CS6101 FINAL PROJECT REPORT
Probabilistic Pedagogy for Accelerated Learning

Martin Nguyen (Nguyen Quang Duc, A0333912J)

November 2025

1 Introduction

In traditional classrooms, the teaching–learning process unfolds through multi-round interactions between teachers and students (Fan et al., 2018; Rafferty et al., 2016). We consider the setting where a teacher aims to guide a student toward mastering a specific hypothesis by sequentially providing examples. Each example contributes differently to the student's understanding and may influence how subsequent examples are perceived. Thus, the careful selection and ordering of examples can significantly accelerate student learning.

Our initial objective is to design a probabilistic pedagogy model that captures this iterative teaching–learning process. Building on this model, we propose to evaluate the contribution of individual examples to the student's evolving knowledge state, thereby identifying the most informative example to present at each interaction round.

Recent shifts in education indicate that students are increasingly active, inquisitive, and capable of requesting clarifications (Jääskelä et al., 2020; Mameli et al., 2023). Prior research also suggests that students' questions or requests can serve as indicators of their understanding (Chin and Osborne, 2008). However, students do not always ask the "right" questions; in some cases, their inquiries may lead to further confusion. While asking the "right" question is an art, students may not have been trained to do so, leading them to produce both insightful and unhelpful questions, ranging from unnecessary or irrelevant to unfocused or aimless (Vale, 2013; Olmo, 1975). In such situations, teachers may need to override student requests to provide more pedagogically effective guidance. Building on this perspective, we extend the original model of pedagogical reasoning proposed by (Shafto et al., 2014) to incorporate both student agency and teacher override ability. In our model, after learning from provided examples, students may request the label of an unobserved example. The teacher then considers both the pedagogical value of the example and the student's request, ultimately deciding whether to comply with the request or override it by selecting an alternative example that, in the teacher's estimation, will better support the student's learning. We simulate this process and develop strategies for adaptive example selection, with the goal of accelerating student learning through principled probabilistic teaching.

2 Related Works

Pedagogical reasoning. Humans are continuously inferring the beliefs and intentions of others to reason about their actions in daily life. In the educational domain, this process is frequently observed in classrooms, where teachers attempt to understand what their students know and do not know in order to take appropriate teaching actions. Many studies have explored different strategies for modeling the interactions between students and teachers to (i) understand the reasons behind their actions and (ii) adapt those actions to achieve a goal (e.g., students fully mastering a concept) (Shafto et al., 2014, 2012; Fan et al., 2018). Recent studies have also introduced the notion of the *student agent*, emphasizing that students are more actively involved in the learning process (Fan et al., 2018; Rafferty et al., 2016; Jääskelä et al., 2020; Mameli et al., 2023). These works consider different scenarios: the teacher providing examples without student interaction (Shafto et al., 2014), the teacher providing an example and the student responding with some form of reward signal (Fan et al., 2018), or the teacher posing a question and the student answering (Rafferty et al., 2016). In contrast, our project focuses on the case where the teacher provides examples and the student must take actions regarding unseen examples, which differs from the settings above. Nevertheless, the underlying mechanism remains similar: the teacher maintains a model of the student and selects examples accordingly, while updating their belief based on the student's actions.

Learning acceleration. Accelerating the learning process of students has been a long-standing challenge in education (Carbonell, 1970; Silcock, 2003; Lee and Horsfall, 2010; Wang et al., 2023; Jiang, 2025). The main idea is to build a learner model and then perform inference or optimization on it (Jiang, 2025). From a computer science perspective, this is analogous to optimizing a model so that it achieves better performance in less time (i.e., with fewer learning data). These techniques can be categorized into two main concepts: adaptive learning and active learning. Adaptive learning involves providing the learner (or model) with appropriate data to facilitate learning, while active learning refers to the learner requesting particular data samples to learn from (Carbonell, 1970; Wang et al., 2023; Settles, 2009; Bonwell and Eison, 1991). In this project, we aim to incorporate both adaptive and active learning ideas into the probabilistic model, where the learner can request specific samples and the teacher can either accept or override the learner's request to provide more suitable examples, with the expectation that the learner will improve faster.

3 Proposed Method

3.1 Generative Model

We formalize the teaching–learning process as a probabilistic generative model. First, we define some notations that will be used in the model description.

- Hypothesis: $\theta^* \in \Theta$ — the true hypothesis or task parameter the student is expected to learn (e.g., a concept, weight vector).

- Student state: $S_t \in \mathcal{S}$ — the student's internal belief/posterior over θ at time t.
- Teacher state: $B_t \in \mathcal{B}$ — the teacher's belief about the student's belief at time t.
- Data sample: $x_t \in \mathcal{X}$ is an example input shown at time t (e.g., an input or problem instance); $y_t \in \mathcal{Y}$ is the corresponding output/label, determined by the task parameter. In general, the label for an example has a distribution conditioned on θ and the input ($y_t \sim p(y_t \mid \theta, x_t)$), but in this project, for simplicity, we assume that the label is deterministic, i.e. $y_t = f_\theta(x_t)$.
- Dataset: $D_t = \{(x_i, y_i)\}_{i=1}^t$ denotes the data observed up to time t. We assume a maximum of T examples.
- Student action: $a_t \in \mathcal{A}_t = \{x | x \in \mathcal{X} \wedge x \notin d, \forall d \in D_t\} \cup \{\text{passive}\}$ is the student's action at time t: either requesting a specific example from the remaining pool or passively waiting for the teacher's choice in the next round.

Original model This model is inspired by (Shafto et al., 2014). This model is effective for cases in which students passively receive examples from the teacher. This model captures the recursive interaction between teacher and student: the teacher selects examples based on their belief about the student's knowledge, while the student updates their belief based on provided examples.

$$
\begin{aligned}
S_t &\sim p(S_t \mid S_{t-1}, x_t, y_t) && \text{(student belief update)} \\
x_t &\sim p(x_t \mid S_t, \theta^*) && \text{(teacher selects example)} \\
y_t &= f_\theta(x_t) && \text{(label generation).}
\end{aligned}
$$

Refined model We extend the original model by incorporating the student actions into the teacher's belief about the student's current understanding. It allows students to take actions (asking for the next example or just waiting for the teacher) that influence subsequent rounds. The dynamics of the system are provided below.

$$
\begin{aligned}
S_t &\sim p(S_t \mid S_{t-1}, x_t, y_t) && \text{(student belief update)} \\
a_t &\sim p(a_t \mid S_t) && \text{(student action generation)} \\
B_t &\sim p(B_t \mid B_{t-1}, a_t) && \text{(teacher belief update)} \\
x_{t+1} &\sim p(x_{t+1} \mid B_t, \theta^*) && \text{(teacher selects example)} \\
y_{t+1} &\sim p(y_{t+1})|x_{t+1}, \theta) && \text{(label generation).}
\end{aligned}
$$

3.1.1 Teacher's Model of the Student

The teacher maintains a probabilistic model of the student's learning and querying behavior. Specifically, the teacher is assumed to know how the student updates her belief upon receiving a new example (i.e., the student's belief-update rule), but remains uncertain about how the student decides whether to request an example. This reflects a realistic setting in educational interactions, where students often lack precise knowledge of what constitutes the "right" question. Accordingly, we consider three candidate models of the student's action strategy: *random*, *uncertainty-reduction*, and *hypothesis-driven* strategies. The teacher also assumes that the student reasons about how examples are selected.

The teacher's initial priors are specified as:

$$\begin{aligned} p(S_0) &= \text{Dirichlet}(\Theta), \\ p(B_0) &= \text{Uniform}(\mathcal{S}), \\ p(a_0) &= p(\text{passive}) = 1. \end{aligned}$$

At each time step, after observing the student's query drawn from $p(a_t \mid S_t)$, the teacher updates her belief about the student's understanding in two stages:

- **Step 1: Updating the student's belief about the hypothesis.** We consider two types of students: (i) a **naive student**, who only updates her belief based on received examples, and (ii) a **rational student**, who recognizes that the teacher selects examples purposefully to facilitate learning. During this step, the distribution of the teacher's belief over possible student beliefs remains unchanged:

$$B_{t-1}(S_t \mid S_{t-1}, x_t, y_t) = B_{t-1}(S_{t-1}).$$

- **Step 2: Updating the teacher's belief about the student's belief.** Here, the student's belief over hypotheses is held fixed, and the teacher updates her belief about the student's internal state upon observing the student's action:

$$B_t(S_t) = B_{t-1}(S_t \mid S_{t-1}, x_t, y_t, a_t) \propto B_{t-1}(S_t \mid S_{t-1}, x_t, y_t) \cdot p(a_t \mid S_t).$$

For each candidate $x \in \mathcal{X} \setminus D_t$, the teacher computes the expected posterior probability that the student assigns to the true hypothesis θ^* after observing the new sample (x, y):

$$\begin{aligned} U(x; S_t, \theta^*) &= \mathbb{E}_{y \sim p(\cdot \mid x, \theta^*)} \left[S_{t+1}\big(\theta^* \mid S_t, x, y\big) \right], \\ p(x_{t+1} \mid B_t, \theta^*) &= \frac{\exp\big(\alpha\, U_{S_t \sim B_t}(x_{t+1}; S_t, \theta^*)\big)}{\sum\limits_{x' \in \mathcal{X} \setminus D_t} \exp\big(\alpha\, U_{S_t \sim B_t}(x'; S_t, \theta^*)\big)}, \end{aligned}$$

where α is a hyperparameter controlling the teacher's level of rationality.

