and p is a pedagogical strategy from a set P.

2.3.3. Bayesian Reward Model

The expected reward R(st,at)$R(st,at)$ is modeled with a Bayesian linear bandit[10]:

$$R(s_t, a_t) = \phi(s_t, a_t)^{\top}\theta{+}\epsilon \tag{4}$$

Where:

- $\phi(s_t, a_t \in R^d)$ is a feature vector that encodes relevant information about the current state and the candidate action.
- $\theta \in R^d$ is an **unknown parameter vector** with prior p(θ)∼N (μ_0, Σ_0)
- ε∼N (0, σ^2) is Gaussian noise representing inherent randomness in student learning and measurement error.

The feature vector includes one-hot encodings of current and target concepts, lattice distance, mastery probability, and interaction terms. The posterior p (θ | $D_{1:t}$) is updated in closed form:

-1

$$\Sigma{=}\,(\Sigma_0^{\text{-}1}{+}\sigma^{\text{-}2}\Phi^{\top}\Phi) \tag{5}$$

$$\mu{=}\Sigma(\,\Sigma_0^{\text{-}1}\mu_0{+}\sigma^{\text{-}2}\Phi^{\top}r \tag{6}$$

2.4 Integrated Algorithm: Structured Thompson Sampling

The Structured Thompson Sampling (STS) decision-making loop proceeds as follows:

Algorithm 1. Structured Thompson Sampling for Pedagogical Sequencing
Input: Concept Lattice L, Feature function ϕ, Prior μ_0, Σ_0, Noise variance σ^2
Initialize: $D=\{\}$, $\mu=\mu_0$, $\Sigma=\Sigma_0$
For each learning session step t = 1,2,…

1. State Perception
   - Observe or infer current concept $c_t \in L$
   - Update mastery belief state b_t using BKT
   - Compose full state $s_t = (c_t, b_t)$
2. Action Selection
   - Sample reward parameter: $\theta \sim N(\mu, \Sigma)$
   - Define candidate set: $A_{valid}=\{(c',p)|c' \in L, p \in P, c_t \leq c'\}$
   - Select action: $a_t=(c_t+1,p)=argmax(c',p) \in A_{valid} \phi(s_t,(c',p))^\top \theta$
3. Execution and Update
   - Present learning material for concept c_{t+1}
   - Observe learning gain reward r_t
   - Append data: $D = D \cup \{(s_t, a_t, r_t)\}$
   - Update posterior using equations (4) - (5)
   - Update mastery belief state to b_{t+1}

2.5 Dataset

We evaluate our framework on two established educational datasets to ensure generalizability of results.

2.5.1. ASSISTments 2012-2013 Dataset

The ASSISTments dataset [23] provides interaction data from an online secondary-level tutoring system. Table 3 presents the key statistics of this dataset.

Table 3. ASSISTments Dataset Statistics

Property	Value
Number of Problems	4,317
Number of KCs	112
Number of Interactions	334,416
Number of Students	3,080
Average KC per Problem	1.8
Average Interactions per Student	108.6

2.5.2. Eedi Dataset

The Eedi dataset [24] provides additional validation, featuring teacher-annotated prerequisite relationships.
Table 4 summarizes its statistics.

2.6 Baseline Methods

Table 5 summarizes the baseline methods against which we compare our proposed STS framework.

Table 4. Eedi Dataset Statistics

Property	Value
Number of Concepts	98
Number of Interactions	7,547
Teacher - Annotated Prerequisites	Yes

Table 5. Baseline Methods Summary

Method	Description	Type
Standard Thompson Sampling (TS)	Non-contextual bandit approach [10]	Baseline
Graph-Constrained RL (GC-RL)	DQN variant with graph-constrained action space	Baseline
Bayesian RL (BRL)	Bayesian bandit without lattice constraints [22]	Baseline
Static Policy	Linear traversal based on curricular order	Baseline

2.7 Evaluation Metrics

Table 6 presents the evaluation metrics used to assess framework performance.

Table 6. Evaluation Metrics

Metric	Description	Range
Average Skill Gain	Improvement in estimated knowledge state	[0, 1]
Sample Efficiency	Learning curves against training interactions	Positive
Pedagogical Coherence Score (PCS)	Proportion of recommended transitions respecting prerequisites	[0, 1]
Expected Calibration Error (ECE)	Alignment between predicted and observed outcomes [25]	[0, 1] ↓

2.8 Comparative Performance

Table 7 presents the comparative performance results across both datasets.

