COGNITIVE STRUCTURE GENERATION VIA DIFFUSION MODELS WITH POLICY OPTIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Cognitive structure (CS), a student's construction of concepts and inter-concept relations, has long been recognized as a foundational notion in educational psychology, yet remains largely unassessable in practice. Existing approaches such as knowledge tracing (KT) and cognitive diagnosis (CD) simplify and indirectly approximate CS, but they intertwine representation learning with prediction objectives, limiting generalization, interpretability, and reuse across tasks. To address this gap, we propose Cognitive Structure Generation (CSG), a task-agnostic framework that explicitly models CS through generative modeling. Based on educational theories, CSG first pretrains a Cognitive Structure Diffusion Probabilistic Model (CSDPM) and then applies reinforcement learning with SOLO-based hierarchical rewards to align generation with genuine cognitive development. By decoupling cognitive structure representation from downstream prediction, CSG produces interpretable and transferable cognitive structures that can be seamlessly integrated into diverse student modeling tasks. Experiments on four real-world datasets show that CSG yields more comprehensive representations, substantially improving performance while offering enhanced interpretability and modularity.

1 INTRODUCTION

Cognitive structure, originally conceived in topological psychology and later embraced by cognitive psychology in education (Piaget, 1952; Bruner, 2009; Ausubel, 1968), denotes the knowledge system within a student's mind, manifested as an inherent learning state. Through the learning processes, students continually integrate new concepts and reorganize existing ones to refine their cognitive structures for further learning. Formally, a cognitive structure can be modeled as an evolving *graph* (Novak & Gowin, 1984), with nodes and edges representing the student's construction of concepts and inter-concept relations, respectively (Steffe & Gale, 1995).

Cognitive structure assessment, has long been a central topic in psychometrics (Lord & Novick, 2008). Traditional methods primarily relied on expert-defined educational principles to directly calculate cognitive structure but lacked sufficient accuracy (Tatsuoka, 2009; Lin et al., 2016b). Considering that cognitive structure is an inherent learning state, researchers have shifted to indirectly measuring it based on students' responses to test items. Knowledge tracing (KT) (Corbett & Anderson, 1994) and cognitive diagnosis (CD) (Leighton & Gierl, 2007) are prototypical tasks. KT predicts the response r_t at time t as $P_{KT}(r_t) = f_{KT}\big(\boldsymbol{h}_t, \boldsymbol{\beta}_t; \Phi\big)$, where $\boldsymbol{h}_t$ is the student's latent state inferred from historical interactions before t, $\boldsymbol{\beta}_t$ is the tested item's features, and Φ denotes the model parameters (Abdelrahman et al., 2023). CD models the association between response r and student's cognitive state or ability $\boldsymbol{\theta}$ based on tested item β as $P_{CD}(r) = f_{CD}\big(\boldsymbol{\theta}, \boldsymbol{\beta}; \Omega\big)$,where Ω denotes the model parameters (Wang et al., 2024). Although recently emerged KT (Piech et al., 2015; Choi et al., 2020; Zhang et al., 2017) and CD (Cheng et al., 2019; Wang et al., 2020) models have achieved remarkable performance, they still face two foundational limitations.

First, both the student's latent state $\boldsymbol{h}_t$ in KT and the cognitive state or ability $\boldsymbol{\theta}$ in CD are typically narrowed to the student's construction of individual concepts, i.e. $\boldsymbol{h}_t, \boldsymbol{\theta} \rightarrow \mathbb{R}^L$ (where L is the number of concepts), and thus cannot model the student's construction of inter-concept relations necessary for modeling a complete cognitive structure and its holistic evolution during the real learning process. Although some studies have applied graph learning methods on static concept maps (Liu et al., 2019; Nakagawa et al., 2019; Tong et al., 2020) or heterogeneous interaction graphs (Gao et al., 2021; Yang et al., 2024) to obtain enhanced representations of $\boldsymbol{h}_t$ and $\boldsymbol{\theta}$, they only model students' construction on individual concepts and still do not explicitly model students' construction of inter-concept relations. Therefore, our core motivation is to explicitly and comprehensively model cognitive structure (CS), the states of the students' construction of concepts and inter-concept relations (Ausubel, 1968), which remains a foundational yet unassessable concept in educational practice.

Second, by definition, students' responses are only an external manifestation or an indirect indicator of their underlying learning state—namely, the cognitive structure in this paper, $\boldsymbol{h}_t$ in KT, and $\boldsymbol{\theta}$ in CD. Yet most existing models have become increasingly preoccupied with maximizing response prediction accuracy, often through extensive domain feature integration (Liu et al., 2021; Xu et al., 2023; Zhou et al., 2021), ever more sophisticated network designs and optimizations (Yang et al., 2023a;b; Li et al., 2024; Liu et al., 2024b; Chen et al., 2023), and so forth. While such directions improve accuracy, they still tightly couple state inference with response prediction, intertwining representation learning with prediction objectives, which restricts generalization, particularly when models are applied in cold-start or uncertain settings, and limits interpretability and modular reuse.

To bridge this gap, we propose **Cognitive Structure Generation (CSG)**, a task-agnostic framework that explicitly models CS through generative modeling, which decouples cognitive structure representation from downstream prediction. Guided by cognitive structure theory (Ausubel, 1968) and constructivism (Steffe & Gale, 1995), CSG aims to produce interpretable and transferable cognitive structures that can be seamlessly integrated into diverse student modeling tasks, thereby enhancing generalization, interpretability, and modularity. Specifically:

First, consider that a cognitive structure is manifested as a graph, we naturally cast *cognitive structure generation* as a *graph generation* task, and propose a *Cognitive Structure Diffusion Probabilistic Model* (CSDPM), whose forward diffusion and reverse denoising processes can learn the underlying distribution of real cognitive structures and produce novel ones. However, since real cognitive structures cannot be directly observed, we devise a rule-based method to infer students' construction of concepts and inter-concept relations from interaction logs, yielding a set of simulated cognitive structures, which is then used to pretrain the CSDPM and initialize its basic capability for CSG.

Second, although the cognitive structures sampled from the pretrained CSDPM match the distribution over simulated cognitive structures, they are insufficient to reflect the genuine levels of cognitive development (Flavell, 1977; Keil, 1992) that students achieve through their learning processes. To fill this gap, inspired by *the Structure of the Observed Learning Outcome (SOLO) taxonomy* (Biggs & Collis, 2014) that characterizes five levels of cognitive development, we define a fine-grained, hierarchical reward function. Using these reward signals, we optimize the policy of the denoising process via reinforcement learning to better align generation with genuine cognitive development.

To the end, the pretrained and fine-tuned CSDPM, has been fully equipped for cognitive structure generation, and the generated cognitive structures can be leveraged for diverse downstream student modeling tasks in the educational domain. To the best of our knowledge, we are the **first** to:

- Reformulate cognitive structure modeling as a cognitive structure generation task;
- Decouples cognitive structure representation from downstream prediction;
- Propose a CSDPM with a two-stage design, pretraining on simulated structures and fine-tuning via reinforcement learning with SOLO-based hierarchical rewards.

Experimental results on four popular real-world education datasets show that cognitive structures generated by CSG offer more comprehensive and effective representations for student modeling, substantially improving performance on KT and CD tasks while enhancing interpretability.

2 Related Works

We organize related works into three strands. **Cognitive Structure Modeling** has been rooted in psychology and education (Piaget, 1952; Ausubel, 1968), where traditional psychometric approaches construct rule-based graphs of students' concepts and relations but lack personalization. With the rise of learning analytics, researchers approximate cognitive structures from student responses via knowledge tracing (Piech et al., 2015; Choi et al., 2020) and cognitive diagnosis

(Leighton & Gierl, 2007; Cheng et al., 2019). KT methods employ hidden-state models, classifiers, or encoder–decoders, sometimes augmented with concept maps or heterogeneous graphs (Liu et al., 2019; Yang et al., 2024), while CD methods focus on fine-grained attributes (Xu et al., 2023). However, they tend to focus on the mastery of individual concepts, overlooking the holistic evolution of cognitive structures. They focus solely on students' mastery of individual concepts while overlooking their mastery of inter-concept relations, thereby hindering the modeling of their holistic evolution of cognitive structures. Recent attempts still rely on predefined graphs (Chen et al., 2024), leaving the task of holistic cognitive structure generation largely unexplored. **Graph Diffusion Probabilistic Models (DPMs)** extend deep generative frameworks such as autoregressive models, VAEs, GANs, and normalizing flows. Continuous-time DPMs (Jo et al., 2022) denoise Gaussian-corrupted graphs, whereas discrete variants (Vignac et al., 2023) use categorical transitions to better preserve sparsity. These advances demonstrate the potential of diffusion models for complex graph generation, yet their mechanisms remain to be adapted for the unique challenges of cognitive structure generation. **Optimization of DPMs** has increasingly leveraged reinforcement learning to align generative models with external objectives. Recent approaches in vision (Fan et al., 2023; Black et al., 2024) and graphs (Liu et al., 2024c) treat reverse diffusion as a Markov decision process optimized via policy gradients. Building on this line of work, we propose a SOLO-based reward to optimize the graph diffusion model for CSG, thereby aligning the generated structures more effectively with cognitive development levels. For a more comprehensive discussion of related studies, please refer to Appendix A.

3 The CSG Framework

3.1 Problem Formulation

Suppose a learning system is defined as $\mathcal{L} = \langle S, Q, K, R \rangle$, where $S = \{s_i\}_{i=1}^{N}$ is the set of N students, $Q = \{q_j\}_{j=1}^{M}$ the set of M questions, and $K = \{k_l\}_{l=1}^{L}$ the set of L knowledge concepts. Students answer questions from Q, generating response logs $R = \{r_{ij} \mid \text{student } s_i \text{ answered question } q_j\}$, where $r_{ij} = 1$ if s_i answers q_j correctly and $r_{ij} = 0$ otherwise. For each student s_i, the sequence of historical interactions up to timestamp T is denoted as $X_i^T = \{(q_j, r_{ij})^t\}_{t=1}^{T}$, where $(q_j, r_{ij})^t$ is the question–response pair at time step t.

A student s_i's cognitive structure at time T is defined as a graph $\mathcal{G}_i^T = (\mathcal{V}_i^T, \mathcal{E}_i^T)$. The node set $\mathcal{V}_i^T \in \mathbb{R}^{L \times c}$ represents s_i's construction states for the L concepts in K, and the edge set $\mathcal{E}_i^T \in \mathbb{R}^{L \times L \times c}$ represents the construction states of inter-concept relations, where c is the size of the discrete construction state space (e.g., "constructed" vs. "unconstructed"). Since we treat the cognitive structure as an undirected graph, all subsequent operations are applied to the upper-triangular entries $\mathcal{E}^+$ of $\mathcal{E}$, after which the matrix is symmetrized. Our goal is to generate $\mathcal{G}_i^T$ from X_i^T, formally defined as a mapping function $f_{CSG} : X_i^T \to \mathcal{G}_i^T$.

To implement this mapping, we propose the *Cognitive Structure Diffusion Probabilistic Model* (CSDPM). The CSDPM is first pretrained on simulated cognitive structures to initialize its generative capacity, and then fine-tuned via policy optimization to align generation with genuine cognitive development. The holistic structures produced by the optimized CSDPM can then be used in downstream tasks such as knowledge tracing (KT) and cognitive diagnosis (CD): $P_{KT}(r_{ij}^{T+1}) = f_{KT}\big(\mathcal{G}_i^T, \boldsymbol{\beta}(q_j^{T+1}); \Phi\big)$ and $P_{CD}(r_{ij}) = f_{CD}\big(\mathcal{G}_i^T, \boldsymbol{\beta}(q_j); \Omega\big)$, where $\boldsymbol{\beta}(q)$ denotes the embedding of question q, and Φ, Ω are model parameters.

The overall architecture of CSG is illustrated in Fig.1. The CSG framework consists of two stages: pretraining CSDPM and optimizing CSDPM, which we will detail in the following subsections.