3.1.2 Student's Model of the Teacher

We consider two types of students as described above: a *naive student* and a *rational student*.

Naive student. The naive student updates her belief solely based on the teacher's provided examples:

$$S_t(\theta) \propto S_{t-1}(\theta) \cdot p(y_t \mid x_t, \theta).$$

Rational student. The rational student maintains a model of the teacher and assumes that the teacher acts pedagogically, *i.e.,* selecting examples to optimize learning. This student, therefore, updates her belief as:

$$S_t(\theta) \propto S_{t-1}(\theta) \cdot p(x_t \mid B_{t-1}, \theta) \cdot p(y_t \mid x_t, \theta),$$

where B_{t-1} denotes the student's belief about the teacher's internal state, initialized as $p(B_0) = \text{Uniform}(\mathcal{S})$.

Both types of students choose actions according to a Boltzmann policy based on a utility function $U(a; S_t)$:

$$p(a_t \mid S_t) = \frac{\exp\big(\beta\, U(a_t; S_t)\big)}{\sum_{a \in \mathcal{A}_t} \exp\big(\beta\, U(a; S_t)\big)},$$

where β is a temperature parameter controlling stochasticity. We consider three possible utility functions:

- **Random strategy:** The student selects an unobserved example uniformly at random:

  $$U(a; S_t) = 1, \quad \forall a \in \mathcal{A}_t.$$

- **Hypothesis-driven strategy:** The student selects the example that maximizes the posterior probability of a particular hypothesis:

  $$U(a; S_t) = \max_{\theta}\ \mathbb{E}_{y \sim p(y|S_t, a)}\big[p(\theta \mid D_t \cup \{(a, y)\})\big].$$

- **Uncertainty-reduction strategy:** The student selects the example expected to minimize the posterior uncertainty about the correct hypothesis:

  $$U(a; S_t) = -\mathbb{H}_{\theta \sim \Theta}\left[\mathbb{E}_{y \sim p(y|S_t, a)}\big[p(\theta \mid D_t \cup \{(a, y)\})\big]\right],$$

  where $\mathbb{H}$ denotes the entropy operator.

3.2 Inference Process

Joint distribution (teacher's perspective). Conditioned on the true hypothesis θ^*, the teacher's joint distribution is defined as

$$p(S_{0:T}, a_{1:T}, B_{0:T}, x_{1:T}, y_{1:T} \mid \theta^*) = p(S_0)\, p(B_0) \prod_{t=1}^{T} \Big[p(S_t \mid S_{t-1}, x_t, y_t)\, p(a_t \mid S_t) \cdot p(B_t \mid B_{t-1}, a_t)\, p(x_t \mid B_{t-1}, \theta^*)\, p(y_t \mid x_t, \theta^*) \Big].$$

Joint Distribution from the Student's Perspective The student does not know the true hypothesis θ^* and begins with a uniform prior, $p(\theta)$, over the hypothesis space. The **naive student** does not model the teacher's pedagogical reasoning and assumes examples are not chosen strategically (e.g., sampled uniformly). The joint distribution over the student's beliefs $S_{0:T}$, actions $a_{1:T}$, observations $(x_{1:T}, y_{1:T})$, and the hypothesis θ is:

$$p(\theta, S_{0:T}, a_{1:T}, x_{1:T}, y_{1:T}) = p(\theta)p(S_0)\prod_{t=1}^{T}\Big[p(a_t \mid S_{t-1})\; p(x_t)\; p(y_t \mid x_t, \theta)\; p(S_t \mid S_{t-1}, x_t, y_t)\Big].$$

In contrast, the **rational student** actively models the teacher's intentions. This requires the student to maintain a belief B_t about the teacher's model of the student. This recursive reasoning is reflected in the joint distribution:

$$\begin{aligned} p(\theta, S_{0:T}, B_{0:T}, a_{1:T}, x_{1:T}, y_{1:T}) = p(\theta)p(S_0)p(B_0)\prod_{t=1}^{T}\Big[&p(a_t \mid S_{t-1})p(x_t \mid B_{t-1}, \theta)p(y_t \mid x_t, \theta) \\ &\cdot p(S_t \mid S_{t-1}, x_t, y_t)p(B_t \mid B_{t-1}, a_t)\Big]. \end{aligned}$$

4 Experiments

4.1 Clustering Environment

To evaluate the efficacy of our proposed model, we construct a synthetic clustering environment. The environment is defined within a d-dimensional Euclidean space. First, we generate a hypothesis space, Θ, which consists of H distinct hypotheses. Each hypothesis, θ_i, comprises a set of C clusters, each defined by a centroid c and a corresponding radius r:

$$\Theta = \{\theta_i\}_{i=1}^{H}, \quad \text{where} \quad \theta_i = \{(c_j, r_j)\}_{j=1}^{C}.$$

We randomly select a hypothesis in the list as the true hypothesis. Subsequently, a dataset of N points is sampled. We employ uniform sampling for data generation, in which points are drawn from a uniform distribution across the hypercube $(-C, C)^d$. Figure 1 illustrates an example of constructing this clustering environment.

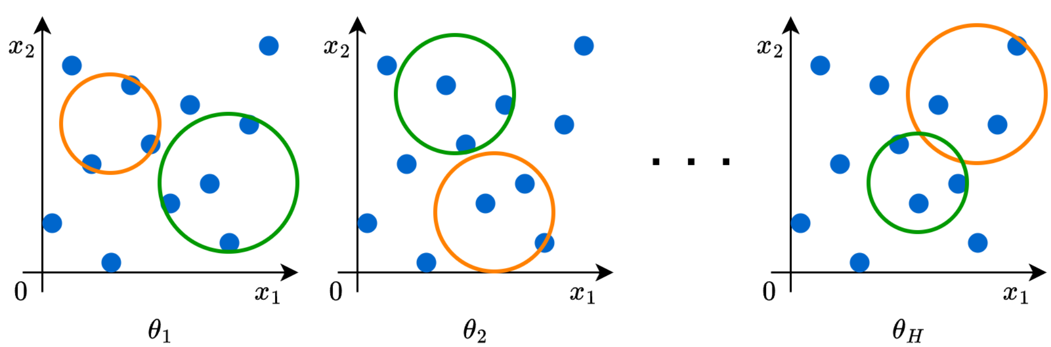

Figure 1: Example of the clustering environment in 2-dimensional space with 2 clusters.

To simulate reality, a point x can belong (be labelled) to a cluster or belong to no cluster. For any given point x and a hypothesis θ, the conditional probability of the point belonging to a specific cluster is proportional to a Gaussian function of the Euclidean distance between the point and the cluster's centroid. Specifically, the unnormalized probability for a cluster i defined as (c_i, r_i) is given by:

$$p(y = j|x, \theta) \propto \exp\left(-\frac{\|x - c_j\|^2}{2r_j^2}\right),$$

$$p(y = \texttt{no cluster}|x, \theta) \propto \begin{cases} 0 & \texttt{if } \exists (c_j, r_j) \in \theta : \|x - c_j\| \leq r_j \\ 1 & \texttt{otherwise} \end{cases}$$

The final label distribution is obtained by normalizing these values across all labels.

4.2 Learning Process Simulation

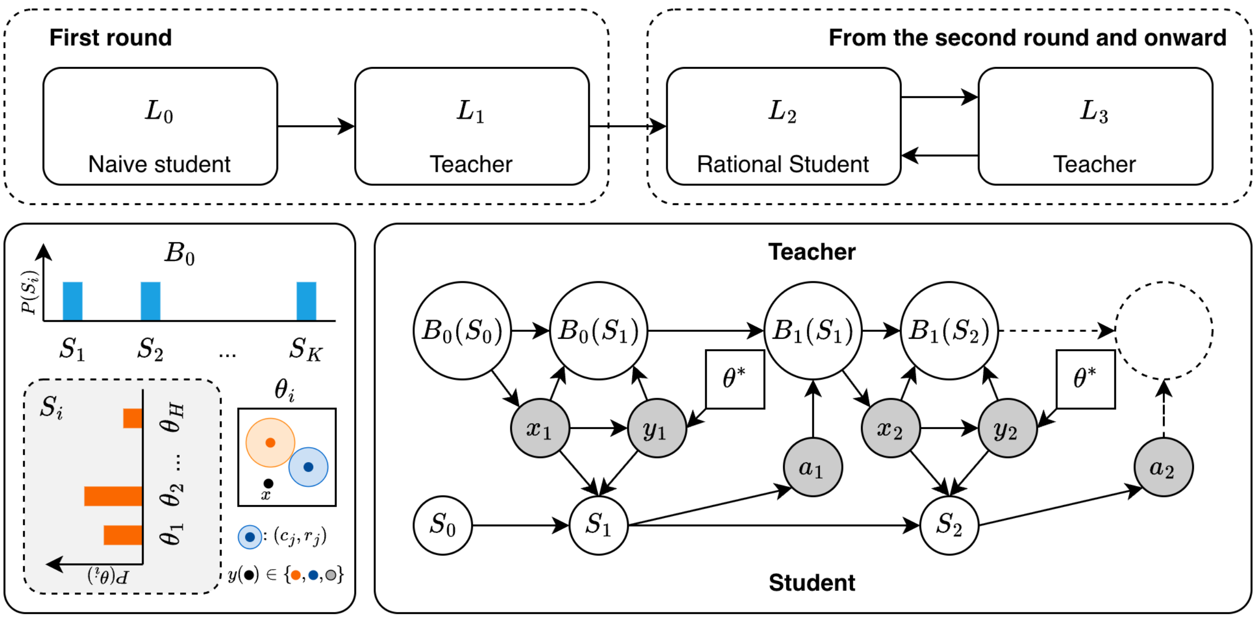

Figure 2: An overview of the learning simulation. **Top:** The level-k reasoning structure, alternating between a Naive Student (L_0), a pragmatic Teacher (L_1), a Rational Student (L_2), and a higher-level Teacher (L_3). **Bottom Left:** Insets showing the Teacher's prior belief B_0 and the internal structure of a student belief S. **Bottom Right:** The probabilistic graphical model of the inference process. The Teacher maintains a belief B_t over the Student's latent state S_t, provides a sample (x_t, y_t), and updates their belief based on the Student's subsequent action a_t.