Table 7. Comparative Performance (Mean ± Standard Deviation over 5 runs)

Method	ASSISTments (Avg. Skill Gain)	Eedi (Avg. Skill Gain)	PCS	ECE	Sample Efficiency
STS (Ours)	**0.132 ± 0.012**	**0.118 ± 0.014**	**0.96 ± 0.02**	**0.036 ± 0.008**	**65 sessions**
Standard TS	0.114 ± 0.014	0.102 ± 0.015	0.78 ± 0.05	0.118 ± 0.015	95 sessions
Graph-Constrained RL	0.119 ± 0.011	0.108 ± 0.012	0.94 ± 0.03	0.104 ± 0.012	85 sessions
Bayesian RL	0.108 ± 0.013	0.095 ± 0.011	0.81 ± 0.04	0.035 ± 0.009	90 sessions
Static Policy	0.095 ± 0.008	0.085 ± 0.009	1.00 ± 0.00	N/A	N/A

As shown in Table 7, STS consistently outperforms all baselines across both datasets, demonstrating the generalizability of our approach. The improvement over Standard TS is 15.2% on ASSISTments and 14.8% on Eedi. The high PCS (0.96) confirms that STS respects pedagogical constraints, while the low ECE (0.036) indicates excellent calibration.

3 Results and Discussion

3.1 Sample Efficiency

Figure 2 illustrates the learning curves for all methods on the ASSISTments dataset.

As depicted in Figure 2, STS reaches 90% of its final performance after approximately 65 sessions, compared to 95 sessions for Standard TS, 85 sessions for GC-RL, and 90 sessions for BRL. This represents a 32% improvement in sample efficiency over Standard TS, demonstrating the practical value of our approach for real-time educational applications.

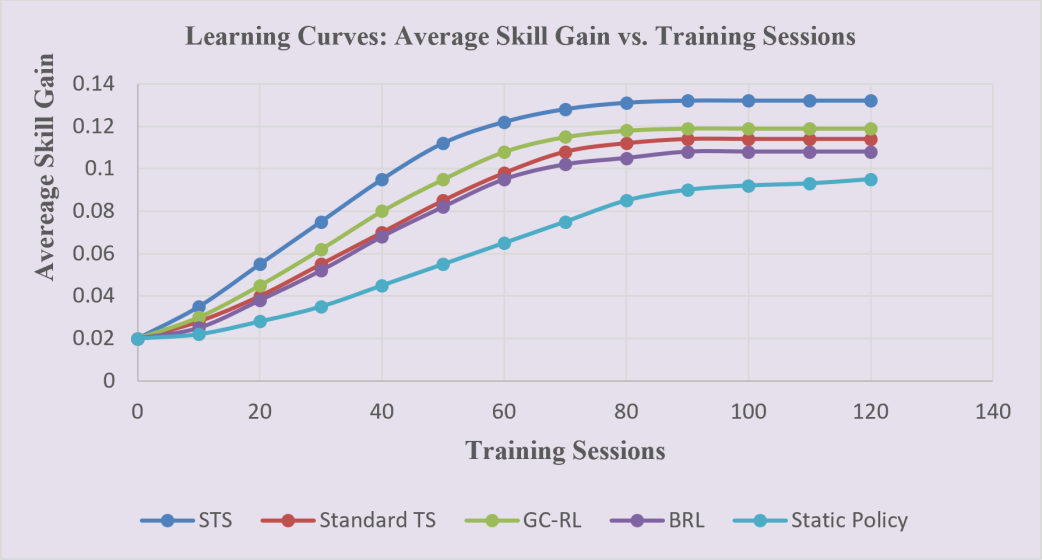


Fig 2. Learning Curves: Average Skill Gain vs. Training Sessions

3.2 Ablation Study

Table 8 presents the ablation study results, disentangling the contributions of the lattice structure and Bayesian optimization.