3.2 Stage I: Pretraining CSDPM with Simulated Cognitive Structures

The goal of Stage I is to initialize the CSDPM so that it captures meaningful inductive biases about how students construct knowledge. Unlike other graph generation domains (Liu et al., 2024a; Zhang et al., 2024; Trivedi et al., 2024; Zhao et al., 2021), training here ideally requires access to ground-truth cognitive structures, which are not directly observable in practice. To address this, we design

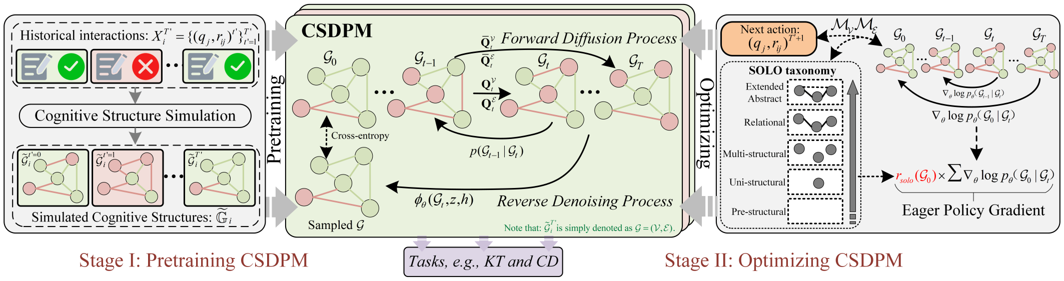


Figure 1: **(Overview).** The CSG includes two stages: pretraining CSDPM with Simulated Cognitive Structures and optimizing CSDPM via SOLO-based Hierarchical Reward. In stage I, the Cognitive Structure Simulation module (left) produces simulated cognitive structures that are used to pretrain the CSDPM. In stage II, a SOLO-based reward is introduced to optimize the CSDPM's policy via RL (right). Once pretrained and optimized, the CSDPM can generates cognitive structures, whose effectiveness is validated on KT and CD tasks through response prediction.

a simple *rule-based simulation process* grounded in theories of cognitive structure (Ausubel, 1968) and constructivist learning (Steffe & Gale, 1995), which serves as a proxy for pretraining.

Cognitive Structure Simulation. For each student s_i and interaction history X_i^T, we simulate a cognitive structure $\tilde{\mathcal{G}}_i^T = (\mathcal{V}_i^T, \mathcal{E}_i^T)$ by defining rule-based functions for concept states and relation states. Inspired by Lin et al. (2016a), we compute the construction state of concept k_l by

$$f_{UOC}(k_l, X_i^T) = \frac{\sum_{(q_j, r_{ij})^t \in X_i^T} \omega_{l,j} \cdot r_{ij}}{\sum_{(q_j, r_{ij})^t \in X_i^T} \omega_{l,j}}, \tag{1}$$

and the construction state of the relation between concepts k_a and k_b by

$$f_{UOR}(k_a, k_b, X_i^T) = \frac{\sum_{(q_j, r_{ij})^t \in X_i^T} \mathbf{1}\{\omega_{a,j} > 0 \wedge \omega_{b,j} > 0\}\,(\omega_{a,j} + \omega_{b,j}) r_{ij}}{\sum_{(q_j, r_{ij})^t \in X_i^T} \mathbf{1}\{\omega_{a,j} > 0 \wedge \omega_{b,j} > 0\}\,(\omega_{a,j} + \omega_{b,j})}. \tag{2}$$

Here, $\omega_{l,j}$ denotes the weight of concept k_l in question q_j, obtained by normalizing the Q-matrix across concepts that a question involves. This ensures that if a question taps multiple concepts, each receives a proportional share of weight. To better reflect real-world data and improve robustness, we also add small Gaussian perturbations to the Q-matrix entries. In Appendix H, we also provide a detailed example with full calculation steps.

Intuition. Equations 1 and 2 can be viewed as *weighted accuracies* that approximate the likelihood a student has constructed a given concept or relation. Eq. 1 averages the student's correctness on all questions involving concept k_l, weighted by how strongly the question tests k_l. Intuitively, if a student answers many k_l-related questions correctly, the ratio will approach 1, signaling that the concept is well constructed. Eq. 2 measures co-construction: it averages correctness on questions that involve *both* k_a and k_b, weighted by their combined relevance. Thus, if a student tends to succeed on joint questions, the relation between the two concepts is considered constructed.

From probabilities to discrete states. The f_{UOC} and f_{UOR} are empirical probabilities in $[0, 1]$. To map them into the discrete construction space Δ^c, we round the values and apply a one-hot encoding, yielding $\tilde{v}_{i,l}^T$ and $\tilde{e}_{i,a\text{–}b}^T$. By repeating this process for all students s_i and timestamps T, we obtain a set of simulated cognitive structures $\tilde{\mathbb{G}}$, which provides the training data to pretrain the CSDPM through forward diffusion and reverse denoising. For clarity, we drop student subscripts and time superscripts when unambiguous, writing $\mathcal{G}, v, e$ in place of $\tilde{\mathcal{G}}_i^T, \tilde{v}_{i,l}^T, \tilde{e}_{i,a\text{–}b}^T$. To avoid confusion between interaction timestamps and diffusion steps, we denote the former by T' now and reserve T for diffusion steps.

Forward Diffusion Process. Our CSDPM uses a forward diffusion process $q(\mathcal{G}_{1:T} \mid \mathcal{G}_0) = \prod_{t=1}^T q(\mathcal{G}_t \mid \mathcal{G}_{t-1})$ that gradually corrupts an initial simulated cognitive structure $\mathcal{G}_0 \sim q(\mathcal{G}_0)$ into near–uniform noise $q(\mathcal{G}_T)$ after T steps. The transition admits a node/edge factorization over the discrete construction state space:

$$q(\mathcal{V}_t \mid \mathcal{V}_{t-1}) = \prod_{v \in \mathcal{V}} q(v_t \mid v_{t-1}), \qquad q(\mathcal{E}_t \mid \mathcal{E}_{t-1}) = \prod_{e \in \mathcal{E}^+} q(e_t \mid e_{t-1}), \tag{3}$$

where $\mathcal{E}^+$ denotes the upper–triangular edge set (the graph is symmetrized afterwards). For each categorical node state $v \in \Delta^c$, we use the discrete noising kernel $q(v_t|v_{t-1}) = Cat(v_t; v_{t-1}\boldsymbol{Q}_t^v)$, $\boldsymbol{Q}_t^v = \alpha_t \boldsymbol{I} + (1-\alpha_t)\frac{\mathbf{1}_c\mathbf{1}_c^\top}{c}$ with schedule $\alpha_t \in [0,1]$ decreasing as t increases (Austin et al., 2021). Here, $\mathbf{1}_c$ is the c-dimensional all-ones vector and $\frac{\mathbf{1}_c\mathbf{1}_c^\top}{c}$ is the uniform transition over Δ^c. Thus, $\alpha_t = 1$ leaves the signal unchanged ($\boldsymbol{Q}_t^v = \boldsymbol{I}$), while smaller α_t mixes in more uniform noise. Let $\bar{\boldsymbol{Q}}_t^v = \boldsymbol{Q}_1^v\boldsymbol{Q}_2^v \cdots \boldsymbol{Q}_t^v$. Then the marginal and one–step posteriors admit closed forms:

$$q(v_t \,|\, v_0) \;=\; \mathrm{Cat}\big(v_t;\, v_0\, \bar{\boldsymbol{Q}}_t^v\big), q(v_{t-1} \,|\, v_t, v_0) \;=\; \mathrm{Cat}\left(v_{t-1};\, \frac{\big(v_t(\boldsymbol{Q}_t^v)^\top\big) \odot \big(v_0\bar{\boldsymbol{Q}}_{t-1}^v\big)}{v_0\bar{\boldsymbol{Q}}_t^v v_t^\top}\right), \tag{4}$$

where $\odot$ denotes element-wise product and all vectors are row-stochastic. As t grows and $\prod_{s=1}^t \alpha_s \to 0$, each node approaches the uniform distribution $q(v_T \mid v_0) \approx \mathrm{Cat}(v_T; \frac{\mathbf{1}_c}{c})$; edge transitions are defined analogously.

Reverse Denoising Process. Given the forward corruption, we learn a parametric reverse process $p_\theta(\mathcal{G}_{0:T}) = p(\mathcal{G}_T)\prod_{t=1}^T p_\theta(\mathcal{G}_{t-1} \mid \mathcal{G}_t)$ to recover cognitive structures from near–uniform noise $p(\mathcal{G}_T) \approx q(\mathcal{G}_T)$. We factor the reverse transition into nodes and edges:

$$p_\theta(\mathcal{G}_{t-1} \,|\, \mathcal{G}_t) \;=\; \prod_{v \in \mathcal{V}} p_\theta(v_{t-1} \,|\, \mathcal{G}_t) \prod_{e \in \mathcal{E}^+} p_\theta(e_{t-1} \,|\, \mathcal{G}_t). \tag{5}$$

Following the standard x_0-parameterization in discrete diffusion (Hasselt, 2010; Karras et al., 2022), each conditional can be expressed by marginalizing the exact posterior with a prediction of the clean state:

$$p_\theta(v_{t-1} \,|\, \mathcal{G}_t) = \sum_{v_0 \in \Delta^c} q(v_{t-1} \,|\, v_t, v_0)\; p_\theta(v_0 \,|\, \mathcal{G}_t)\,, \quad p_\theta(e_{t-1} \,|\, \mathcal{G}_t) = \sum_{e_0 \in \Delta^c} q(e_{t-1} \,|\, e_t, e_0)\; p_\theta(e_0 \,|\, \mathcal{G}_t)\,, \tag{6}$$

where a neural network predicts $p_\theta(v_0 \,|\, \mathcal{G}_t)$ and $p_\theta(e_0 \,|\, \mathcal{G}_t)$ given the noisy graph $\mathcal{G}_t$.

Training Objective. We pretrain on the simulated dataset $\tilde{\mathbb{G}}$ by maximizing the expected log-likelihood of clean structures conditioned on noisy ones:

$$J_{\mathrm{CSDPM}}(\theta) = \mathbb{E}_{\mathcal{G}_0 \sim \tilde{\mathbb{G}}, t \sim \mathcal{U}[\![1,T]\!]}\left[\mathbb{E}_{q(\mathcal{G}_t|\mathcal{G}_0)}\left[\log p_\theta(\mathcal{G}_0|\mathcal{G}_t)\right]\right], \tag{7}$$

with t sampled uniformly from $[\![1,T]\!]$. At generation time, we sample $\mathcal{G}_T \sim p(\mathcal{G}_T)$ and iteratively draw $\mathcal{G}_{t-1} \sim p_\theta(\mathcal{G}_{t-1} \,|\, \mathcal{G}_t)$ to obtain the trajectory $(\mathcal{G}_T, \mathcal{G}_{T-1}, \ldots, \mathcal{G}_0)$ for CSG.

Parametrization. We instantiate p_θ with an extended Graph Transformer Dwivedi & Bresson (2020); Vignac et al. (2023) that takes a noisy cognitive structure $\mathcal{G}_t = (\mathcal{V}_t, \mathcal{E}_t)$ as input and outputs distributions over clean node and edge states. Following (Vignac et al., 2023), we retain graph-theoretic feature integration and additionally condition the model on two auxiliary features: (i) a diffusion-step embedding that encodes the current noise level t, and (ii) an embedding of the student's interaction history $X^{T'}$, which provides task-specific guidance. An algorithmic summary is provided in Appendix B.

3.3 Stage II: Optimizing CSDPM via SOLO-based Hierarchical Reward

Building on the pretrained CSDPM, we further optimize its reverse denoising process to better align generation with genuine cognitive development. Inspired by the SOLO taxonomy (Biggs & Collis, 2014), we introduce a fine-grained hierarchical reward function and cast the denoising process as a reinforcement learning problem.