The learning simulation starts with the construction of the student's and teacher's prior beliefs. The student's actual prior belief is defined as a Dirichlet distribution over the hypothesis space Θ. From the teacher's perspective, the space of possible student beliefs is infinite. Therefore, we approximate it with a finite set of K candidate student priors, denoted $B = \{S_1, \ldots, S_K\}$. Each student's belief S_k, which is a probability distribution over

Θ, is generated by sampling from an unbiased Dirichlet distribution. The teacher is assumed to initially hold a uniform prior over this set, considering each potential student's belief in B_0 to be equally likely. In the case where the student is rational (i.e., maintains a model of the teacher), we assume this model is perfectly aligned. Specifically, the set of candidate priors that the student assumes the teacher is using is identical to the teacher's actual set.

The simulation begins with an initial step where the L_0 student is passive, awaiting the first example from the teacher. Subsequently, the L_1 teacher actively selects examples designed to maximize the student's probability of identifying the correct hypothesis. This selection process is guided by the teacher's assumptions about the student's reasoning model (i.e., whether the student is naive or rational). The resulting chosen example, (x_{t+1}, y_{t+1}), is then presented to the student.

From the second round onward, the interaction becomes dynamic. The L_2 student may either actively request a specific example or remain passive, while the L_3 teacher is aware that the student can now be rational. The inference process is largely symmetric for both agents, with the crucial distinction lying between the student's true action-selection policy, $p_{\text{student}}(a_t|S_t)$, and the teacher's internal model of that policy, $p_{\text{teacher}}(a_t|S_t)$. The teacher ultimately provides the example deemed most informative, which may or may not coincide with the student's request. This interactive learning process continues until one of two termination conditions is met: (1) the student's belief in the true hypothesis, θ^*, exceeds a high-confidence threshold (e.g., $S_{\text{student}}(\theta^*) \geq 0.95$), or (2) the number of allowed examples exceeds a predefined number. We present an overview of the simulation process in Figure 2. Our code is available on Github[1].

4.3 Results and Discussion

We conducted experiments within a two-dimensional clustering environment consisting of two clusters and a total of 1,000 data samples. The number of hypotheses was fixed at 20. We approximate the teacher's belief B with $K = 100$. The primary evaluation metric across all experiments is learning efficiency, defined as the number of rounds (i.e., examples presented) required for the student's posterior belief in the true hypothesis θ^* to exceed a confidence threshold of 0.95, as well as the number of iterations required for the true hypothesis to achieve 1^{st} rank in the student's belief distribution. To ensure statistical robustness, each experimental configuration was simulated over 20 independent runs with different random seeds, and we report the mean and 95% confidence interval for each metric. The proposed method was evaluated through seven experimental settings as described below.

4.3.1 Experiment 1: Baseline and Core Model Efficacy

This experiment evaluates the effectiveness of our model by comparing it against simpler, non-strategic teaching baselines to demonstrate the importance of an adaptive teacher in accelerating learning. We hypothesize that a strategic teacher interacting with a student will achieve the fastest learning, outperforming a random, non-strategic teacher and also a lazy teacher who just provides labels for the student's queries. To isolate the impact of teacher

[1] `https://github.com/martinakaduc/teaching_simulation`

actions, we perform this experiment on a rational, hypothesis-driven student. The results are presented in Figure 3 and an example visualization is provided in Figure 4. We can observe that a teacher who strategically selects samples to help the student will help the student learn faster. Furthermore, when the student's queries are informative (*i.e.*, the student uses a hypothesis-driven strategy to choose actions), a lazy teacher who only provides labels for the student's queries performs slightly better than one who provides random examples.

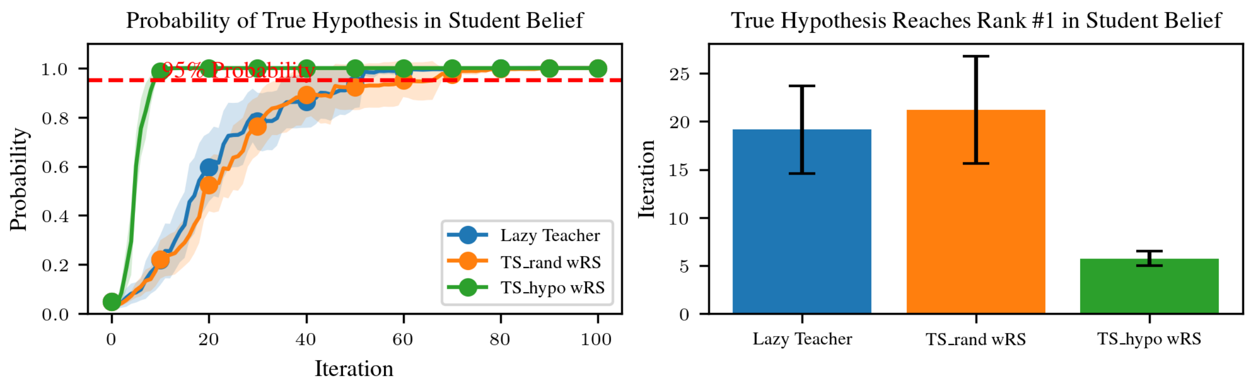

Figure 3: Comparison of two teaching strategies, random and hypothesis-driven, with a rational, hypothesis-driven student.

4.3.2 Experiment 2: Impact of Student Rationality

This experiment aims to determine whether a rational student learns more efficiently than a naive student. We hypothesize that the rational student will converge on the true hypothesis in fewer rounds, as reasoning about the teacher's strategy allows them to extract more information from each example and accelerate belief updates. To test this, we compare the performance of both student types under the assumption that the teacher's model of the student is accurate. Specifically, the teacher interacts with a hypothesis-driven student in this setup. As shown in Figure 5, the rational student consistently reaches 95% confidence in the correct hypothesis significantly faster than the naive student. We visualize the data space at the iteration when the student identified the correct hypothesis in Figure 6.

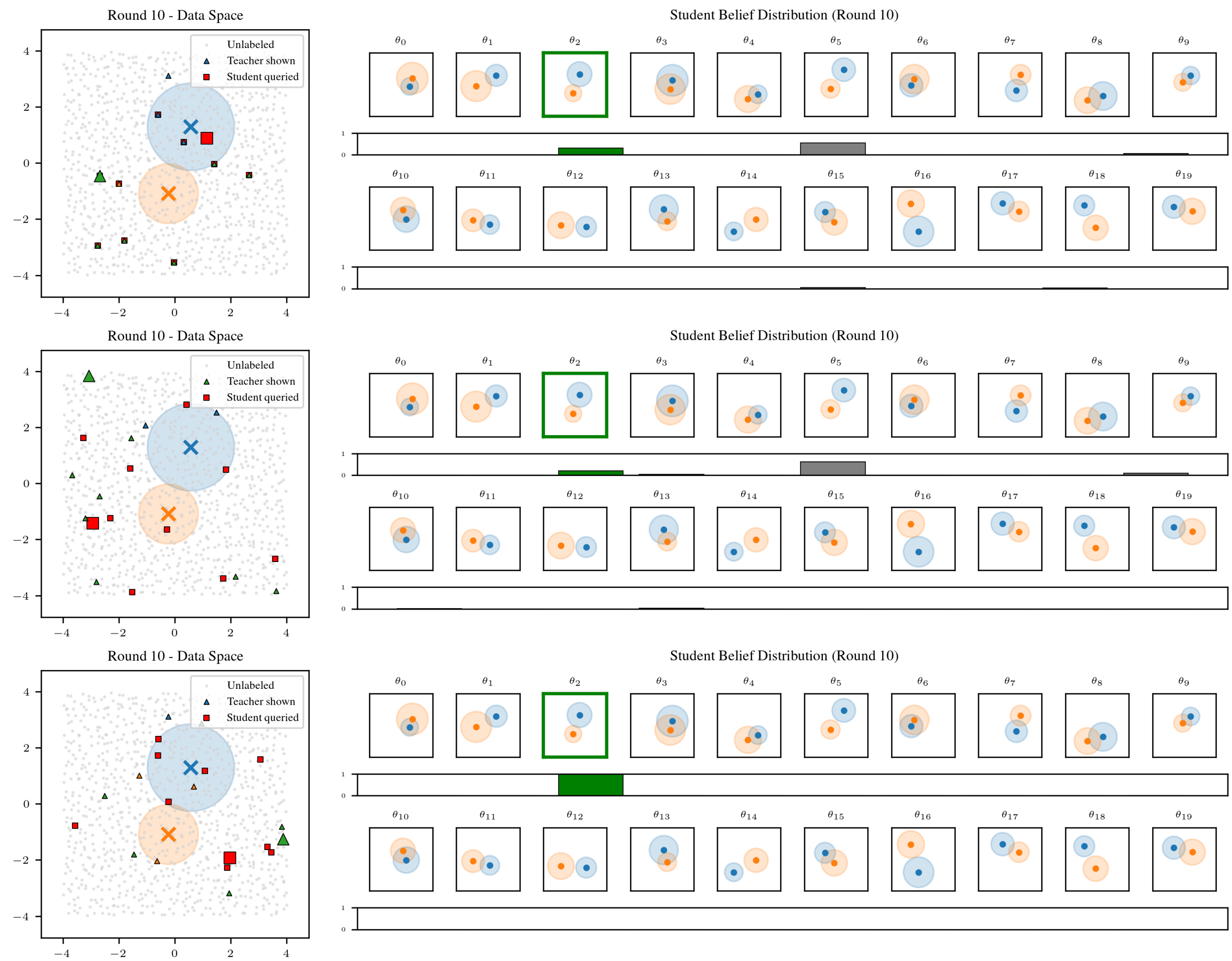

Figure 4: Visualization of the data space and the student's belief distribution. From top to bottom: Results of the lazy teacher, the random teacher, and the hypothesis-driven teacher at iteration #10. In the data space, triangles represent the teacher's examples (green indicates no cluster), and squares represent the student's queries. The hypothesis-driven teacher tends to choose boundary examples and informative negative examples.