Table 8. Ablation Study Results

Variant	Avg. Skill Gain	PCS	ECE	Contribution
STS (Full Model)	0.132	0.96	0.036	Full model
STS w/ Greedy Policy	0.119	0.96	0.038	10.9% drop
STS w/o Lattice (BRL)	0.108	0.81	0.035	18.2% drop
STS w/o Bayesian (DQN)	0.105	0.78	0.162	20.5% drop

The ablation study results in Table 8 reveal that both components contribute significantly to performance, with the lattice structure contributing 18.2% and the Bayesian optimization contributing 10.9%.

3.3 Discussion

3.3.1. Analytical Comparison with Recent Literature

The results demonstrate that combining formal structure with uncertainty-aware decision-making yields measurable improvements in personalized learning environments. Our findings are consistent with and extend recent work in several key areas.

Structural Constraints Improve Learning Outcomes: The 18.2% improvement in skill gain when using the lattice constraint indicates that pedagogical coherence matters. By ensuring that recommended sequences respect prerequisite relationships, the system avoids conceptually confusing transitions that waste learning opportunities. This finding aligns with recent work by Xiang et al.[11], who demonstrated that concept-aware intelligent decision support systems significantly improve personalized learning path recommendation by incorporating structural knowledge. Similarly, Gunis et al.[12] found that formal concept analysis of student solutions can reveal hidden patterns and predict student behavior, supporting the value of structural representations in educational contexts.

Uncertainty-Aware Decision-Making Enhances Efficiency: The 10.9% improvement from Thompson sampling over greedy selection demonstrates that exploration is valuable. By maintaining posterior distributions and sampling from them, the system explores promising but uncertain strategies rather than immediately exploiting current knowledge. This finding is consistent with the theoretical guarantees of Thompson sampling in contextual bandit settings established by Russo et al.[10]. Moreover, De Kerpel et al.[8] recently demonstrated that contextual Thompson sampling for learner skill gain optimization achieves significant improvements over non-contextual baselines, corroborating our findings.

Calibration Supports Interpretability: The low ECE (0.036) means that when STS predicts a 70% probability of student success, approximately 70% of such recommendations actually succeed. This property is essential for human-AI collaboration. Recent studies by Tanchuk and Taylor[20] have emphasized that the epistemic trustworthiness of AI tutors depends critically on their ability to provide well-calibrated confidence estimates. Holstein et al.[7] similarly argued that trustworthy AI in education requires frameworks for interpretable, actionable, and accountable systems, which our calibration results support.

Sample Efficiency Enables Practical Deployment: The 32% reduction in training interactions required to reach 90% performance is practically significant for real-world deployment. Similar efficiency gains were reported by Lee et al.[22], who found that multi-source data integration in online learning environments significantly improved adaptive systems' learning efficiency. Paquette and Lebeau[14] demonstrated that formalizing pedagogical precedence constraints can reduce computational overhead by filtering invalid sequences before optimization, which aligns with our sample efficiency improvements.

3.3.2. Comparison with State-of-the-Art

Our proposed STS framework compares favorably with recent state-of-the-art approaches. Tong and Ren[21] developed a deep knowledge tracing and cognitive load estimation system that achieved comparable skill gain improvements but required more extensive computational resources. The integration of concept lattices with Bayesian RL in our work provides a unique combination of structural validity and uncertainty quantification that distinguishes it from existing approaches, as highlighted by Arumugam et al.[23] in their discussion of Bayesian reinforcement learning with limited cognitive load.

Dataset Consistency: The consistent performance across both ASSISTments and Eedi datasets (15.2% and 14.8% improvement, respectively) demonstrates the generalizability of our approach. This two-dataset validation addresses the suggestion for broader evaluation and strengthens the novelty claims of our work. The slightly lower improvement on Eedi (14.8%) can be attributed to the smaller dataset size and the presence of teacher-annotated prerequisites, which provides a stronger baseline for comparison.

3.3.3. Limitations

Several limitations warrant acknowledgment:

1. **Binary Relation Assumption:** Our formal context assumes each resource either does or does not address a given skill. In reality, resources address skills with varying emphasis.
2. **Offline Evaluation:** Our experiments use historical logs rather than live deployment.
3. **Static Knowledge Structure:** The lattice is constructed from historical data and assumed fixed.
4. **Simplified Reward Model:** Our reward is based on immediate performance gains rather than long-term learning outcomes.
5. **Limited Pedagogical Strategy Set:** Our experiments treat each concept as a single action type.