Standard Markov Decision Process Formulation. A standard MDP is specified by $(\mathcal{S}, \mathcal{A}, \mathcal{P}, r, \rho_0)$, where $\mathcal{S}$ is the state space, $\mathcal{A}$ the action space, $\mathcal{P}(\boldsymbol{s}'|\boldsymbol{s},\boldsymbol{a})$ the transition kernel, $r(\boldsymbol{s},\boldsymbol{a})$ the reward, and ρ_0 the initial-state distribution. Under a parameterized policy $\pi_\theta(\boldsymbol{a}|\boldsymbol{s})$, an agent generates a trajectory $\boldsymbol{\tau} = (\boldsymbol{s}_0, \boldsymbol{a}_0, \ldots, \boldsymbol{s}_T)$ by sampling $\boldsymbol{s}_0 \sim \rho_0$, then repeatedly choosing $\boldsymbol{a}_t \sim \pi_\theta(\cdot|\boldsymbol{s}_t)$, receiving reward $r(\boldsymbol{s}_t,\boldsymbol{a}_t)$, and transitioning via $\boldsymbol{s}_{t+1} \sim \mathcal{P}(\cdot|\boldsymbol{s}_t,\boldsymbol{a}_t)$. The return is $\mathcal{R}(\boldsymbol{\tau}) = \sum_{t=0}^T r(\boldsymbol{s}_t,\boldsymbol{a}_t)$, and the RL objective is to maximize $\mathcal{J}_{\mathrm{RL}}(\theta) = \mathbb{E}_{\boldsymbol{\tau} \sim p(\boldsymbol{\tau}|\pi_\theta)}[\mathcal{R}(\boldsymbol{\tau})]$. By the policy-gradient theorem (Grondman et al., 2012), this objective can be optimized using REINFORCE algorithm (Sutton et al., 1998):

$$\nabla_\theta \mathcal{J}_{\mathrm{RL}}(\theta) = \mathbb{E}_{\boldsymbol{\tau}\sim p(\boldsymbol{\tau}|\pi_\theta)}\Big[\sum_{t=0}^{T} \nabla_\theta \log \pi_\theta(\boldsymbol{a}_t|\boldsymbol{s}_t)\,\mathcal{R}(\boldsymbol{\tau})\Big]. \tag{8}$$

Mapping the Reverse Denoising Process to a T-step MDP. The pretrained CSDPM defines samples via its reverse denoising chain $p_\theta(\mathcal{G}_{0:T})$, but the marginal $p_\theta(\mathcal{G}_0)$ is intractable (Ho et al., 2020), and the reward $r(\mathcal{G}_0)$ is a black box with no gradient signal (Black et al., 2024). Following Fan et al. (2023); Liu et al. (2024c), we reformulate the denoising process as a T-step MDP:

$$\begin{aligned}
&\boldsymbol{s}_t \triangleq (\mathcal{G}_{T-t}, T-t), \quad \boldsymbol{a}_t \triangleq \mathcal{G}_{T-t-1},\\
&\pi_\theta(\boldsymbol{a}_t|\boldsymbol{s}_t) \triangleq p_\theta(\mathcal{G}_{T-t-1}|\mathcal{G}_{T-t}, T-t), \quad \mathcal{P}(\boldsymbol{s}_{t+1}|\boldsymbol{s}_t,\boldsymbol{a}_t) \triangleq \delta\big(\boldsymbol{s}_{t+1} - (\mathcal{G}_{T-t-1}, T-t-1)\big),\\
&r(\boldsymbol{s}_t,\boldsymbol{a}_t) \triangleq r(\mathcal{G}_0) \text{ if } t = T, \quad r(\boldsymbol{s}_t,\boldsymbol{a}_t) \triangleq 0 \text{ if } t < T,
\end{aligned} \tag{9}$$

where $\delta(\cdot)$ denotes a Dirac distribution, capturing the fact that transitions are deterministic: given $\boldsymbol{s}_t$ and $\boldsymbol{a}_t$, the next state is exactly $\boldsymbol{s}_{t+1} = (\mathcal{G}_{T-t-1}, T-t-1)$. The initial state $\boldsymbol{s}_0 = (\mathcal{G}_T, T)$ is the fully noised structure, and the terminal state $\boldsymbol{s}_T = (\mathcal{G}_0, 0)$ is the fully denoised structure.

SOLO-based Hierarchical Reward Function. After formulating the reverse denoising process of CSDPM as a MDP, we can optimize it for specific reward signals, which should ideally reflect the genuine levels of cognitive development that students achieve through their learning processes. Inspired by the SOLO taxonomy (Biggs & Collis, 2014), we propose a fine-grained, hierarchical reward function that scores the generated cognitive structures according to their alignment with the five levels of SOLO, which correspond to progressively better construction of concepts and inter-concept relations within more sophisticated cognitive structure.

Given a sampled structure $\mathcal{G}_0 = (\mathcal{V}_0, \mathcal{E}_0)$ and the next real interaction $(q_j, r_{ij})^{T'+1}$, we compare the predicted construction of relevant concepts and relations against the observed response. The matching degrees are

$$\mathcal{M}_{\mathcal{V}} = \frac{1}{|\mathcal{V}_{q_j}|}\sum_{v\in\mathcal{V}_{q_j}}(r_{ij} \veebar v), \qquad \mathcal{M}_{\mathcal{E}} = \frac{1}{|\mathcal{E}_{q_j}|}\sum_{e\in\mathcal{E}_{q_j}}(r_{ij} \veebar e), \tag{10}$$

where $\veebar$ denotes the XNOR operation. The SOLO-based reward is then

$$r_{solo}(\mathcal{G}_0) = \begin{cases} r_1, & \mathcal{M}_{\mathcal{V}} = 0,\\ r_2, & 0 < \mathcal{M}_{\mathcal{V}} < \kappa,\\ r_3, & \mathcal{M}_{\mathcal{V}} \geq \kappa \wedge \mathcal{M}_{\mathcal{E}} < \kappa,\\ r_4, & \kappa \leq \mathcal{M}_{\mathcal{V}} < 1 \wedge \kappa \leq \mathcal{M}_{\mathcal{E}} < 1,\\ r_5, & (\mathcal{M}_{\mathcal{V}} = 1 \wedge \mathcal{M}_{\mathcal{E}} \geq \kappa) \vee (\mathcal{M}_{\mathcal{V}} \geq \kappa \wedge \mathcal{M}_{\mathcal{E}} = 1), \end{cases} \tag{11}$$

with $r_1 < r_2 < r_3 < r_4 < r_5$ corresponding to SOLO levels: (i) *Pre-structural*: No meaningful concept alignment; (ii) *Uni-structural*: Alignment of a single or few concepts; (iii) *Multi-structural:* Alignment of multiple concepts, few relations; (iv) *Relational*: Alignment of multiple concepts and multiple relations; (v) *Extended abstract*: Alignment of almost all concepts and relations.

Since $\mathcal{M}_{\mathcal{V}}, \mathcal{M}_{\mathcal{E}} \in [0, 1]$, we adopt $\kappa = 0.5$ as the default threshold to distinguish "few" from "multiple" alignments. For instance, $0 < \mathcal{M}_{\mathcal{V}} < 0.5$ maps to the uni-structural level and is rewarded with r_2. Sensitivity analyses on thresholds and reward scales are reported in Appendix F.

Policy Gradient Estimation. With the reverse denoising process formulated as a T-step MDP, an agent generates a CSG trajectory $\boldsymbol{\tau} = (\mathcal{G}_T, \mathcal{G}_{T-1}, \ldots, \mathcal{G}_0)$, where $\boldsymbol{\tau} \sim p(\boldsymbol{\tau}|\pi_\theta) = p_\theta(\mathcal{G}_{0:T})$. Since rewards are only assigned at the terminal state, the cumulative return of any trajectory reduces to

$$\mathcal{R}(\boldsymbol{\tau}) = \sum_{t=0}^{T} r(\boldsymbol{s}_t, \boldsymbol{a}_t) = r_{solo}(\mathcal{G}_0). \tag{12}$$

The learning objective is therefore $\mathcal{J}_{\mathrm{RL}}(\theta) = \mathbb{E}_{\boldsymbol{\tau}\sim p(\boldsymbol{\tau}|\pi_\theta)}[\mathcal{R}(\boldsymbol{\tau})] = \mathbb{E}_{\mathcal{G}_{0:T}\sim p_\theta}[r_{solo}(\mathcal{G}_0)]$, which coincides with the end-structure objective $\mathcal{J}_{\mathcal{G}_0}(\theta)$.

A standard REINFORCE estimator gives the gradient

$$\nabla_\theta \mathcal{J}_{\mathrm{RL}}(\theta) = \mathbb{E}_{\mathcal{G}_{0:T}\sim p_\theta}\Big[r_{solo}(\mathcal{G}_0)\sum_{t=1}^{T}\nabla_\theta \log p_\theta(\mathcal{G}_{t-1}|\mathcal{G}_t)\Big], \tag{13}$$

but this estimator suffers from high variance on discrete graph diffusion. Following Liu et al. (2024c), we instead adopt the *eager policy gradient*, which directly reinforces the likelihood of high-reward terminal structures (i.e., the clean cognitive structures after T reverse denoising steps), rather than distributing credit iteratively via the term $\nabla_\theta \log p_\theta(\mathcal{G}_{t-1}|\mathcal{G}_t)$. With Monte Carlo estimation, the policy gradient can be modified as follows:

$$\nabla_\theta \mathcal{J}_{\mathrm{RL}}(\theta) \;\approx\; \frac{1}{|\mathcal{D}|}\sum_{d=1}^{|\mathcal{D}|}\frac{T}{|\mathcal{T}_d|}\sum_{t\in\mathcal{T}_d} r_{solo}\big(\mathcal{G}_0^{(d)}\big)\,\nabla_\theta \log p_\theta\big(\mathcal{G}_0^{(d)}\mid\mathcal{G}_t^{(d)}\big), \tag{14}$$

where $\mathcal{D}$ is the set of sampled trajectories, and $\mathcal{T}_d \subseteq [\![1,T]\!]$ is a random subset of timesteps for trajectory d. This estimator treats all trajectories ending at the same $\mathcal{G}_0$ as an equivalence class and reinforces them jointly, which significantly improves stability and sample efficiency. The full policy optimization procedure is summarized in Appendix C.

Two-Stage Training Paradigm. Overall, CSG training follows a two-stage paradigm reminiscent of pretraining–finetuning in LLMs (Devlin et al., 2019). In Stage I, simulated cognitive structures grounded in educational principles bootstrap the model with a meaningful prior, avoiding training from pure noise. In Stage II, the SOLO-based hierarchical reward evaluates generated structures by their alignment with progressively deeper levels of cognitive development, allowing the model to refine its pretrained representations and surpass handcrafted assumptions.

4 Experiments

Downstream Modeling for CSG. Because no ground-truth cognitive structures are directly observable, we evaluate the quality of the generated structures *indirectly* through downstream student modeling tasks. The intuition is that if the generated structures truly capture students' latent cognitive states, they should provide useful representations that improve prediction accuracy in standard benchmarks. We focus on two widely studied tasks: *knowledge tracing* (KT), which predicts a student's future performance, and *cognitive diagnosis* (CD), which infers a student's fine-grained mastery of concepts. These tasks serve as natural testbeds for assessing how well CSG captures interpretable and transferable cognitive information.

From Structures to Representations. To make the generated cognitive structures usable in downstream models, we apply the *edge-aware hard-clustering graph pooling* method from Zhu et al. (2023). This pooling jointly summarizes node- and edge-level construction features into a compact cognitive state vector for each student, which preserves both concept mastery and inter-concept relation information. The resulting vector is then concatenated with the embedding of the tested question and passed into task-specific output layers.

CSG-KT. For knowledge tracing, we use the pooled structure representation to augment a standard DKT (Piech et al., 2015) model. The prediction function is

$$P_{KT}(r_{ij}^{T'+1}) = f_{KT,\Phi}:\ \sigma\Big(\mathrm{FC}\Big(\mathrm{Pooling}(\mathcal{G}_i^{T'})\ \oplus\ emb(\boldsymbol{\beta}(q_j^{T'+1}))\Big)\Big), \tag{15}$$

where T' is the current interaction timestamp, $emb(\cdot)$ denotes the question embedding, $\oplus$ is concatenation, FC is a fully-connected layer, and σ is the sigmoid activation. This formulation allows the model to predict whether student s_i will answer question $q_j^{T'+1}$ correctly, informed by their generated cognitive structure.