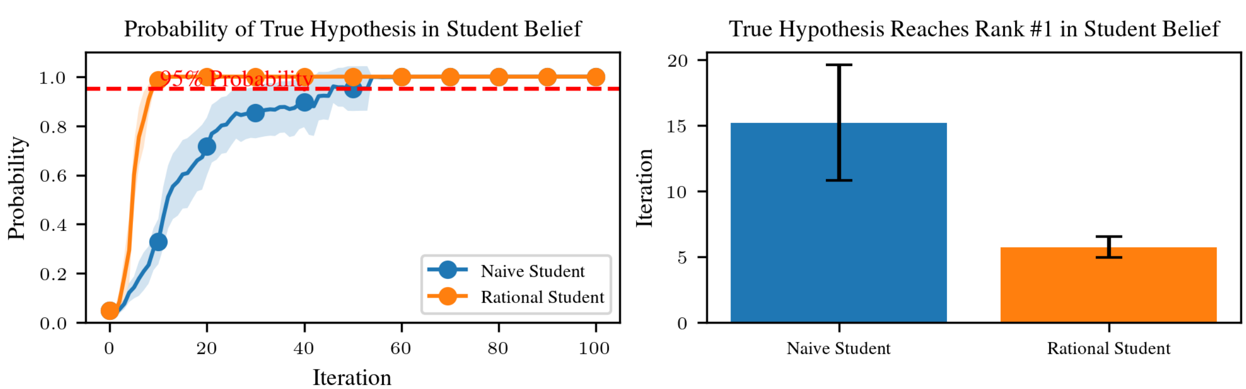

Figure 5: The effect of student type, naive and rational, on the learning process.

Figure 6: Visualization of the data space and the student's belief distribution. Top: Results of the naive student at iteration #50. Bottom: Results of the rational student at iteration #10. In the data space, triangles represent the teacher's examples (green indicates no cluster), and squares represent the student's queries. With naive students, the teacher tends to give more examples, especially negative ones.

4.3.3 Experiment 3: Evaluating Student Action-Selection Strategies

This experiment aims to identify which asking strategy most effectively enables the student to learn the correct hypothesis. Using the best-performing student model from Experiment 2 (assumed to be rational), we vary the student's actual query strategy while ensuring that the teacher's assumption about the strategy remains accurate. We also include a scenario in which the student is lazy and does not ask any questions.

The results in Figure 7 show that the hypothesis-driven strategy is the most effective for the student to cooperate with the active teacher, but the gaps between students' strategies are not large. Figure 8 is an illustration of data at the tenth iteration of all student strategies. We examined the reasons and found that this is because our setting involves an active teacher, which means the teacher can overwrite the student queries to provide more suitable learning examples. This allows the teacher's student beliefs (i.e., $S_i \in B$) to converge faster than the teacher's belief distribution B. Another reason is that we model the student's queries and the teacher's updates of the student's belief probabilities in a naive way, which means that neither the student nor the teacher models the other during this process. We support this finding by visualizing the top-3 students' beliefs in the teacher's belief in Figure 9. To further understand the importance of the active teacher, we perform the seventh experiment, which is described in the following section.

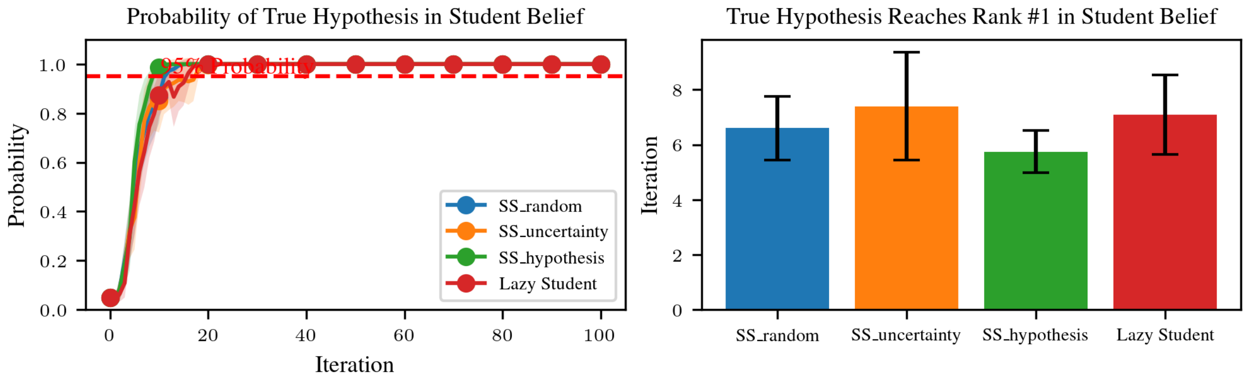

Figure 7: Comparison between students' action-selection strategy on active teacher.

4.3.4 Experiment 4: Robustness to Teacher's Misconceptions

This experiment conducts an ablation study to evaluate the model's robustness when the teacher's internal model of the student is incorrect. We hypothesize that a mismatch between the teacher's assumptions and the student's actual behavior will reduce learning efficiency. As shown in Figure 10, when the teacher mistakenly assumes that the student is naive or that the student's strategy is not hypothesis-driven, the learning process becomes less effective, resulting in longer learning time and greater uncertainty about the true hypothesis in the student's belief. We illustrate the data space at the tenth iteration for all settings in this experiment in Figure 11.

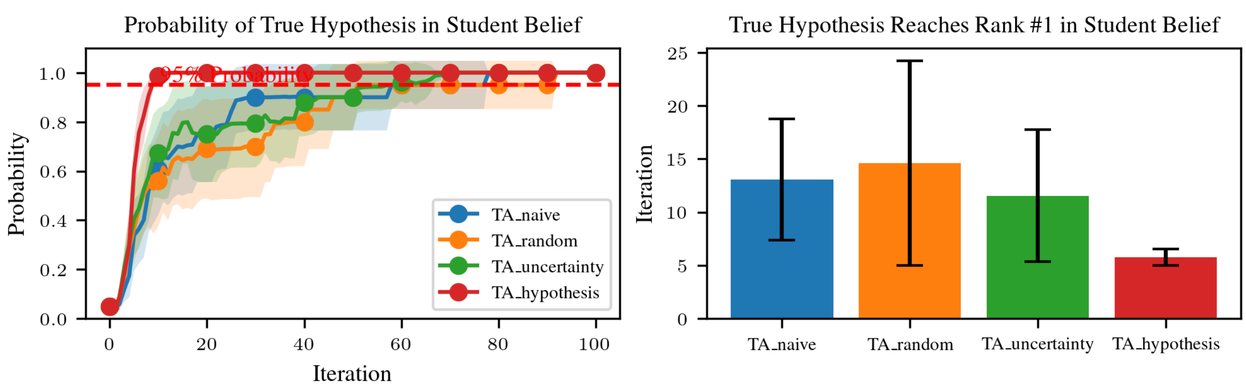

Figure 10: The effect of the teacher's assumption on the student learning process.