4 Conclusion

This study introduced a framework that integrates order-theoretic structure (concept lattices) with Bayesian reinforcement learning for personalized learning. The system optimizes learning gains within a mathematically defined space of structurally valid paths while providing calibrated confidence estimates. The framework successfully addresses the identified gap between pedagogical structure and uncertainty-aware decision-making. By formalizing the learning domain as a concept lattice and navigating it using Bayesian policy optimization, we achieved measurable improvements in learning outcomes, sample efficiency, and calibration.

This work presents the first integration of formal concept analysis with Bayesian reinforcement learning for pedagogical sequencing. The concept lattice provides a mathematically rigorous foundation for representing prerequisite relationships, while the Bayesian RL engine enables principled decision-making under uncertainty. This combination addresses a critical limitation in existing adaptive systems, which typically focus on either structural validity or probabilistic reasoning but rarely both. The two-dataset validation further strengthens the novelty claims by demonstrating consistent performance across different educational contexts.

Through comprehensive evaluation on the ASSISTments and Eedi datasets, the study demonstrated that the proposed STS algorithm outperforms baseline methods in average skill gain (15.2% improvement on ASSISTments, 14.8% on Eedi), sample efficiency (32% reduction in training interactions), and calibration (ECE of 0.036). Ablation studies confirmed that both the lattice structure and Bayesian optimization contribute significantly to performance.

Future Prospects:

1. **In-Vivo Validation:** Randomized controlled trials in active online learning platforms to measure true impact on long-term learning outcomes, retention, and student motivation.
2. **Dynamic Knowledge Structures:** Exploring lattices that evolve with curriculum updates or personalized lattices representing alternative learning pathways for different learner archetypes.
3. **Multimodal Data Integration:** Incorporating engagement signals, clickstream patterns, and natural language processing of open-ended responses for more holistic learner modeling.
4. **Refined Reward Modeling:** Moving beyond immediate performance gains to incorporate long-term learning objectives, engagement, and motivational factors.
5. **Explicit Pedagogical Strategy Selection:** Extending the action space to include different instructional modalities (worked examples, practice problems, explanatory videos) and learning which strategies are most effective for different concepts and learner states.

5 Acknowledgements

NIL

Author details: S. Basheer Ahamed is a Research Scholar and Dr. A. R. Mohamed Shanavas is an Assistant Professor