CSG-CD. For cognitive diagnosis, we integrate the pooled structure representation into the NCD framework (Wang et al., 2020). The prediction function is

$$P_{CD}(r_{ij}) = f_{CD,\Omega}:\ \sigma\Big(\boldsymbol{\mathcal{Q}}_j \odot \Big((\mathrm{Pooling}(\mathcal{G}_i^{T'}) - \boldsymbol{h}_{diff}) \times \boldsymbol{h}_{disc}\Big)\Big), \tag{16}$$

where $\boldsymbol{\mathcal{Q}}_j$ is one row of the Q-matrix that specifies which concepts question q_j assesses. The vectors $\boldsymbol{h}_{diff}$ and $\boldsymbol{h}_{disc}$ are transformations of the question embedding $emb(\boldsymbol{\beta}(q_j))$, following Wang

Table 1: Performance comparison between CSG-KT and CSG-CD with their baselines on different datasets, averaged over five-fold cross-validation. Statistical significance is assessed via the Wilcoxon rank-sum test, with * ($p < 0.05$), ** ($p < 0.01$), and *** ($p < 0.001$).

Category	**Model**	**Math1**			**Math2**			**FrcSub**			**NIPS**		
	Metrics	AUC↑	ACC↑	RMSE↓	AUC↑	ACC↑	RMSE↓	AUC↑	ACC↑	RMSE↓	AUC↑	ACC↑	RMSE↓
KT	DKT	0.7735	0.7082	0.4524	0.7381	0.6678	0.4600	0.8202	0.7529	0.3392	0.6593	0.6214	0.4690
	SAKT	0.7612	0.7017	0.4552	0.7250	0.6583	0.4618	0.8113	0.7513	0.3419	0.6531	0.6176	0.4710
	GKT	0.7843	0.7147	0.4493	0.7463	0.6759	0.4519	0.8247	0.7608	0.3360	0.6841	0.6339	0.4645
	SKT	0.7895	0.7181	0.4489	0.7529	0.6842	0.4492	0.8385	0.7696	0.3338	0.6985	0.6429	0.4637
	GRKT	0.7943	0.7242	0.4461	0.7618	0.6976	0.4448	0.8418	0.7754	0.3280	0.7070	0.6501	0.4601
	MIKT	0.8030	0.7281	0.4412	0.7701	0.7017	0.4426	0.8472	0.7804	0.3253	0.7147	0.6570	0.4583
	ENAS-KT	0.8103	0.7326	0.4334	0.7722	0.7120	0.4405	0.8506	0.7865	0.3207	0.7233	0.6634	0.4565
	simpleKT	0.8074	0.7304	0.4390	0.7713	0.7083	0.4411	0.8485	0.7844	0.3232	0.7191	0.6618	0.4542
	PSI-KT	0.8118	0.7392	0.4317	0.7759	0.7140	0.4403	0.8533	0.7908	0.3309	0.7260	0.6687	0.4520
	CSG-KT	**0.8220***	**0.7412****	**0.4283****	**0.7772***	**0.7197***	**0.4390***	**0.8636***	**0.8022***	**0.3192****	**0.7413****	**0.6757****	**0.4511****
CD	IRT	0.7356	0.7179	0.4279	0.7589	0.6981	0.4516	0.7414	0.7091	0.3944	0.7489	0.6907	0.4516
	MIRT	0.7482	0.7347	0.4256	0.7699	0.7038	0.4478	0.8086	0.7745	0.3589	0.7589	0.7017	0.4483
	NCD	0.7691	0.7459	0.4084	0.7781	0.7182	0.4456	0.8250	0.8042	0.3498	0.7697	0.7113	0.4412
	RCD	0.7861	0.7584	0.4033	0.7911	0.7275	0.4406	0.8321	0.8178	0.3419	0.7736	0.7171	0.4345
	HyperCDM	0.7876	0.7599	0.4016	0.7972	0.7320	0.4383	0.8417	0.8239	0.3387	0.7821	0.7209	0.4301
	DisenGCD	0.7983	0.7628	0.4001	0.8039	0.7457	0.4324	0.8559	0.8375	0.3342	0.7886	0.7311	0.4275
	CSG-CD	**0.8133***	**0.7710****	**0.3987*****	**0.8179****	**0.7521***	**0.4270*****	**0.8699*****	**0.8451***	**0.3152*****	**0.8036***	**0.7507****	**0.4242*****

et al. (2020). Here, $\odot$ denotes element-wise product, and $\times$ is scalar-vector multiplication. This formulation enables diagnosis of whether the student's structure $\mathcal{G}_i^{T'}$ is consistent with their actual response r_{ij}. Both CSG-KT and CSG-CD are trained using cross-entropy loss between the predicted response probabilities and the ground-truth responses.

Experimental Settings. We conduct experiments on four real-world datasets of varying scales: Math1, Math2, FrcSub, and NIPS34[1], with dataset statistics provided in Appendix D. To evaluate the usefulness of the generated cognitive structures, we compare against representative baselines in both KT and CD. For KT, we include DKT (Piech et al., 2015), SAKT (Pandey & Karypis, 2019), GKT (Nakagawa et al., 2019), SKT (Tong et al., 2020), GRKT (Cui et al., 2024), MIKT (Sun et al., 2024), ENAS-KT (Yang et al., 2023a), simpleKT (Liu et al., 2023), PSI-KT (Zhou et al., 2024), and our CSG-KT. For CD, we include IRT (Cai et al., 2016), MIRT (Ackerman et al., 2003), NCD (Wang et al., 2020), RCD (Gao et al., 2021), HyperCDM (Shen et al., 2024), DisenGCD (Yang et al., 2024), and our CSG-CD.

Our goal is not to exhaustively benchmark every KT/CD model, but rather to assess how well CSG-generated structures serve as general-purpose representations across tasks. Accordingly, we select representative baselines from three categories: (i) classical models (e.g., DKT, IRT), (ii) structural models (e.g., GKT, SKT, GRKT, RCD), and (iii) recent state-of-the-art methods (e.g., simpleKT, PSI-KT, ENAS-KT, HyperCDM, DisenGCD). For evaluation, we follow prior work and report AUC (Bradley, 1997), accuracy (ACC), and root mean square error (RMSE). Further implementation details are provided in Appendix E.

In all experiments, student interaction data are randomly split at the level of individual interaction records using an 8:1:1 ratio. Importantly, the split is *disjoint*: no test interaction ever appears in the training set, and no model is trained on test data. For evaluation, CSG generates cognitive structures from a student's interaction history up to time T'. For KT, these structures are used to predict the response at $T' + 1$. For CD, the model is never exposed to the target response r_{ij} for the item it is asked to predict. This ensures that evaluation strictly measures generalization rather than memorization, and that no information leaks from training to testing.

Overall Performance. Table 1 reports the performance of CSG-KT, CSG-CD, and all KT/CD baselines on four public datasets, measured by average AUC, ACC, and RMSE over 5-fold cross-validation, with the best scores highlighted in bold. We observe: **(i)** CSG-KT not only substantially outperforms classical knowledge tracing models (e.g., DKT, SAKT), but also delivers clear gains over graph-based methods that model only concept construction without inter-concept relations (e.g., GKT, SKT, GRKT), and even surpasses recent SOTA approaches (e.g., PSI-KT, MIKT, ENAS-KT). Similarly, CSG-CD markedly improves upon classical parameter-estimation models (e.g., IRT, MIRT), achieves significant gains over neural models that represent student ability only at the concept level (e.g., NCD), and also exceeds heterogeneous graph-based SOTA methods. These results indicate that generative cognitive structures provide more comprehensive and accurate representations of student learning states, while capturing their dynamic evolution over time. **(ii)** Across

[1]Math1, Math2, and FrcSub are available at `http://staff.ustc.edu.cn/~qiliuql/data/math2015.rar`. NIPS34 is available at `http://ednet-leaderboard.s3-website-ap-northeast-1.amazonaws.com/`

Table 3: Ablation study on the impact of CSG variants for KT and CD across multiple datasets.

Category	Model	Math1			Math2			FrcSub			NIPS		
	Metrics	AUC↑	ACC↑	RMSE↓	AUC↑	ACC↑	RMSE↓	AUC↑	ACC↑	RMSE↓	AUC↑	ACC↑	RMSE↓
KT	V_1-KT	0.7842	0.7050	0.4496	0.7276	0.6745	0.4571	0.8144	0.7486	0.3455	0.6807	0.6504	0.4697
	V_2-KT	0.7991	0.7196	0.4433	0.7421	0.6887	0.4543	0.8288	0.7630	0.3397	0.6951	0.6647	0.4674
	V_3-KT	0.8042	0.7343	0.4472	0.7567	0.6930	0.4511	0.8433	0.7775	0.3241	0.7196	0.6691	0.4663
	V_4-KT	0.8085	0.7351	0.4413	0.7614	0.6974	0.4491	0.8479	0.7821	0.3287	0.7242	0.6697	0.4604
	V_5-KT	0.8111	0.7387	0.4322	0.7758	0.7184	0.4379	0.8598	0.7882	0.3262	0.7318	0.6730	0.4528
	CSG-KT	**0.8220**	**0.7412**	**0.4283**	**0.7772**	**0.7197**	**0.4390**	**0.8636**	**0.8022**	**0.3192**	**0.7413**	**0.6757**	**0.4511**
CD	V_1-CD	0.7870	0.7477	0.4218	0.7967	0.7277	0.4508	0.8210	0.8063	0.3475	0.7671	0.7068	0.4411
	V_2-CD	0.7913	0.7520	0.4157	0.8008	0.7319	0.4471	0.8354	0.8138	0.3309	0.7713	0.7210	0.4371
	V_3-CD	0.7958	0.7665	0.4098	0.8051	0.7463	0.4406	0.8601	0.8385	0.3276	0.7857	0.7254	0.4313
	V_4-CD	0.7965	0.7669	0.4041	0.8086	0.7469	0.4395	0.8650	0.8434	0.3275	0.7903	0.7300	0.4257
	V_5-CD	0.7985	0.7673	0.4030	0.8169	0.7473	0.4377	0.8661	0.8438	0.3205	0.7997	0.7392	0.4353
	CSG-CD	**0.8133**	**0.7710**	**0.3987**	**0.8179**	**0.7521**	**0.4270**	**0.8699**	**0.8451**	**0.3152**	**0.8036**	**0.7507**	**0.4242**

datasets of very different scales and interaction densities, both CSG-KT and CSG-CD consistently deliver robust performance, underscoring the general applicability of our framework. We note that we employed simple KT/CD models with CSG to demonstrate effectiveness and reduce confounding factors, leaving adaptation to advanced methods for future work.

Ablation Study. We evaluate several variants of our framework by comparing their prediction performance on sampled cognitive structures, as summarized in Table 2: **(i)** $\mathbf{V}_1$ uses only the rule-based simulated structures without any learning; **(ii)** $\mathbf{V}_2$ pretrains CSDPM on simulated structures but does not apply RL optimization; **(iii)** $\mathbf{V}_3$ skips pretraining and applies RL with a generic reward $r(\cdot)$; **(iv)** $\mathbf{V}_4$ skips pretraining and applies RL with the SOLO-based reward $r_{solo}(\cdot)$; **(v)** $\mathbf{V}_5$ combines pretraining with RL under the generic reward; and **(vi)** CSG is our complete framework with both pretraining and SOLO-based optimization. The generic reward $r(\cdot)$ does not differentiate developmental levels and simply sums $\mathcal{M}_{\mathcal{V}}$ and $\mathcal{M}_{\mathcal{E}}$ into a single scalar.

Table 2: Detailed configurations of CSG variants used in the ablation study.

Variants	Pretraining	Optimization	
		$r(\cdot)$	$r_{solo}(\cdot)$
V_1	✗	✗	✗
V_2	✓	✗	✗
V_3	✗	✓	✗
V_4	✗	✗	✓
V_5	✓	✓	✗
CSG	✓	✗	✓

For a fair comparison, we use the rule-based simulated set $\tilde{\mathbb{G}}$ for $\mathbf{V}_1$, and sample the corresponding generated set $\mathbb{G}_0$ for variants $\mathbf{V}_2$–$\mathbf{V}_5$. Each variant is then used to independently train and evaluate downstream KT and CD models, denoted as $\mathbf{V}_i$**-KT** and $\mathbf{V}_i$**-CD**, respectively, for $i = 1, \ldots, 5$.