4.3.5 Experiment 5: Sensitivity Analysis

This experiment examines the model's sensitivity to key hyperparameters by varying the teacher's rationality parameter (α) and the student's rationality parameter (β) across three settings: 0.1, 1, and 10. The results (Figure 12) show that when both teacher and student

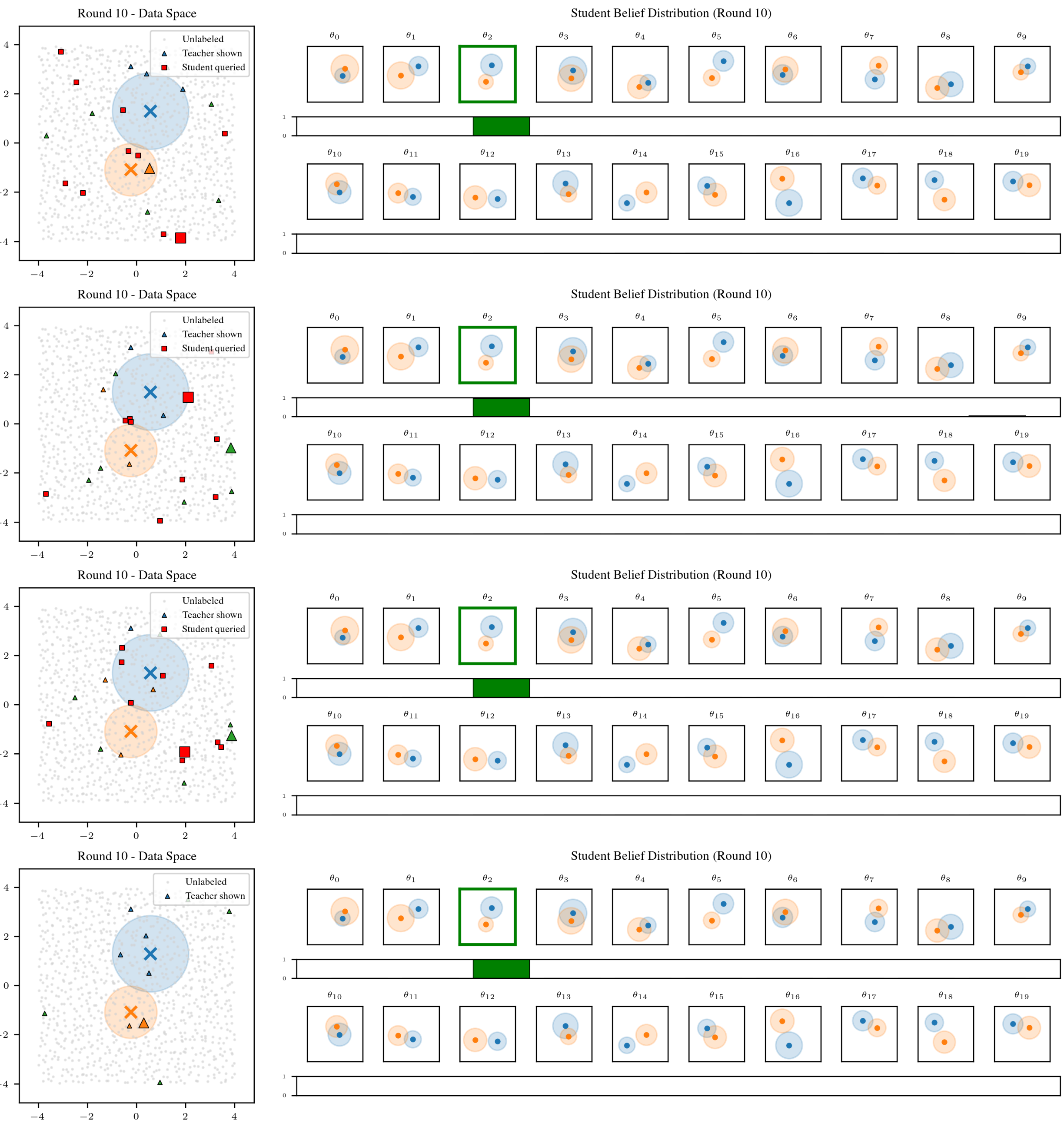

Figure 8: Visualization of the data space and the student's belief distribution. From top to bottom: results of students with random, uncertainty, and hypothesis-driven strategies at #10; and results of a lazy student. In the data space, triangles represent the teacher's examples (green indicates no cluster), and squares represent the student's queries. For active students, the teacher tends to prefer negative and boundary examples, whereas with a lazy student, the strategy leans more balance between positive and negative examples.

become more rational ($\alpha, \beta = 10$), the learning process become significantly faster because both teacher and student aare strongly believe in each other. Moreover, the illustration in

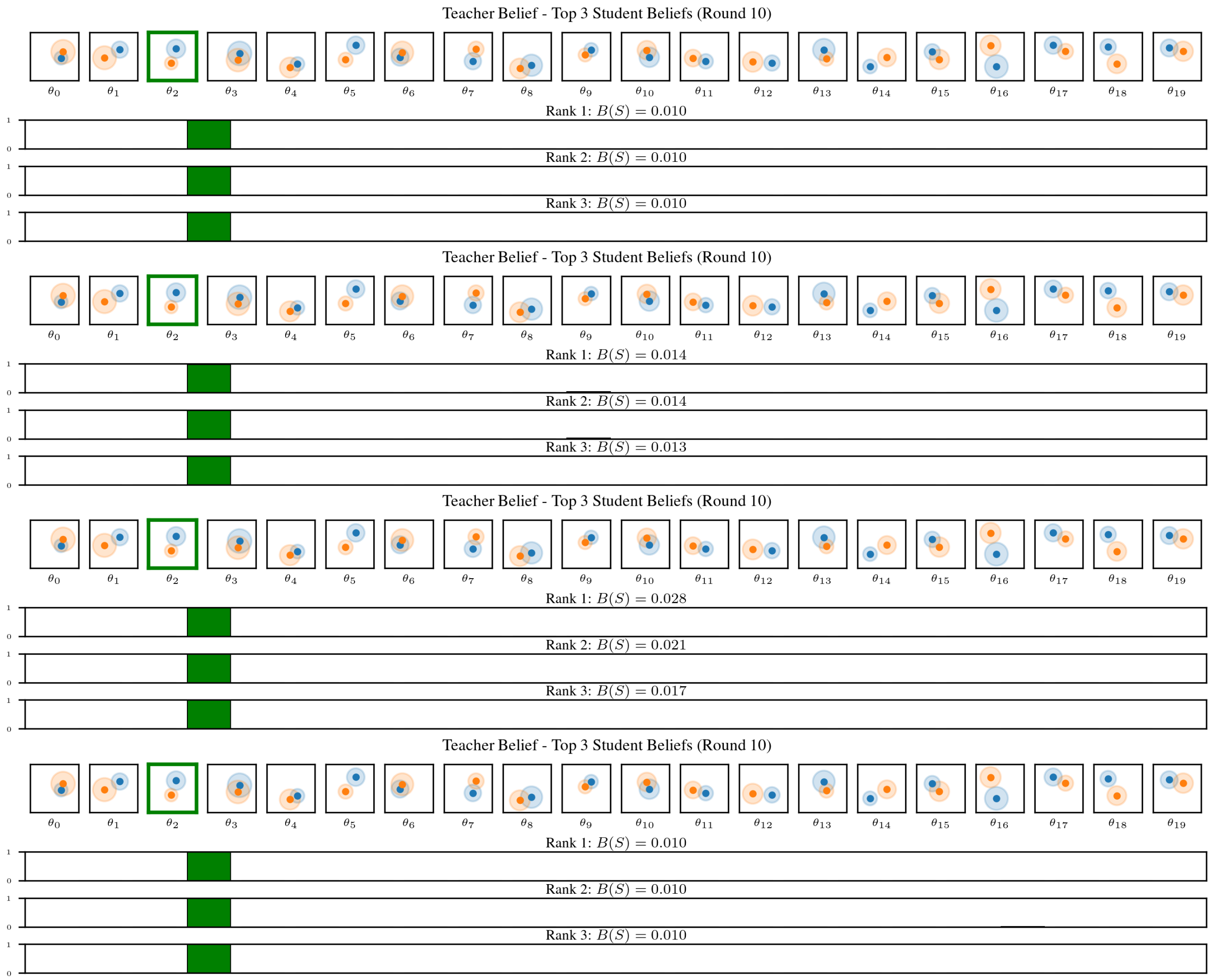

Figure 9: Visualization of the top-3 students' belief in the teacher's belief. From top to bottom: results of students with random, uncertainty, and hypothesis-driven strategies at #10; and results of a lazy student. We can see that the top-3 students' beliefs converged on the true hypothesis within 10 iterations. The probabilities of the students' beliefs under uncertainty and hypothesis-driven strategies were updated, but the convergence was slow.

Figure 13 shows that the student and teacher with $\alpha, \beta = 10$ only need two iteration rounds for the student to identify the true hypothesis correctly.

4.3.6 Experiment 6: Ablation Study on Teacher's Number of Beliefs

In the above experiments, the number of the teacher's beliefs is approximated by randomly generating a fixed number $K = 100$ of possible student beliefs. To examine the impact of this approximation on our method, we vary K across three settings: 10, 50, and 100. Figure 14 presents the results of this experiment. The findings indicate that there is no significant difference when varying this hyperparameter. With the results in this experiment