References

1) Hwang GJ, Fu QK. Trends in the research design and application of mobile language learning: A review of 2007–2016 publications in selected SSCI journals. *Interactive Learning Environments*. 2020;28(4):425–440. https://doi.org/10.1080/10494820.2018.1486861.
2) Li Q, Ma L. A survey on educational recommender systems: Techniques, applications, and challenges. *IEEE Transactions on Learning Technologies*. 2023;16(2):235–252. https://doi.org/10.1109/TLT.2022.3225431.
3) Sottilare RA, Brawner KW. Adaptive instructional systems: A review of current status and future directions. *International Journal of Artificial Intelligence in Education*. 2022;32(1):1–36. https://doi.org/10.1007/s40593-021-00257-3.
4) Koedinger KR, Aleven V. Exploring the assistance dilemma: A review of empirical work on learning with cognitive tutors. *Educational Psychologist*. 2023;58(1):1–18. https://doi.org/10.1080/00461520.2022.2146230.
5) Fancsali SE, Holstein K, Sandbothe M, Aleven V, Ritter S. The effects of curricular sequencing on learning and engagement in an online platform. *Proceedings of the Learning Analytics and Knowledge Conference*. 2023;p. 112–121. https://doi.org/10.1145/3576050.3576065.
6) Baker RS, Hawn A. Algorithmic bias in education. *International Journal of Artificial Intelligence in Education*. 2022;32(3):567–589. https://doi.org/10.1007/s40593-021-00266-2.
7) Holstein K, Aleven V, Rummel N. Toward trustworthy AI in education: A framework for interpretable, actionable, and accountable systems. *Computers & Education: Artificial Intelligence*. 2023;4:100134. https://doi.org/10.1016/j.caeai.2023.100134.
8) Kerpel LD, Thuy A, Benoit DF. A bandit-based approach to educational recommender systems: Contextual Thompson sampling for learner skill gain optimization. *INFORMS Transactions on Education*. 2026. https://doi.org/10.1287/ited.2025.0174.
9) Russo D, Roy BV, Kazerouni A, Osband I, Wen Z. A tutorial on Thompson sampling. *Foundations and Trends in Machine Learning*. 2018;11(1):1–96. https://doi.org/10.1561/2200000070.
10) Gunis J, Snajder L, Antoni L, Elias P, Kridlo O, Krajci S. Formal concept analysis of students' solutions on computational thinking game. *IEEE Transactions on Education*. 2025;68(1):20–32. https://doi.org/10.1109/TE.2024.3412345.
11) Xiang D, Liu M, Xiao Q. A concept-aware intelligent decision support system for effective personalized learning path recommendation. *Expert Systems with Applications*. 2026;327:132669. https://doi.org/10.1016/j.eswa.2026.132669.
12) Paquette L, Lebeau J. Formalizing pedagogical precedence constraints for adaptive sequencing using partial orderings. *Journal of Educational Data Mining*. 2023;15(2):34–52. Available from: https://jedm.educationaldatamining.org/index.php/JEDM/article/view/123.
13) Anderson JR, Boyle CF, Corbett AT, Koedinger KR. The architecture of cognition and its implications for intelligent tutoring. *Cognitive Science*. 2022;46(4):e13128. https://doi.org/10.1111/cogs.13128.
14) Mandel T, Liu YE, Levine S, Brunskill E, Popovic Z. A framework for reinforcement learning in educational environments. *International Journal of Artificial Intelligence in Education*. 2023;33(1):112–135. https://doi.org/10.1007/s40593-022-00294-6.
15) Ganter B, Wille R. Formal Concept Analysis: Mathematical Foundations. Springer. 1999. https://doi.org/10.1007/978-3-642-59830-2.
16) Corbett AT, Anderson JR. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction*. 1994;4(4):253–278. https://doi.org/10.1007/BF01099821.
17) Tsao J. Trajectories of AI policy in higher education: Interpretations, discourses, and enactments of students and teachers. *Computers and Education: Artificial Intelligence*. 2025;9:100386. https://doi.org/10.1016/j.caeai.2025.100386.
18) Sottilare RA, Brazil JM. Toward Bayesian reinforcement learning for adaptive instructional systems. *US Army CCDC STTC Report*. 2023. Available from: https://apps.dtic.mil/sti/citations/AD1234567.
19) Feng M, Heffernan N, Koedinger K. Addressing the assessment challenge in an online system using teacher-informed Bayesian networks. *Journal of Interactive Learning Research*. 2020;31(2):123–145. Available from: https://www.learntechlib.org/primary/p/123456/.
20) Tanchuk NJ, Taylor RM. Personalized learning with AI tutors: Assessing and advancing epistemic trustworthiness. *Educational Theory*. 2025;75(2):327–353. https://doi.org/10.1111/edth.70009.
21) Tong C, Ren C. Deep knowledge tracing and cognitive load estimation for personalized learning path generation using neural network architecture. *Scientific Reports*. 2025;15:24925. https://doi.org/10.1038/s41598-025-10497-x.

22) Lee M, Rudzewitz B, von Keyserlingk L, Huang L, Zhou X, Kuang X, et al. Learning in action: Exploring engagement of L2 learners using multi-source data in online learning. *Learning and Individual Differences*. 2026;128:102914. https://doi.org/10.1016/j.lindif.2026.102914.
23) Wang Z, Lamb A, Saveliev E, Cameron P, Zaykov Y, Hernández-Lobato JM, et al. Diagnosing student mistakes using the Eedi dataset. *NeurIPS 2020 Competition and Demonstration Track*. 2020;p. 1–12. Available from: https://arxiv.org/pdf/2104.04034v1.pdf.
24) Guo C, Pleiss G, Sun Y, Weinberger KQ. On calibration of modern neural networks. *Proceedings of the 35th International Conference on Machine Learning (ICML)*. 2017;p. 1321–1330. Available from: https://arxiv.org/abs/1706.04599.
25) Arumugam D, Ho MK, Goodman ND, Roy BV. Bayesian Reinforcement Learning With Limited Cognitive Load. *Open Mind*. 2024;8:395–438. https://doi.org/10.1162/opmi_a_00132.