Results in Table 3 show several key findings: **(i)** Overall, performance steadily improves from the simplest variant $\mathbf{V}_1$ through $\mathbf{V}_5$ to our full CSG, for both KT and CD tasks. **(ii)** Despite its simplicity, $\mathbf{V}_1$ performs competitively with classical baselines (e.g., DKT for KT, IRT and NCD for CD), validating that our rule-based simulation already provides a strong approximation of students' learning states. On Math1, Math2, and FrcSub, where sequences are short but coverage is high, this simulation is especially effective; on NIPS34, longer interaction sequences offset lower coverage, yielding similarly strong outcomes. **(iii)** $\mathbf{V}_3$ generally outperforms $\mathbf{V}_2$, suggesting that task-driven RL optimization can capture hidden learning patterns and incorporate them into generated structures. **(iv)** The improvements of $\mathbf{V}_4$ over $\mathbf{V}_3$, and of full CSG over $\mathbf{V}_5$, highlight the value of explicitly modeling developmental levels and confirm the effectiveness of SOLO-based hierarchical rewards. Additional analyses on hyperparameters, inference time, and visualization are provided in Appendix F, G, and I.

5 Conclusion

In this work, we introduced Cognitive Structure Generation (CSG), a framework for modeling students' evolving cognitive structures with a graph diffusion model. By decoupling structure representation from downstream prediction, CSG produces explicit cognitive structures that align with genuine developmental patterns. Our two-stage design first pretrains on simulated structures grounded in educational theory, then optimizes with reinforcement learning guided by a SOLO-based hierarchical reward to capture authentic levels of cognitive growth. Experiments on four real-world datasets show that CSG consistently improves performance on knowledge tracing (KT) and cognitive diagnosis (CD), while also enhancing generalizability, interpretability, and modular design. These results highlight the promise of holistic cognitive structure modeling as a foundation for more effective and transparent educational intelligence systems.

REFERENCES

Ghodai Abdelrahman, Qing Wang, and Bernardo Nunes. Knowledge tracing: A survey. *ACM Computing Surveys*, 55(11):1–37, 2023.

Terry A Ackerman, Mark J Gierl, and Cindy M Walker. Using multidimensional item response theory to evaluate educational and psychological tests. *Educational Measurement: Issues and Practice*, 22(3):37–51, 2003.

Jacob Austin, Daniel D. Johnson, Jonathan Ho, Daniel Tarlow, and Rianne van den Berg. Structured denoising diffusion models in discrete state-spaces. In Marc'Aurelio Ranzato, Alina Beygelzimer, Yann N. Dauphin, Percy Liang, and Jennifer Wortman Vaughan (eds.), *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual*, pp. 17981–17993, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/958c530554f78bcd8e97125b70e6973d-Abstract.html.

David Paul Ausubel. Educational psychology: A cognitive view. *Holt, Rinehart and Winston*, 1968. URL https://archive.org/details/in.ernet.dli.2015.112045/page/n397/mode/2up.

John B Biggs and Kevin F Collis. *Evaluating the quality of learning: The SOLO taxonomy (Structure of the Observed Learning Outcome)*. Academic press, 2014.

Kevin Black, Michael Janner, Yilun Du, Ilya Kostrikov, and Sergey Levine. Training diffusion models with reinforcement learning. In *The Twelfth International Conference on Learning Representations, ICLR 2024, Vienna, Austria, May 7-11, 2024*. OpenReview.net, 2024. URL https://openreview.net/forum?id=YCWjhGrJFD.

Andrew P Bradley. The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern recognition*, 30(7):1145–1159, 1997.

Derek C Briggs and Ruhan Circi. Challenges to the use of artificial neural networks for diagnostic classifications with student test data. *International Journal of Testing*, 17(4):302–321, 2017.

Jerome S Bruner. *The process of education*. Harvard university press, 2009.

Li Cai, Kilchan Choi, Mark Hansen, and Lauren Harrell. Item response theory. *Annual Review of Statistics and Its Application*, 3(1):297–321, 2016.

Xiangzhi Chen, Le Wu, Fei Liu, Lei Chen, Kun Zhang, Richang Hong, and Meng Wang. Disentangling cognitive diagnosis with limited exercise labels. In Alice Oh, Tristan Naumann, Amir Globerson, Kate Saenko, Moritz Hardt, and Sergey Levine (eds.), *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023*, 2023. URL http://papers.nips.cc/paper_files/paper/2023/hash/3a14ae9951e8153a8fc814b5f506b5b7-Abstract-Conference.html.

Zhifu Chen, Hengnian Gu, Jin Peng Zhou, and Dongdai Zhou. Enhancing cognitive diagnosis by modeling learner cognitive structure state, 2024. URL https://arxiv.org/abs/2412.19759.

Song Cheng, Qi Liu, Enhong Chen, Zai Huang, Zhenya Huang, Yiying Chen, Haiping Ma, and Guoping Hu. Dirt: Deep learning enhanced item response theory for cognitive diagnosis. In *Proceedings of the 28th ACM international conference on information and knowledge management*, pp. 2397–2400, 2019.

Weihua Cheng, Hanwen Du, Chunxiao Li, Ersheng Ni, Liangdi Tan, Tianqi Xu, and Yongxin Ni. Uncertainty-aware knowledge tracing. In Toby Walsh, Julie Shah, and Zico Kolter (eds.), *AAAI-25, Sponsored by the Association for the Advancement of Artificial Intelligence, February 25 - March 4, 2025, Philadelphia, PA, USA*, pp. 27905–27913. AAAI Press, 2025. doi: 10.1609/AAAI.V39I27.35007. URL https://doi.org/10.1609/aaai.v39i27.35007.

Youngduck Choi, Youngnam Lee, Junghyun Cho, Jineon Baek, Byungsoo Kim, Yeongmin Cha, Dongmin Shin, Chan Bae, and Jaewe Heo. Towards an appropriate query, key, and value computation for knowledge tracing. In *Proceedings of the seventh ACM conference on learning@ scale*, pp. 341–344, 2020.

Albert T Corbett and John R Anderson. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User modeling and user-adapted interaction*, 4:253–278, 1994.

Jiajun Cui, Hong Qian, Bo Jiang, and Wei Zhang. Leveraging pedagogical theories to understand student learning process with graph-based reasonable knowledge tracing. In Ricardo Baeza-Yates and Francesco Bonchi (eds.), *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD 2024, Barcelona, Spain, August 25-29, 2024*, pp. 502–513. ACM, 2024. doi: 10.1145/3637528.3671853. URL `https://doi.org/10.1145/3637528.3671853`.

Ying Cui, Mark Gierl, and Qi Guo. Statistical classification for cognitive diagnostic assessment: An artificial neural network approach. *Educational Psychology*, 36(6):1065–1082, 2016.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pp. 4171–4186, 2019.

Vijay Prakash Dwivedi and Xavier Bresson. A generalization of transformer networks to graphs. *CoRR*, abs/2012.09699, 2020. URL `https://arxiv.org/abs/2012.09699`.

Ying Fan, Olivia Watkins, Yuqing Du, Hao Liu, Moonkyung Ryu, Craig Boutilier, Pieter Abbeel, Mohammad Ghavamzadeh, Kangwook Lee, and Kimin Lee. Reinforcement learning for fine-tuning text-to-image diffusion models. In Alice Oh, Tristan Naumann, Amir Globerson, Kate Saenko, Moritz Hardt, and Sergey Levine (eds.), *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023*, 2023. URL `http://papers.nips.cc/paper_files/paper/2023/hash/fc65fab891d83433bd3c8d966edde311-Abstract-Conference.html`.

John H Flavell. *Cognitive development.* Prentice-Hall, 1977.

Weibo Gao, Qi Liu, Zhenya Huang, Yu Yin, Haoyang Bi, Mu-Chun Wang, Jianhui Ma, Shijin Wang, and Yu Su. RCD: relation map driven cognitive diagnosis for intelligent education systems. In Fernando Diaz, Chirag Shah, Torsten Suel, Pablo Castells, Rosie Jones, and Tetsuya Sakai (eds.), *SIGIR '21: The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Virtual Event, Canada, July 11-15, 2021*, pp. 501–510. ACM, 2021.

Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pp. 249–256. JMLR Workshop and Conference Proceedings, 2010.

Ivo Grondman, Lucian Busoniu, Gabriel AD Lopes, and Robert Babuska. A survey of actor-critic reinforcement learning: Standard and natural policy gradients. *IEEE Transactions on Systems, Man, and Cybernetics, part C (applications and reviews)*, 42(6):1291–1307, 2012.

Hengnian Gu, Guoqian Luo, Xiaoxiao Dong, Shulin Li, and Dongdai Zhou. Revisiting cognition in neural cognitive diagnosis. In Yizhou Sun, Flavio Chierichetti, Hady W. Lauw, Claudia Perlich, Wee Hyong Tok, and Andrew Tomkins (eds.), *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining, V.1, KDD 2025, Toronto, ON, Canada, August 3-7, 2025*, pp. 402–412. ACM, 2025. doi: 10.1145/3690624.3709319. URL `https://doi.org/10.1145/3690624.3709319`.

Hado Hasselt. Double q-learning. *Advances in neural information processing systems*, 23, 2010.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

Jaehyeong Jo, Seul Lee, and Sung Ju Hwang. Score-based generative modeling of graphs via the system of stochastic differential equations. In *International conference on machine learning*, pp. 10362–10383. PMLR, 2022.

Jaehyeong Jo, Dongki Kim, and Sung Ju Hwang. Graph generation with diffusion mixture. *arXiv preprint arXiv:2302.03596*, 2023.

Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. In Sanmi Koyejo, S. Mohamed, A. Agarwal, Danielle Belgrave, K. Cho, and A. Oh (eds.), *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022*, 2022. URL http://papers.nips.cc/paper_files/paper/2022/hash/a98846e9d9cc01cfb87eb694d946ce6b-Abstract-Conference.html.

Frank C Keil. *Concepts, kinds, and cognitive development*. mit Press, 1992.

Jacqueline Leighton and Mark Gierl. *Cognitive diagnostic assessment for education: Theory and applications*. Cambridge University Press, 2007.

Kurt Lewin. *Principles of topological psychology*. Read Books Ltd, 2013.

Jiatong Li, Qi Liu, Fei Wang, Jiayu Liu, Zhenya Huang, Fangzhou Yao, Linbo Zhu, and Yu Su. Towards the identifiability and explainability for personalized learner modeling: An inductive paradigm. In Tat-Seng Chua, Chong-Wah Ngo, Ravi Kumar, Hady W. Lauw, and Roy Ka-Wei Lee (eds.), *Proceedings of the ACM on Web Conference 2024, WWW 2024, Singapore, May 13-17, 2024*, pp. 3420–3431. ACM, 2024. doi: 10.1145/3589334.3645437. URL https://doi.org/10.1145/3589334.3645437.

Renjie Liao, Yujia Li, Yang Song, Shenlong Wang, William L. Hamilton, David Duvenaud, Raquel Urtasun, and Richard S. Zemel. Efficient graph generation with graph recurrent attention networks. In Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, pp. 4257–4267, 2019. URL https://proceedings.neurips.cc/paper/2019/hash/d0921d442ee91b896ad95059d13df618-Abstract.html.

Yu-Shih Lin, Yi-Chun Chang, Keng-Hou Liew, and Chih-Ping Chu. Effects of concept map extraction and a test-based diagnostic environment on learning achievement and learners' perceptions. *Br. J. Educ. Technol.*, 47(4):649–664, 2016a. doi: 10.1111/BJET.12250. URL https://doi.org/10.1111/bjet.12250.

Yu-Shih Lin, Yi-Chun Chang, Keng-Hou Liew, and Chih-Ping Chu. Effects of concept map extraction and a test-based diagnostic environment on learning achievement and learners' perceptions. *British Journal of Educational Technology*, 47(4):649–664, 2016b.

Gang Liu, Jiaxin Xu, Tengfei Luo, and Meng Jiang. Graph diffusion transformers for multi-conditional molecular generation. In Amir Globersons, Lester Mackey, Danielle Belgrave, Angela Fan, Ulrich Paquet, Jakub M. Tomczak, and Cheng Zhang (eds.), *Advances in Neural Information Processing Systems 38: Annual Conference on Neural Information Processing Systems 2024, NeurIPS 2024, Vancouver, BC, Canada, December 10 - 15, 2024*, 2024a. URL http://papers.nips.cc/paper_files/paper/2024/hash/0f6931a9e339a012a9909306d7c758b4-Abstract-Conference.html.

Luping Liu, Yi Ren, Zhijie Lin, and Zhou Zhao. Pseudo numerical methods for diffusion models on manifolds. *arXiv preprint arXiv:2202.09778*, 2022.