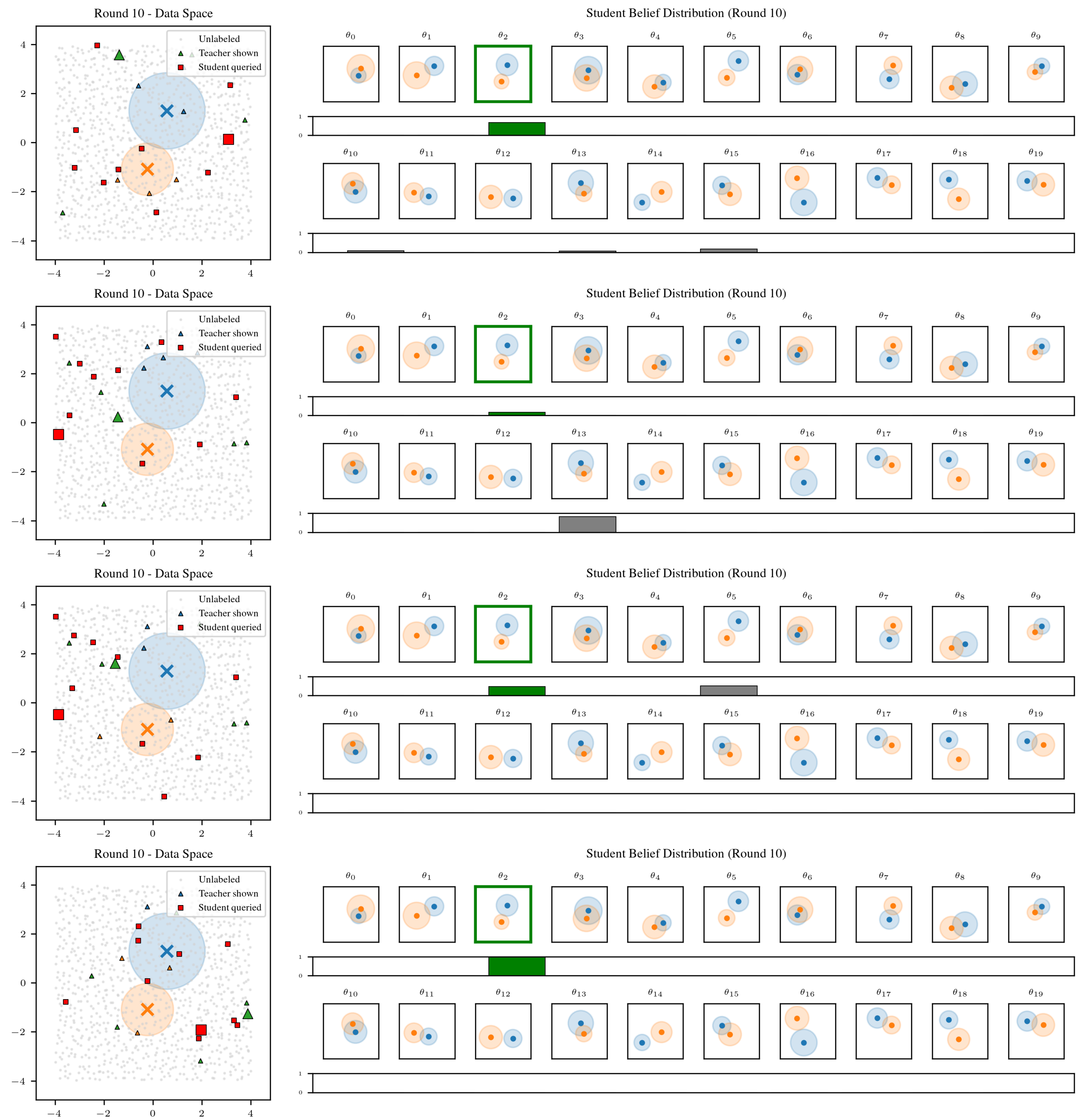

Figure 11: Visualization of the data space and the student's belief distribution. From top to bottom: results of incorrect teacher assumptions about naive, random, and uncertainty students at iteration #10. In the data space, triangles represent the teacher's examples (green indicates no cluster), and squares represent the student's queries. When the teacher has incorrect assumptions about the student's strategy, the provided examples may not be very helpful, leading to similar student beliefs and actions, which in turn result in similar teacher actions in subsequent steps.

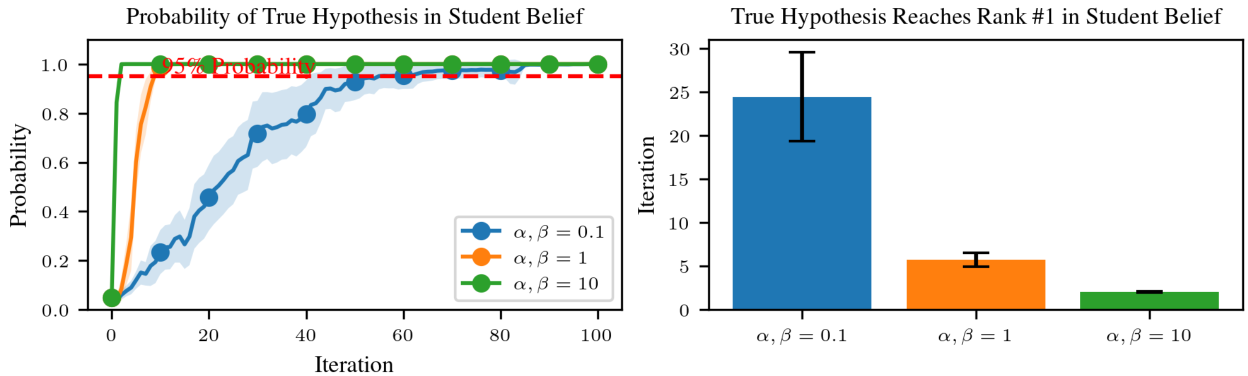

Figure 12: The effect of rationality in the learning process.

and the previous Experiment 3, an interesting observation is that a good teaching strategy (rational and having a correct model of the student to overwrite the student's queries) is more important than relying solely on the student's feedback to identify their state. The example runs in Figure 15 also confirm the above observation, as the teacher's provided examples appear similar regardless of the value of K.

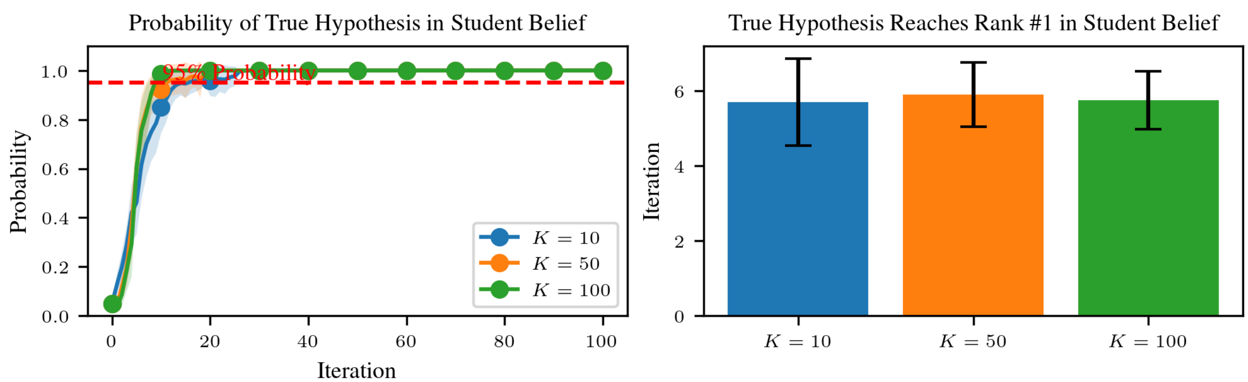

Figure 14: The effect of the number of teachers' beliefs.

4.3.7 Experiment 7: Effects of Student Strategies with Lazy Teacher

This experiment is similar to Experiment 3 but involves a lazy teacher instead of an active one. A lazy teacher only actively selects samples for the student when the student does not ask any questions; otherwise, she simply provides labels for the examples the student requests. The motivation for this experiment is to understand the effect of student strategies on learning under a lazy teacher and to further confirm the importance of an active, rational teacher. Figure 16 shows that without an active teacher, the student requires more time to learn and identify the correct hypothesis, with the random strategy being the least effective. The difference between different students' strategies is not too much for the same reason as