Qi Liu, Miltiadis Allamanis, Marc Brockschmidt, and Alexander L. Gaunt. Constrained graph variational autoencoders for molecule design. In Samy Bengio, Hanna M. Wallach, Hugo Larochelle, Kristen Grauman, Nicolò Cesa-Bianchi, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3-8, 2018, Montréal, Canada*, pp.

7806–7815, 2018. URL https://proceedings.neurips.cc/paper/2018/hash/b8a03c5c15fcfa8dae0b03351eb1742f-Abstract.html.

Qi Liu, Shiwei Tong, Chuanren Liu, Hongke Zhao, Enhong Chen, Haiping Ma, and Shijin Wang. Exploiting cognitive structure for adaptive learning. In Ankur Teredesai, Vipin Kumar, Ying Li, Rómer Rosales, Evimaria Terzi, and George Karypis (eds.), *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2019, Anchorage, AK, USA, August 4-8, 2019*, pp. 627–635. ACM, 2019. doi: 10.1145/3292500.3330922. URL https://doi.org/10.1145/3292500.3330922.

Qi Liu, Zhenya Huang, Yu Yin, Enhong Chen, Hui Xiong, Yu Su, and Guoping Hu. EKT: exercise-aware knowledge tracing for student performance prediction. *IEEE Trans. Knowl. Data Eng.*, 33(1):100–115, 2021. doi: 10.1109/TKDE.2019.2924374. URL https://doi.org/10.1109/TKDE.2019.2924374.

Shuo Liu, Junhao Shen, Hong Qian, and Aimin Zhou. Inductive cognitive diagnosis for fast student learning in web-based intelligent education systems. In Tat-Seng Chua, Chong-Wah Ngo, Ravi Kumar, Hady W. Lauw, and Roy Ka-Wei Lee (eds.), *Proceedings of the ACM on Web Conference 2024, WWW 2024, Singapore, May 13-17, 2024*, pp. 4260–4271. ACM, 2024b. doi: 10.1145/3589334.3645589. URL https://doi.org/10.1145/3589334.3645589.

Yijing Liu, Chao Du, Tianyu Pang, Chongxuan Li, Min Lin, and Wei Chen. Graph diffusion policy optimization. In Amir Globersons, Lester Mackey, Danielle Belgrave, Angela Fan, Ulrich Paquet, Jakub M. Tomczak, and Cheng Zhang (eds.), *Advances in Neural Information Processing Systems 38: Annual Conference on Neural Information Processing Systems 2024, NeurIPS 2024, Vancouver, BC, Canada, December 10 - 15, 2024*, 2024c. URL http://papers.nips.cc/paper_files/paper/2024/hash/124256ed80af5d4bf4c4de17b66c4298-Abstract-Conference.html.

Zitao Liu, Qiongqiong Liu, Jiahao Chen, Shuyan Huang, and Weiqi Luo. simplekt: a simple but tough-to-beat baseline for knowledge tracing. *arXiv preprint arXiv:2302.06881*, 2023.

Frederic M Lord and Melvin R Novick. *Statistical theories of mental test scores*. IAP, 2008.

Youzhi Luo, Keqiang Yan, and Shuiwang Ji. Graphdf: A discrete flow model for molecular graph generation. In *International Conference on Machine Learning*, pp. 7192–7203. PMLR, 2021.

Karolis Martinkus, Andreas Loukas, Nathanaël Perraudin, and Roger Wattenhofer. SPECTRE: spectral conditioning helps to overcome the expressivity limits of one-shot graph generators. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvári, Gang Niu, and Sivan Sabato (eds.), *International Conference on Machine Learning, ICML 2022, 17-23 July 2022, Baltimore, Maryland, USA*, volume 162 of *Proceedings of Machine Learning Research*, pp. 15159–15179. PMLR, 2022. URL https://proceedings.mlr.press/v162/martinkus22a.html.

Hiromi Nakagawa, Yusuke Iwasawa, and Yutaka Matsuo. Graph-based knowledge tracing: Modeling student proficiency using graph neural network. In Payam M. Barnaghi, Georg Gottlob, Yannis Manolopoulos, Theodoros Tzouramanis, and Athena Vakali (eds.), *2019 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2019, Thessaloniki, Greece, October 14-17, 2019*, pp. 156–163. ACM, 2019. doi: 10.1145/3350546.3352513. URL https://doi.org/10.1145/3350546.3352513.

Alexander Quinn Nichol and Prafulla Dhariwal. Improved denoising diffusion probabilistic models. In *International Conference on Machine Learning*, pp. 8162–8171. PMLR, 2021.

Chenhao Niu, Yang Song, Jiaming Song, Shengjia Zhao, Aditya Grover, and Stefano Ermon. Permutation invariant graph generation via score-based generative modeling. In *International conference on artificial intelligence and statistics*, pp. 4474–4484. PMLR, 2020.

Joseph D Novak and D Bob Gowin. *Learning how to learn*. cambridge University press, 1984.

Shalini Pandey and George Karypis. A self attentive model for knowledge tracing. In Michel C. Desmarais, Collin F. Lynch, Agathe Merceron, and Roger Nkambou (eds.), *Proceedings of the 12th International Conference on Educational Data Mining, EDM 2019, Montréal, Canada, July*

2-5, 2019. International Educational Data Mining Society (IEDMS), 2019. URL https://drive.google.com/file/d/18d_X6AXkPMhiHFQ2POarstVbX_7oMdFM.

John Piaget. The origins of intelligence in children. *International University*, 1952.

Chris Piech, Jonathan Bassen, Jonathan Huang, Surya Ganguli, Mehran Sahami, Leonidas J Guibas, and Jascha Sohl-Dickstein. Deep knowledge tracing. *Advances in neural information processing systems*, 28, 2015.

Junhao Shen, Hong Qian, Shuo Liu, Wei Zhang, Bo Jiang, and Aimin Zhou. Capturing homogeneous influence among students: Hypergraph cognitive diagnosis for intelligent education systems. In Ricardo Baeza-Yates and Francesco Bonchi (eds.), *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD 2024, Barcelona, Spain, August 25-29, 2024*, pp. 2628–2639. ACM, 2024. doi: 10.1145/3637528.3672002. URL https://doi.org/10.1145/3637528.3672002.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.

Yang Song, Prafulla Dhariwal, Mark Chen, and Ilya Sutskever. Consistency models. 2023.

Leslie P Steffe and Jerry Edward Gale. *Constructivism in education*. Psychology Press, 1995. URL https://emis.dsd.sztaki.hu/journals/ZDM/zdm982r2.pdf.

Jianwen Sun, Fenghua Yu, Qian Wan, Qing Li, Sannyuya Liu, and Xiaoxuan Shen. Interpretable knowledge tracing with multiscale state representation. In Tat-Seng Chua, Chong-Wah Ngo, Ravi Kumar, Hady W. Lauw, and Roy Ka-Wei Lee (eds.), *Proceedings of the ACM on Web Conference 2024, WWW 2024, Singapore, May 13-17, 2024*, pp. 3265–3276. ACM, 2024. doi: 10.1145/3589334.3645373. URL https://doi.org/10.1145/3589334.3645373.

Richard S Sutton, Andrew G Barto, et al. *Reinforcement learning: An introduction*, volume 1. MIT press Cambridge, 1998.

Richard S. Sutton, David A. McAllester, Satinder Singh, and Yishay Mansour. Policy gradient methods for reinforcement learning with function approximation. In Sara A. Solla, Todd K. Leen, and Klaus-Robert Müller (eds.), *Advances in Neural Information Processing Systems 12, [NIPS Conference, Denver, Colorado, USA, November 29 - December 4, 1999]*, pp. 1057–1063. The MIT Press, 1999. URL http://papers.nips.cc/paper/1713-policy-gradient-methods-for-reinforcement-learning-with-function-approximatio

Kikumi K Tatsuoka. *Cognitive assessment: An introduction to the rule space method*. Routledge, 2009.

Shiwei Tong, Qi Liu, Wei Huang, Zhenya Huang, Enhong Chen, Chuanren Liu, Haiping Ma, and Shijin Wang. Structure-based knowledge tracing: An influence propagation view. In Claudia Plant, Haixun Wang, Alfredo Cuzzocrea, Carlo Zaniolo, and Xindong Wu (eds.), *20th IEEE International Conference on Data Mining, ICDM 2020, Sorrento, Italy, November 17-20, 2020*, pp. 541–550. IEEE, 2020. doi: 10.1109/ICDM50108.2020.00063. URL https://doi.org/10.1109/ICDM50108.2020.00063.

Puja Trivedi, Ryan A. Rossi, David Arbour, Tong Yu, Franck Dernoncourt, Sungchul Kim, Nedim Lipka, Namyong Park, Nesreen K. Ahmed, and Danai Koutra. Editing partially observable networks via graph diffusion models. In *Forty-first International Conference on Machine Learning, ICML 2024, Vienna, Austria, July 21-27, 2024*. OpenReview.net, 2024. URL https://openreview.net/forum?id=2cEhQ4vtTf.

Clément Vignac, Igor Krawczuk, Antoine Siraudin, Bohan Wang, Volkan Cevher, and Pascal Frossard. Digress: Discrete denoising diffusion for graph generation. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net, 2023. URL https://openreview.net/forum?id=UaAD-Nu86WX.

Fei Wang, Qi Liu, Enhong Chen, Zhenya Huang, Yuying Chen, Yu Yin, Zai Huang, and Shijin Wang. Neural cognitive diagnosis for intelligent education systems. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pp. 6153–6161, 2020.

Fei Wang, Weibo Gao, Qi Liu, Jiatong Li, Guanhao Zhao, Zheng Zhang, Zhenya Huang, Mengxiao Zhu, Shijin Wang, Wei Tong, et al. A survey of models for cognitive diagnosis: New developments and future directions. *arXiv preprint arXiv:2407.05458*, 2024.

Bihan Xu, Zhenya Huang, Jiayu Liu, Shuanghong Shen, Qi Liu, Enhong Chen, Jinze Wu, and Shijin Wang. Learning behavior-oriented knowledge tracing. In Ambuj K. Singh, Yizhou Sun, Leman Akoglu, Dimitrios Gunopulos, Xifeng Yan, Ravi Kumar, Fatma Ozcan, and Jieping Ye (eds.), *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD 2023, Long Beach, CA, USA, August 6-10, 2023*, pp. 2789–2800. ACM, 2023. doi: 10.1145/3580305.3599407. URL `https://doi.org/10.1145/3580305.3599407`.

Shangshang Yang, Xiaoshan Yu, Ye Tian, Xueming Yan, Haiping Ma, and Xingyi Zhang. Evolutionary neural architecture search for transformer in knowledge tracing. In Alice Oh, Tristan Naumann, Amir Globerson, Kate Saenko, Moritz Hardt, and Sergey Levine (eds.), *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023*, 2023a. URL `http://papers.nips.cc/paper_files/paper/2023/hash/3e53d82a1113e3d240059a9195668edc-Abstract-Conference.html`.

Shangshang Yang, Cheng Zhen, Ye Tian, Haiping Ma, Yuanchao Liu, Panpan Zhang, and Xingyi Zhang. Evolutionary multi-objective neural architecture search for generalized cognitive diagnosis models. In *2023 5th International Conference on Data-driven Optimization of Complex Systems (DOCS)*, pp. 1–10. IEEE, 2023b.

Shangshang Yang, Mingyang Chen, Ziwen Wang, Xiaoshan Yu, Panpan Zhang, Haiping Ma, and Xingyi Zhang. Disengcd: A meta multigraph-assisted disentangled graph learning framework for cognitive diagnosis. In Amir Globersons, Lester Mackey, Danielle Belgrave, Angela Fan, Ulrich Paquet, Jakub M. Tomczak, and Cheng Zhang (eds.), *Advances in Neural Information Processing Systems 38: Annual Conference on Neural Information Processing Systems 2024, NeurIPS 2024, Vancouver, BC, Canada, December 10 - 15, 2024*, 2024. URL `http://papers.nips.cc/paper_files/paper/2024/hash/a69d7f3a1340d55c720e572742439eaf-Abstract-Conference.html`.

Tianwei Yin, Michaël Gharbi, Richard Zhang, Eli Shechtman, Fredo Durand, William T Freeman, and Taesung Park. One-step diffusion with distribution matching distillation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 6613–6623, 2024.

Jiani Zhang, Xingjian Shi, Irwin King, and Dit-Yan Yeung. Dynamic key-value memory networks for knowledge tracing. In *Proceedings of the 26th international conference on World Wide Web*, pp. 765–774, 2017.

Jiying Zhang, Zijing Liu, Yu Wang, Bin Feng, and Yu Li. Subgdiff: A subgraph diffusion model to improve molecular representation learning. In Amir Globersons, Lester Mackey, Danielle Belgrave, Angela Fan, Ulrich Paquet, Jakub M. Tomczak, and Cheng Zhang (eds.), *Advances in Neural Information Processing Systems 38: Annual Conference on Neural Information Processing Systems 2024, NeurIPS 2024, Vancouver, BC, Canada, December 10 - 15, 2024*, 2024. URL `http://papers.nips.cc/paper_files/paper/2024/hash/3477ca0ce484aa2fa42c1361ab601c25-Abstract-Conference.html`.

Mengchun Zhang, Maryam Qamar, Taegoo Kang, Yuna Jung, Chenshuang Zhang, Sung-Ho Bae, and Chaoning Zhang. A survey on graph diffusion models: Generative ai in science for molecule, protein and material. *arXiv preprint arXiv:2304.01565*, 2023a.

Yunfei Zhang, Chuan Qin, Dazhong Shen, Haiping Ma, Le Zhang, Xingyi Zhang, and Hengshu Zhu. Relicd: A reliable cognitive diagnosis framework with confidence awareness. In Guihai Chen, Latifur Khan, Xiaofeng Gao, Meikang Qiu, Witold Pedrycz, and Xindong Wu (eds.), *IEEE International Conference on Data Mining, ICDM 2023, Shanghai, China, December 1-4, 2023*, pp. 858–867. IEEE, 2023b. doi: 10.1109/ICDM58522.2023.00095. URL `https://doi.org/10.1109/ICDM58522.2023.00095`.

Jialin Zhao, Yuxiao Dong, Ming Ding, Evgeny Kharlamov, and Jie Tang. Adaptive diffusion in graph neural networks. In Marc'Aurelio Ranzato, Alina Beygelzimer, Yann N. Dauphin, Percy Liang, and Jennifer Wortman Vaughan (eds.), *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual*, pp. 23321–23333, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/c42af2fa7356818e0389593714f59b52-Abstract.html.

Hanqi Zhou, Robert Bamler, Charley M Wu, and Álvaro Tejero-Cantero. Predictive, scalable and interpretable knowledge tracing on structured domains. *arXiv preprint arXiv:2403.13179*, 2024.

Yuqiang Zhou, Qi Liu, Jinze Wu, Fei Wang, Zhenya Huang, Wei Tong, Hui Xiong, Enhong Chen, and Jianhui Ma. Modeling context-aware features for cognitive diagnosis in student learning. In Feida Zhu, Beng Chin Ooi, and Chunyan Miao (eds.), *KDD '21: The 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event, Singapore, August 14-18, 2021*, pp. 2420–2428. ACM, 2021. doi: 10.1145/3447548.3467264. URL https://doi.org/10.1145/3447548.3467264.

Zhenpeng Zhou, Steven Kearnes, Li Li, Richard N. Zare, and Patrick Riley. Optimization of molecules via deep reinforcement learning. *CoRR*, abs/1810.08678, 2018. URL http://arxiv.org/abs/1810.08678.

Cheng Zhu, Jiayi Zhu, Lijuan Zhang, Xi Wu, Shuqi Yang, Ping Liang, Honghan Chen, and Ying Tan. Edge-aware hard clustering graph pooling for brain imaging data. *CoRR*, abs/2308.11909, 2023. doi: 10.48550/ARXIV.2308.11909. URL https://doi.org/10.48550/arXiv.2308.11909.

A Additional Discussion of Related Works

As a central topic in educational measurement, modeling cognitive structures has long remained a challenging task. With the advancement of educational data mining techniques, recent progress in graph generation offers promising support. Accordingly, we review related works as follows: *cognitive structure modeling*, *graph diffusion probabilistic models*, and *optimization of DPMs*.

Cognitive Structure Modeling. The students' cognitive structures (Lewin, 2013; Piaget, 1952; Bruner, 2009; Ausubel, 1968) represent their internal knowledge system, an evolving graph whose nodes reflect their construction of concepts and whose edges capture their construction of inter-concept relations (Novak & Gowin, 1984; Steffe & Gale, 1995). Traditional psychometric approaches derive such structures from expert-defined rules, which limit personalization and accuracy (Lord & Novick, 2008; Tatsuoka, 2009; Lin et al., 2016b). Considering that cognitive structure is an inherent learning state, researchers have shifted to indirectly measuring it based on students' responses to test items, e.g., knowledge tracing (KT) and cognitive diagnosis (CD).

From the KT perspective (Piech et al., 2015; Choi et al., 2020; Zhang et al., 2017), cognitive structures are implicitly approximated via students' learning states (also termed hidden states or knowledge states) inferred from response logs. This includes theory-guided state models (Gu et al., 2025; Sun et al., 2024), mastery pattern classifiers (Briggs & Circi, 2017; Cui et al., 2016), and encoder–decoder architectures (Li et al., 2024; Liu et al., 2024b; Chen et al., 2023). Some KT methods enrich these states with static concept maps or heterogeneous interaction graphs (Liu et al., 2019; Nakagawa et al., 2019; Tong et al., 2020; Gao et al., 2021; Yang et al., 2024), yet they typically emphasize concept mastery without modeling the formation of inter-concept relations.

From the CD perspective (Leighton & Gierl, 2007; Cheng et al., 2019; Wang et al., 2020), models aim to identify fine-grained cognitive attributes or abilities underlying observed responses. While some approaches introduce additional features (Liu et al., 2021; Xu et al., 2023; Zhou et al., 2021), address data distribution issues (Cheng et al., 2025; Zhang et al., 2023b), or optimize network structures (Yang et al., 2023a;b), they also tend to focus on the correctness of individual concepts, overlooking the holistic evolution of cognitive structures.

A recent attempt (Chen et al., 2024) to model cognitive structure state still relies on a predefined concept graph and treats node and edge construction independently, failing to capture their coupled dynamics. To our knowledge, we are the first to explicitly formulate the task of cognitive structure generation and present a unified framework for its holistic modeling.

Graph Diffusion Probabilistic Models. Graph generation has long relied on traditional deep generative frameworks (e.g., auto-regressive models (Liao et al., 2019), VAEs (Liu et al., 2018), GANs (Martinkus et al., 2022), and normalizing flows (Luo et al., 2021)) to capture complex graph distributions. More recently, diffusion probabilistic models (DPMs) (Ho et al., 2020) have emerged as a powerful new trend for graph generation (Zhang et al., 2023a). Continuous-time graph DPMs (e.g., EDP-GNN (Niu et al., 2020), GDSS (Jo et al., 2022), DruM (Jo et al., 2023)) learn to denoise Gaussian-corrupted graph representations (Song et al., 2020) but can struggle to preserve graph sparsity. To address this, discrete diffusion methods like DiGress (Vignac et al., 2023) replace continuous noise with categorical transitions, achieving strong results on complex benchmarks. To our knowledge, we are the first to introduce a graph diffusion probabilistic model for CSG.

Optimization of DPMs. Reinforcement learning (RL) has been widely used to steer graph generators toward downstream objectives. Traditional methods (Sutton et al., 1999; Zhou et al., 2018) rely on custom environments and exhibit high computational costs. Diffusion models (DPMs) have been aligned to external rewards in vision: DPO (Fan et al., 2023) and DDPO (Black et al., 2024) treat the reverse diffusion as a Markov decision process and apply policy gradients to optimize black-box reward signals, and DPM alignment has been extended to graphs by GDPO (Liu et al., 2024c), which introduces an eager policy gradient. Thus, we propose a SOLO-based reward to optimize the CSDPM, which is effective for aligning with cognitive development levels.

B THE COMPLETE PROCEDURE OF PRETRAINING CSDPM

Algorithm 1: Pretraining CSDPM

Input: Simulated dataset $\tilde{\mathbb{G}}$, diffusion steps T, loss weight λ_{ve}
while *not converged* **do**
  Sample $(\mathcal{G}_0, X^{T'}) \sim \tilde{\mathbb{G}}$;
  `// Sample a simulated cognitive structure and its interaction sequence`
  Sample $t \sim \mathcal{U}[\![1, T]\!]$;
  Sample $\mathcal{G}_t \sim q(\mathcal{G}_t|\mathcal{G}_0)$;
  $z \leftarrow f(\mathcal{G}_t, t)$; `// Graph-theoretic features`
  $h \leftarrow \mathrm{emb}(X^{T'})$; `// Interaction-guidance features`
  $(\hat{p}^{\mathcal{V}}, \hat{p}^{\mathcal{E}}) \leftarrow \phi_\theta(\mathcal{G}_t, z, h)$; `// Denoising pass`
  optimizer.step $(\mathcal{L}_{CE}(\hat{p}^{\mathcal{V}}, \mathcal{V}_0) + \lambda_{ve}\mathcal{L}_{CE}(\hat{p}^{\mathcal{E}}, \mathcal{E}_0))$; `// Cross-entropy loss`

C THE COMPLETE PROCEDURE OF POLICY OPTIMIZATION

Algorithm 2: Optimizing CSDPM

Input: Pretrained CSDPM p_θ, diffusion steps T, reward function $r_{solo}(\cdot)$, learning rate η, number of trajectories $|\mathcal{D}|$, timestep samples $|\mathcal{T}|$, training steps N
Output: Optimized CSDPM p_θ
for $n = 1, \ldots, N$ **do**
  for $d = 1, \ldots, |\mathcal{D}|$ **do**
    Sample cognitive structure trajectory $\mathcal{G}_{0:T}^{(d)} \sim p_\theta(\mathcal{G}_{0:T})$;
    Compute reward $r_{solo}(\mathcal{G}_0^{(d)})$;
    Sample random timesteps subset $\mathcal{T}_d \subseteq [\![1, T]\!]$;
  `// Estimate reward statistics`
  $\bar{r} \leftarrow \frac{1}{|\mathcal{D}|}\sum_{d=1}^{|\mathcal{D}|} r_{solo}(\mathcal{G}_0^{(d)}), \quad \mathrm{std}[r] \leftarrow \sqrt{\frac{1}{|\mathcal{D}|-1}\sum_{d=1}^{|\mathcal{D}|}(r_{solo}(\mathcal{G}_0^{(d)}) - \bar{r})^2}$;
  `// Estimate eager policy gradient`
  $\nabla_\theta J_{\mathrm{RL}}(\theta) \leftarrow \frac{1}{|\mathcal{D}|}\sum_{d=1}^{|\mathcal{D}|}\frac{T}{|\mathcal{T}_d|}\sum_{t\in\mathcal{T}_d}\frac{r_{solo}(\mathcal{G}_0^{(d)})-\bar{r}}{\mathrm{std}[r]}\nabla_\theta \log p_\theta(\mathcal{G}_0^{(d)}|\mathcal{G}_t^{(d)})$;
  `// Update parameters`
  $\theta \leftarrow \theta + \eta \cdot \nabla_\theta J_{\mathrm{RL}}(\theta)$;

D STATISTICS OF ALL FOUR DATASETS.

Table 4: Statistics of all four datasets.