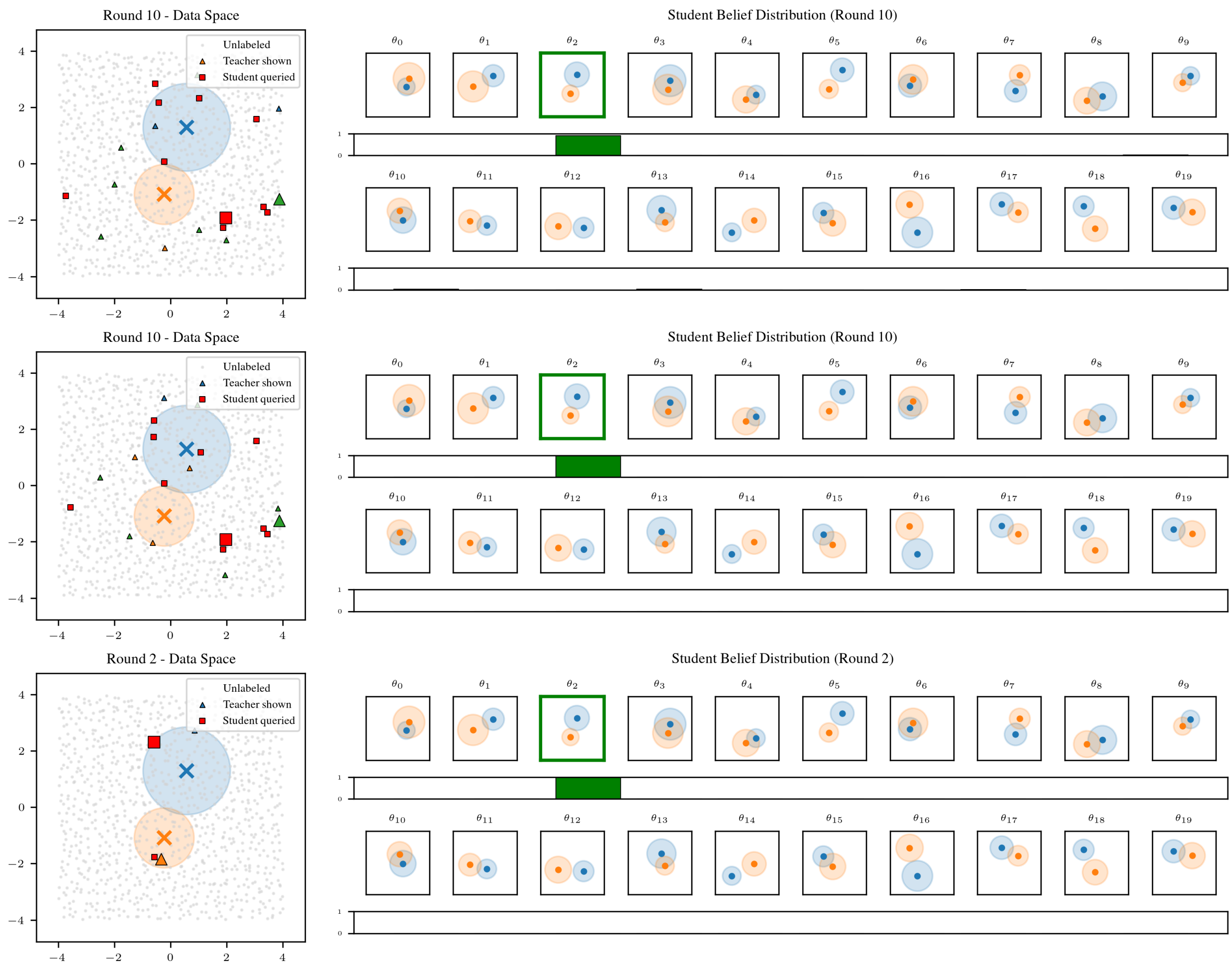

Figure 13: Visualization of the data space and the student's belief distribution. From top to bottom: results of different rational levels, including 0.1, 1, and 10. In the data space, triangles represent the teacher's examples (green indicates no cluster), and squares represent the student's queries. With a high rational level, the student can learn the correct hypothesis within just two boundary positive examples.

in Experiment 3. We visualize example runs at 20^{th} iteration in Figure 17 to demonstrate the acting strategy of the student with a lazy teacher.

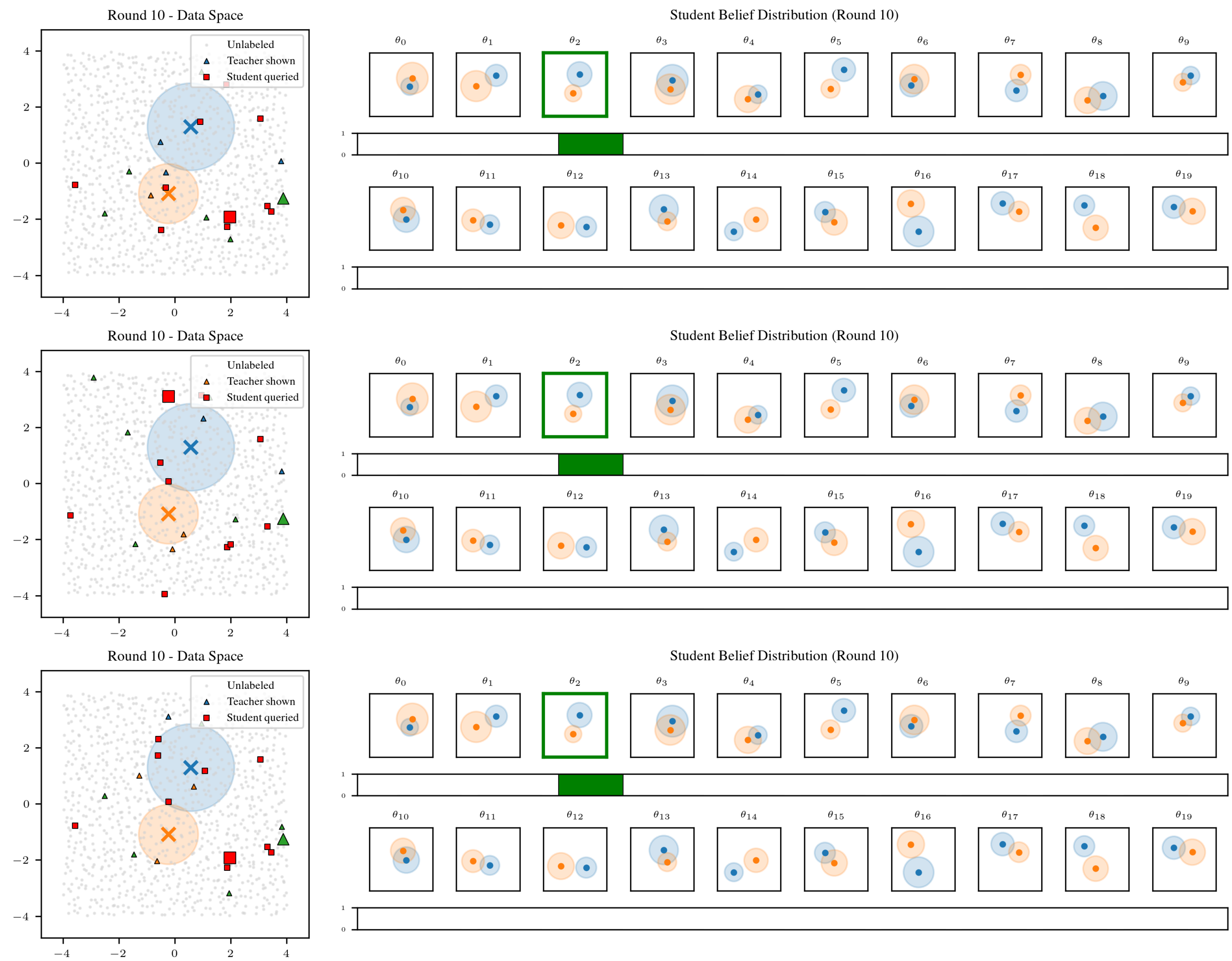

Figure 15: Visualization of the data space and the student's belief distribution. From top to bottom: results for different numbers of students' beliefs in the teacher's belief. In the data space, triangles represent the teacher's examples (green indicates no cluster), and squares represent the student's queries. The teacher appears to follow a similar strategy in this experiment, even though the values of K are different.

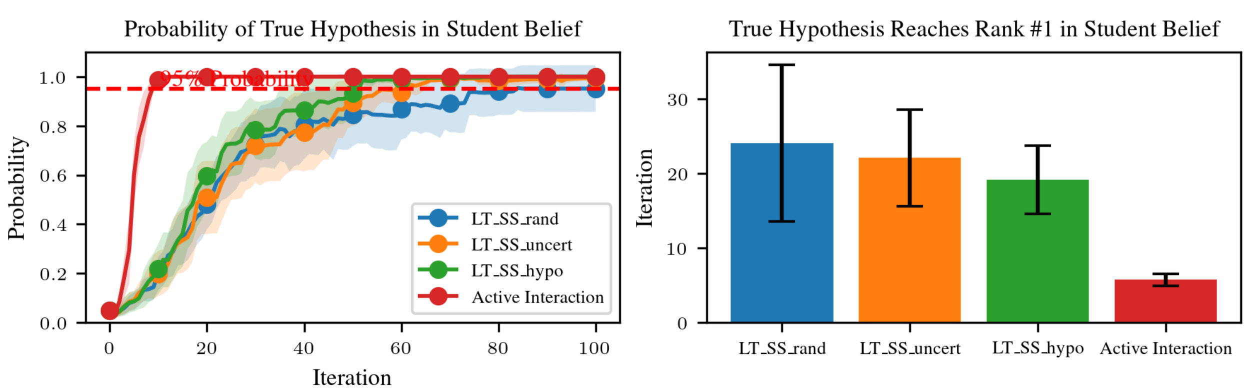

Figure 16: The effect of the number of teachers' beliefs.

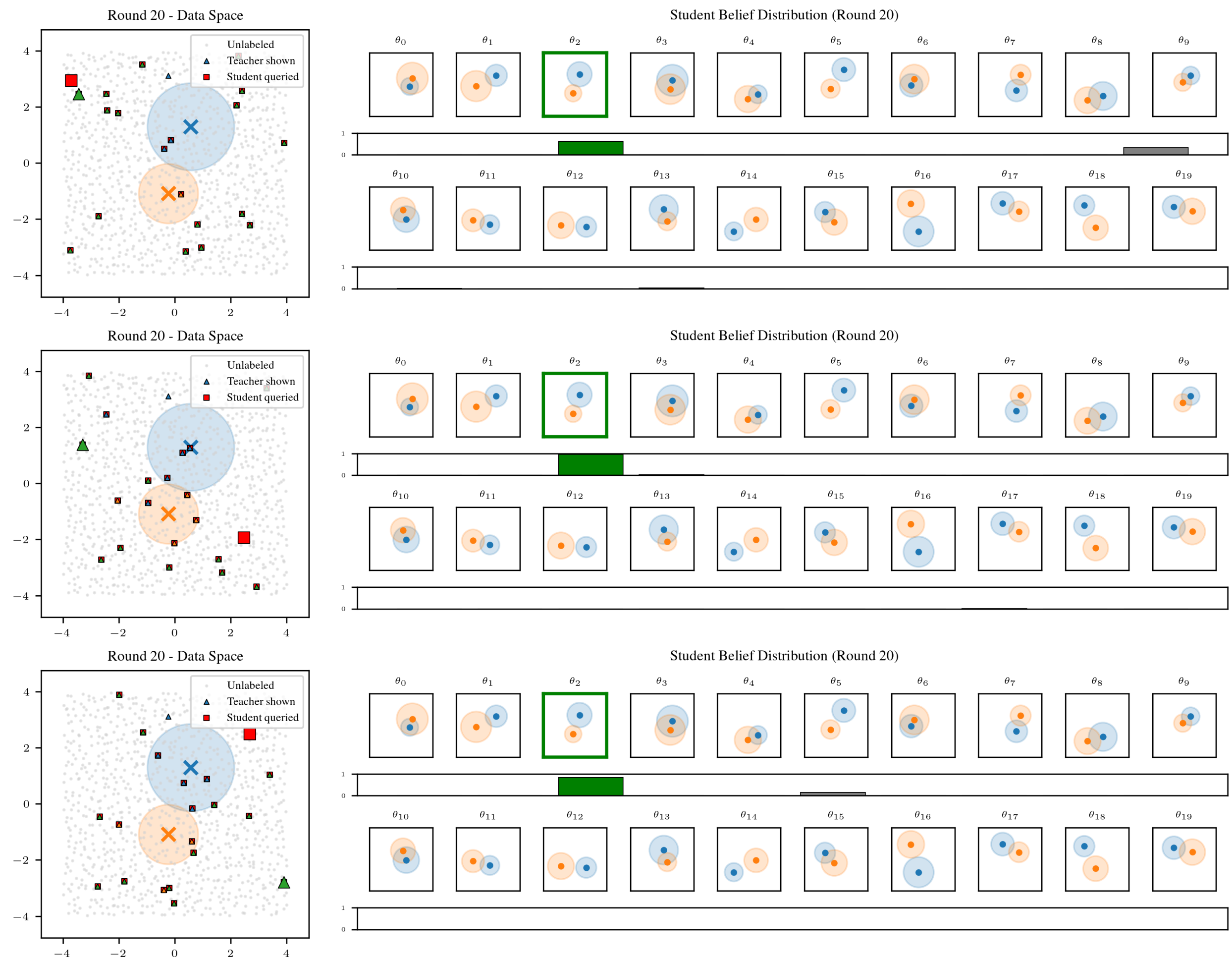

Figure 17: Visualization of the data space and the student's belief distribution. From top to bottom: results of students with random, uncertainty, and hypothesis-driven strategies at iteration #20. In the data space, triangles represent the teacher's examples (green indicates no cluster), and squares represent the student's queries. The uncertainty strategy tends to select samples near cluster boundaries, while the hypothesis-driven strategy maintains a more balanced selection of positive and negative examples, focusing on those most informative for learning.

5 Conclusion

In this project, we proposed and evaluated a probabilistic pedagogy model that extends traditional frameworks by incorporating both active student and active teacher. Our goal was to simulate a more realistic, interactive learning process to identify strategies that better accelerate student learning. Through a series of experiments in a synthetic clustering environment, we demonstrated that the nature of the teacher-student interaction is a critical determinant of learning efficiency. Our key findings include:

- **Student rationality matters.** A rational student, who models the teacher's pedagogical intent, learns significantly faster than a naive student who only updates beliefs based on data (Experiment 2).
- **Strategic interaction is key.** The most effective learning occurs when a strategic, rational teacher interacts with a student who also adopts an optimal query strategy (*i.e.,* hypothesis-driven, as shown in Experiment 3). Mismatches in the teacher's assumptions about the student's strategy can degrade performance (Experiment 4).
- **Active, strategic teaching is paramount.** Our most significant finding is the power of the active teacher. A "lazy" teacher who passively fulfills student requests is far less effective than an active teacher who strategically selects examples (Experiment 7). This ability to override a student's query is so impactful that it can compensate for a less precise model of the student's belief state (Experiment 6), underscoring that a robust pedagogical strategy is more important than perfectly inferring the student's intent at every step.

While this work was limited to a simulated environment, it provides strong quantitative support for a pedagogical model where the teacher acts as an active, rational guider rather than a passive information provider. Future work could explore these dynamics in more complex domains (such as training AI models) or with human-in-the-loop experiments. In summary, our model offers a compelling framework for developing more effective, accelerated, and adaptive learning technologies by balancing student agency with principled pedagogical override.

AI Tool Declaration

I used ChatGPT to revise the writing and correct grammar mistakes. I am responsible for the content and quality of the submitted work.

References

C. C. Bonwell and J. A. Eison. Active Learning: Creating Excitement in the Classroom. 1991 ASHE-ERIC Higher Education Reports. Technical report, ERIC Clearinghouse on Higher Education, The George Washington University, One Dupont Circle, Suite 630, Washington, DC 20036-1183 ($17, 1991. URL `https://eric.ed.gov/?id=ED336049`. ISBN: 9781878380081 ISSN: 0884-0040 ERIC Number: ED336049.

J. R. Carbonell. AI in CAI: An Artificial-Intelligence Approach to Computer-Assisted Instruction. *IEEE Transactions on Man-Machine Systems*, 11(4):190–202, Oct. 1970. ISSN 2168-2860. doi: 10.1109/TMMS.1970.299942. URL `https://ieeexplore.ieee.org/document/4081977`.

C. Chin and J. Osborne. Students' questions: a potential resource for teaching and learning science. *Studies in Science Education*, 44(1):1–39, Mar. 2008. ISSN 0305-7267.

doi: 10.1080/03057260701828101. URL https://doi.org/10.1080/03057260701828101. Publisher: Routledge _eprint: https://doi.org/10.1080/03057260701828101.

Y. Fan, F. Tian, T. Qin, X.-Y. Li, and T.-Y. Liu. Learning to Teach. In *Proceedings of The Sixth International Conference on Learning Representations*, Feb. 2018. URL https://openreview.net/forum?id=HJewuJWCZ.

B. Jiang. *Toward Trustworthy Adaptive Learning: Explainable Learner Models*. Routledge, London, Mar. 2025. ISBN 978-1-003-58517-6. doi: 10.4324/9781003585176.

P. Jääskelä, A.-M. Poikkeus, P. Häkkinen, K. Vasalampi, H. Rasku-Puttonen, and A. Tolvanen. Students' agency profiles in relation to student-perceived teaching practices in university courses. *International Journal of Educational Research*, 103:101604, Jan. 2020. ISSN 0883-0355. doi: 10.1016/j.ijer.2020.101604. URL https://www.sciencedirect.com/science/article/pii/S088303551932748X.

N. Lee and B. Horsfall. Accelerated Learning: A Study of Faculty and Student Experiences. *Innovative Higher Education*, 35(3):191–202, June 2010. ISSN 1573-1758. doi: 10.1007/s10755-010-9141-0. URL https://link.springer.com/article/10.1007/s10755-010-9141-0. Company: Springer Distributor: Springer Institution: Springer Label: Springer Publisher: Springer Netherlands.

C. Mameli, V. Grazia, and L. Molinari. Student agency: Theoretical elaborations and implications for research and practice. *International Journal of Educational Research*, 122:102258, Jan. 2023. ISSN 0883-0355. doi: 10.1016/j.ijer.2023.102258. URL https://www.sciencedirect.com/science/article/pii/S0883035523001222.

B. G. Olmo. Teaching Students to Ask Questions. *Language Arts*, 52(8):1116–1119, 1975. ISSN 0360-9170. URL https://www.jstor.org/stable/41962351. Publisher: National Council of Teachers of English.

A. N. Rafferty, E. Brunskill, T. L. Griffiths, and P. Shafto. Faster Teaching via POMDP Planning. *Cognitive Science*, 40(6):1290–1332, Aug. 2016. ISSN 0364-0213, 1551-6709. doi: 10.1111/cogs.12290. URL https://onlinelibrary.wiley.com/doi/10.1111/cogs.12290.

B. Settles. Active Learning Literature Survey. Technical Report, University of Wisconsin-Madison Department of Computer Sciences, 2009. URL https://minds.wisconsin.edu/handle/1793/60660. Accepted: 2012-03-15T17:23:56Z.

P. Shafto, N. D. Goodman, and M. C. Frank. Learning From Others: The Consequences of Psychological Reasoning for Human Learning. *Perspectives on Psychological Science*, 7 (4):341–351, July 2012. ISSN 1745-6916. doi: 10.1177/1745691612448481. URL https://doi.org/10.1177/1745691612448481. Publisher: SAGE Publications Inc.

P. Shafto, N. D. Goodman, and T. L. Griffiths. A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology*, 71:55–89, June 2014. ISSN 00100285. doi: 10.1016/j.cogpsych.2013.12.004. URL https://linkinghub.elsevier.com/retrieve/pii/S0010028514000024.

P. Silcock. Accelerated learning: A revolution in teaching method? *Education 3-13*, 31(1):48–52, Mar. 2003. ISSN 0300-4279. doi: 10.1080/03004270385200081. URL `https://doi.org/10.1080/03004270385200081`. Publisher: Routledge _eprint: https://doi.org/10.1080/03004270385200081.

R. D. Vale. The value of asking questions. *Molecular Biology of the Cell*, 24(6):680–682, Mar. 2013. ISSN 1059-1524. doi: 10.1091/mbc.e12-09-0660. URL `https://www.molbiolcell.org/doi/10.1091/mbc.e12-09-0660`. Publisher: American Society for Cell Biology (mboc).

S. Wang, C. Christensen, W. Cui, R. Tong, L. Yarnall, L. Shear, and M. Feng. When adaptive learning is effective learning: comparison of an adaptive learning system to teacher-led instruction. *Interactive Learning Environments*, 31(2):793–803, Feb. 2023. ISSN 1049-4820. doi: 10.1080/10494820.2020.1808794. URL `https://doi.org/10.1080/10494820.2020.1808794`. Publisher: Routledge _eprint: https://doi.org/10.1080/10494820.2020.1808794.