Datasets	Math1	Math2	FrcSub	NIPS34
# of students	4,209	3,911	536	4918
# of questions	20	20	20	948
# of knowledge concepts	11	16	8	57
# of interactions	72,359	78,221	10,720	1,399,470
# of interactions per student	17.19	20.00	20.00	284.56

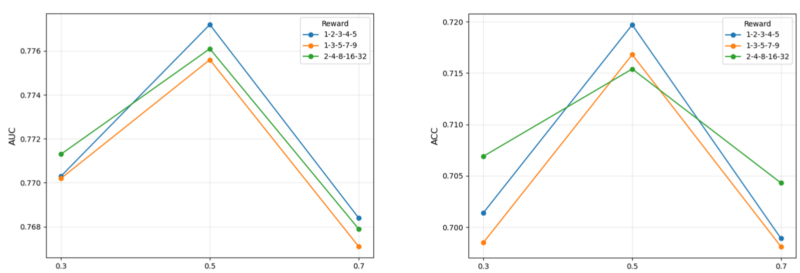
E IMPLEMENTATION DETAILS

For the parameterization of the CSDPM, we employ the extended Graph Transformer architecture from Dwivedi & Bresson (2020); Vignac et al. (2023), configuring it with 8 transformer layers, whose hidden dimensions (e.g., MLP, attention heads, and feed-forward layers) are set identically to those in Vignac et al. (2023). For pretraining the CSDPM, the CSDPM is trained using a uniform transition kernel for diffusion and the AdamW optimizer, with the number of diffusion steps T set as 500, node–edge loss balancing coefficient λ_{ve} $(0, 1)$, the batch size $(64, 512)$, dropout rate $(0, 0.5)$, and initial learning rate [1e-5,1e-2] with weight decay tuned via random or grid search strategy. The number of sampled trajectories $\mathcal{D}$ is searched in $\{128, 256, 512\}$. For CSG-KT and CSG-CD, the dimension of the graph pooling for cognitive state representation is searched in $\{8, 16, 32, 64\}$. To configure the training process, we initialize the parameters using Xavier initialization (Glorot & Bengio, 2010) and employ flexible methods such as random, grid, and bayes search& select strategies. For fairness, the hyper-parameter settings of the baseline models have been further tuned using the same tuning strategies to achieve optimal results. All experiments were run on Linux servers equipped with an Intel Xeon Platinum 8352V CPU and NVIDIA RTX 4090 GPUs.

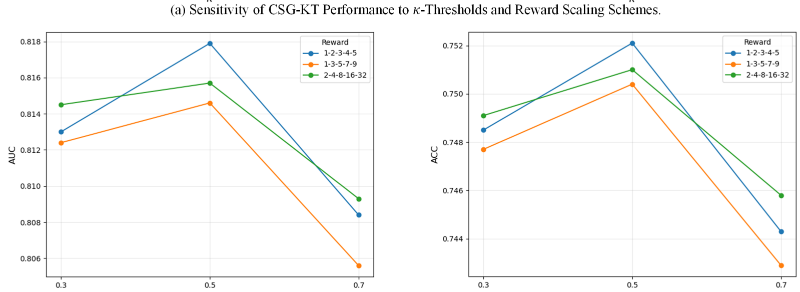
F Hyperparameters Analysis

We conducted a sensitivity analysis of some key parameters. We summarize the following observations and conclusions: The optimal node–edge loss balancing coefficient $\lambda_{ve} \in (0, 1)$ was 0.5 for Math1, Math2, and FrcSub, and 0.6 for NIPS34, which has a larger number of nodes yielding a correspondingly greater number of edges. For both CSG-KT and CSG-CD, the optimal graph pooling dimension was 16 for Math1 and Math2, 8 for FrcSub, and 32 for NIPS34.

To further examine the robustness of our reward design, we conducted an ablation study by systematically varying the threshold parameter κ of the matching degrees $\mathcal{M}_{\mathcal{V}}$ and $\mathcal{M}_{\mathcal{E}}$, as well as the reward scaling schemes. Specifically, we tested three settings of $\kappa \in \{0.3, 0.5, 0.7\}$, and three reward tuples: (i) a simple linear progression $(1, 2, 3, 4, 5)$, (ii) a steeper linear progression $(1, 3, 5, 7, 9)$, and (iii) an exponential progression $(2, 4, 8, 16, 32)$. Figure 2 summarizes the final AUC and ACC results, where we take the Math2 dataset as a representative example. The combination of $\kappa = 0.5$ with the simple linear reward $(1, 2, 3, 4, 5)$ consistently achieves the best balance between performance and stability. In contrast, exponential scaling tends to amplify the contribution of rare high-level cases, leading to unstable optimization, while the steeper linear scheme introduces uneven signals that bias the model toward intermediate levels. The neutral threshold $\kappa = 0.5$ also proved optimal: a looser setting ($\kappa = 0.3$) misclassifies partially aligned structures, whereas a stricter setting ($\kappa = 0.7$) over-penalizes mid-level structures. In practice, the selected reward tuple yields stable training behavior and consistent performance across datasets, and we apply the same values throughout all experiments without dataset-specific tuning.



(a) Sensitivity of CSG-KT Performance to κ-Thresholds and Reward Scaling Schemes.



(b) Sensitivity of CSG-CD Performance to κ-Thresholds and Reward Scaling Schemes.

Figure 2: **(Hyperparameter Study).** Sensitivity of CSG-KT and CSG-CD Performance (AUC and ACC) to κ-Thresholds and Reward Scaling Schemes.

G INFERENCE TIME ANALYSIS

We further report the inference time of CSG for generating a single cognitive structure graph. As shown in Table 5, the inference time remains low across datasets of different sizes, demonstrating the practical feasibility and efficiency of our CSG.

Table 5: Inference time for generating a single cognitive structure graph.

Dataset	Nodes	Inference Time (ms)
Math1	11	2.61
Math2	16	4.24
FrcSub	8	0.74
NIPS34	57	25.65

H SIMPLE EXAMPLE OF COGNITIVE STRUCTURE SIMULATION

Given five questions q_1–q_5 that assess the concepts *Sine Theorem* and *Cosine Theorem*, we make an idealized assumption: if a question involves only one concept, its weight for that concept is set to 1; if it involves both concepts, the weights for each concept are set to 0.5. Suppose a student s_i's responses to these five questions are recorded as X_i^5, as shown in the Table 6 below.

Table 6: Example of question weights and student responses.

Question	Sine Weight	Cosine Weight	Response
q_1	1.0	0.0	Correct
q_2	1.0	0.0	Correct
q_3	0.5	0.5	Correct
q_4	0.5	0.5	Incorrect
q_5	1.0	0.0	Incorrect

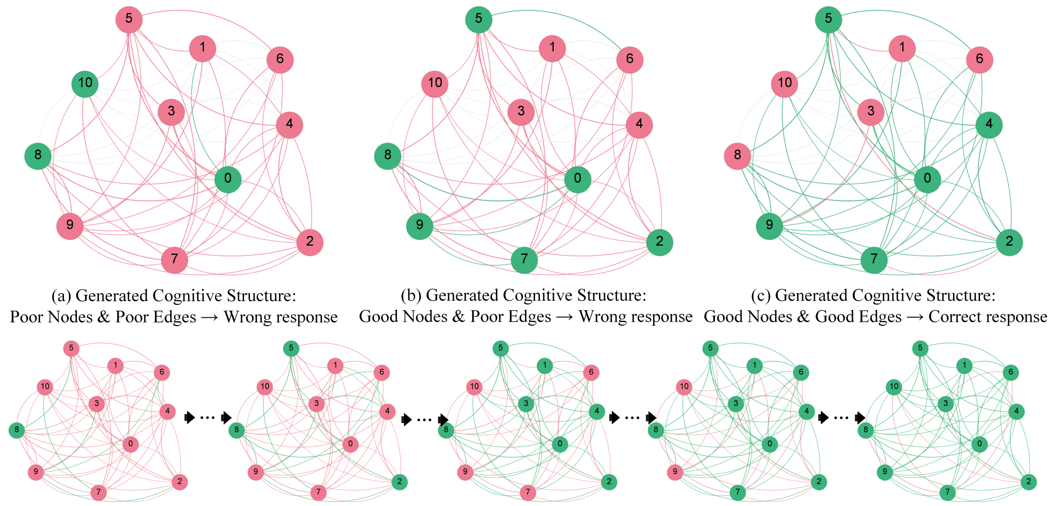
Accordingly, using Eqs.1 and 2, based on interaction records X_i^5, we can calculate the student's construction for the concepts *Sine Theorem* and *Cosine Theorem*, the node-level term $f_{UOC}(\textit{Sine Theorem}, X_i^5)$ and the edge-level term $f_{UOR}(\textit{Sine Theorem}, \textit{Cosine Theorem}, X_i^5)$ in the simulated cognitive structure, as follows:

$$f_{UOC}(\textit{Sine Theorem}, X_i^5) = \frac{1.0{\cdot}1.0+1.0{\cdot}1.0+0.5{\cdot}1.0+0.5{\cdot}0+1.0{\cdot}0}{1.0+1.0+0.5+0.5+1.0} = \frac{2.5}{4.0} = 0.625,$$

$$f_{UOR}(\textit{Sine Theorem}, \textit{Cosine Theorem}, X_i^5) = \frac{0{\cdot}(1.0+0){\cdot}0+0{\cdot}(1.0+0){\cdot}0+1.0{\cdot}(0.5+0.5){\cdot}1.0+1.0{\cdot}(0.5+0.5){\cdot}0+0{\cdot}(1.0+0){\cdot}0}{0{\cdot}(1.0+0)+0{\cdot}(1.0+0)+1.0{\cdot}(0.5+0.5)+1.0{\cdot}(0.5+0.5)+0{\cdot}(1.0+0)} = \frac{1.0}{2.0} = 0.5.$$

I VISUALIZATION AND INTERPRETABILITY ANALYSIS

As shown in Fig. 3, we observe the following: **(i)** Subfigure (a) shows the cognitive structure generated by CSG-CD for student s_5 immediately before answering question q_1 (assessing concepts $k_{0,2,5,7,9}$). The student exhibits weak construction of both individual concepts and their inter-concept relations, so CSG-CD predicts that the student will answer incorrectly. Subfigure (b) shows the structure for student s_{18} before the same question q_1; here the student has strong construction of all five concepts but still weak construction of their relations, and CSG-CD again predicts that the student will answer incorrectly. Subfigure (c) shows the structure for student s_{37} before q_1; in this case, the student demonstrates strong construction of both concepts and relations, so CSG-CD predicts a correct response. **(ii)** Subfigure (d) shows five representative cognitive structures generated by CSG-KT for student s_{15} at different points in their learning trajectory. Over time, s_{15}'s cognitive structure evolves from minimal construction to a fully developed structure that integrates the entire knowledge system in s_{15}'s mind, broadly aligning with the SOLO taxonomy levels of cognitive development. These case studies illustrate that CSG-generated structures not only capture students' subjective construction of the objective knowledge system but also trace its evolution throughout learning. The results are consistent with established findings in educational psychology, thereby providing meaningful explanations for students' response behaviors.



(a) Generated Cognitive Structure: Poor Nodes & Poor Edges → Wrong response

(b) Generated Cognitive Structure: Good Nodes & Poor Edges → Wrong response

(c) Generated Cognitive Structure: Good Nodes & Good Edges → Correct response

(d) Evolution of a student's generated cognitive structure over five time points.

Figure 3: **(Case).** Examples of generated cognitive structures and the evolution process. Each graph depicts a student's generated cognitive structure at a given timestamp. Nodes represent the student's construction of concepts, while edges represent their construction of inter-concept relations. Green indicates fully constructed elements, red indicates elements not yet constructed, and gray denotes low-frequency or irrelevant edges shown for clarity.

Table 7: List of knowledge concepts in Math1.

No.	Concept Name
0	Set
1	Inequality
2	Trigonometric function
3	Logarithm versus exponential
4	Plane vector
5	Property of function
6	Image of function
7	Spatial imagination
8	Abstract summarization
9	Reasoning and demonstration
10	Calculation

J LIMITATIONS

CSG leverages diffusion models, which are generally more computationally intensive than classical architectures used in knowledge tracing and cognitive diagnosis, such as LSTMs and GNNs. However, recent advances in accelerating the denoising process of diffusion models (Nichol & Dhariwal, 2021; Liu et al., 2022; Song et al., 2023; Yin et al., 2024) offer promising avenues to improve efficiency. Moreover, student cognitive structures typically do not require real-time updates, making the added computational cost acceptable in practical settings.

Besides, while we acknowledge that this rule-based simulation simplifies the inherently complex and nonlinear nature of cognitive development, it provides a crucial foundation for the first stage of our two-step training process. Inspired by LLM pretraining, we first learn initial representations using these rule-generated graphs, then refine them through reinforcement learning with a SOLO-taxonomy–based reward. This second stage allows the model to move beyond handcrafted assumptions and better capture authentic patterns of cognitive construction.

K LLMS USAGE

During the preparation of this paper, LLMs (specifically, ChatGPT) were used to assist in generating tables and figures and to support language polishing and proofreading.