Pedagogical Value-Aligned Crowdsourcing: Inspiring the Wisdom of Crowds via Interactive Teaching

Runzhe Yang[1,2], Yexiang Xue[2], Carla P. Gomes[2]
Shanghai Jiao Tong University[1], Cornell University[2]

ABSTRACT

Crowdsourcing offers an economical means to leverage human wisdom for large-scale data annotation. However, the crowdsourced labeled data often suffer from low quality and significant inconsistencies, since the low-cost crowd workers are commonly lacking in corresponding domain knowledge and might make cursory choices. Most research in this area emphasizes the post-processing of the obtained noisy labels, which cannot radically ameliorate the quality of crowdsourcing service. In this paper, we focus on improving the worker's reliability during the label collecting process. We propose a novel game-theoretical framework of crowdsourcing, which formulates the interaction between the annotation system and the crowd workers as an incentivized pedagogical process between the teacher and the students. In this framework, the system is able to infer the worker's belief or prior from their current answers, reward them by performance-contingent bonus, and instruct them accordingly via near-optimal examples. We develop an effective algorithm for the system to select examples, even when the worker's belief is unidentifiable. Also, our mathematical guarantees show that our framework not only ensures a fair payoff to crowd workers regardless of their initial priors but also facilitates value-alignment between the annotation system (requester) and the crowd workers. Our experiments further demonstrate the effectiveness and robustness of our approach among different worker populations and worker behavior in improving the crowd worker's reliability.

KEYWORDS

Game Theory for practical applications; Reasoning about action, plans and change in multi-agent systems; Human-robot/agent interaction; Agents for improving human cooperative activities

1 INTRODUCTION

Recent decades have witnessed a huge benefit provided by crowdsourcing services to various applications of artificial intelligence, such as computer vision [4, 11], natural language processing [1, 23] and citizen science [8, 20], due to the fact that the emergent deep learning and other machine learning tools often heavily rely on huge amounts of manually annotated data. Compared with hiring experts to label, online crowdsourced data annotation is a cheaper and faster means to obtain a massive labeled dataset. However, crowdsourced labels are usually noisy and poor in quality, because of the problems with following two aspects of crowd workers' reliability during the label collection process:

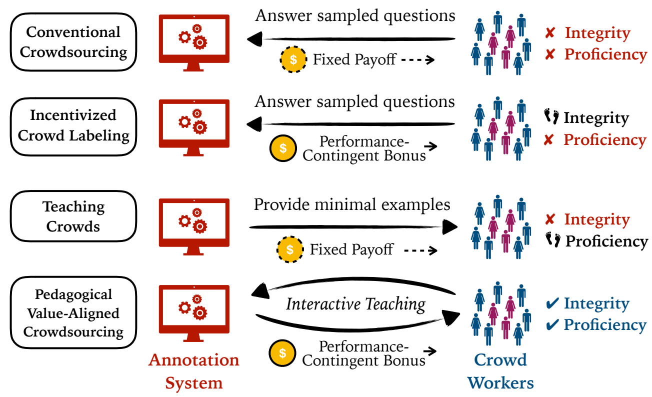


Figure 1: A comparison of crowdsourcing paradigms. The pedagogical value-aligned crowdsourcing provides an integrated solution to improving crowd workers' integrity and proficiency, and thus is more reliable than the others.

- **Integrity** As the annotation tasks are typically tedious, and the goal of workers is to earn rewards, there might be some dishonest workers who act hastily to finish tasks if the payoff is not linked to the accuracy of their responses.
- **Proficiency** Since the crowd workers often have different backgrounds, their competence to perform certain tasks may differ. And some workers' non-professional decisions would be inconsistent with the those of experts.

Most research has improved the quality of the final crowdsourced labels, without tackling these two problems head-on. These studies focused on developing algorithms to aggregate noisy crowdsourced data [3, 13, 19], allocating tasks to different individuals [10, 14] or designing a mechanism to mitigate inadvertent mistakes of crowd workers [17]. The reason for us to tolerate those "spammers" or unqualified workers is that building trust relationships with particular workers is hard, and the majority will provide reliable results in most cases [10]. However, when annotation tasks require specific domain knowledge that people commonly don't have, which is normal in citizen science projects, integrity and proficiency problems will significantly impair the accuracy of crowdsourced data. If we prequalify workers or adopt reputation-based mechanisms [22] to allow only skilled workers or those with good reputations to access tasks, it will increase the cost and still provide no guarantee. Therefore, reducing workers' dishonest behavior and improving their expertise during the label collection process are two inevitable challenges towards effective crowdsourcing.

However, how should we address the workers' integrity and proficiency issues during the label collection process? First, other

Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), M. Dastani, G. Sukthankar, E. Andre, S. Koenig (eds.), July 2018, Stockholm, Sweden

https://doi.org/doi

than engaging workers with intrinsic incentives such as curiosity [12] and enjoyment [7], there is some research on more realizable approaches like *incentivized crowd labeling* [6, 21], which uses performance-contingent bonuses to elicit worker effort under a limited budget. These ideas are creative for facilitating value alignment between the system and the crowd workers to address the integrity issue. Another orthogonal direction is to *teach the crowd* by letting them review expert example solutions [5, 18], where the examples are sampled from a small ground truth dataset labeled by experts in advance. The optimal example selection can be formulated in a submodular optimization problem [18], based on the assumptions that the crowd worker's prior is known and the worker's behavior is honest. However, this progress towards addressing integrity and proficiency problems is still limited because they treat the incentive and the teaching as two separate tools to enhance worker's reliability, and discount the interactive process in crowdsourcing. When incentivizing crowds, they may be confused about how to improve; when teaching crowds, they may be unwilling to follow.

Our main contribution is a general interactive crowdsourcing framework, *Pedagogical Value-aligned Crowdsourcing*, which leverages the system instruction as well as performance-contingent bonuses to address both integrity and proficiency problems. It formulates the interaction between the annotation system and the crowd workers as a multi-round pedagogical game between the "teacher" and the "student". **We establish methods for the teacher's reasoning in interactive teaching settings**, i.e. how the teacher estimates the student belief and which examples the teacher should choose. By analyzing the identifiability of the student belief and performance, we derive an unbiased estimate of the student performance. We provide an effective example suggestion algorithm that maximizes the minimal submodular surrogate objective, guaranteeing a provable improvement for the rational learner to achieve the teaching target even when the teacher is uncertain about student belief. **We further show two good properties of our framework** by investigating the student's pragmatic behavior: 1) ***fairness***– any worker who displays the same effort to learn from examples would be treated fairly regardless of their initial belief of the concept, and 2) ***value-alignment***– the more the student earns, the higher his final performance should be. Jointly, our framework can attract a broader crowd worker population to truly contribute to crowdsourcing tasks with high integrity and proficiency. We also design experiments on simulated workers, which further demonstrate the effectiveness and robustness of our approach among various worker populations and behavioral characteristics to improve crowd workers' reliability.

2 GAME-THEORETIC MODEL OF PEDAGOGICAL CROWDSOURCING

Our proposed pedagogical value-aligned crowdsourcing is a general framework for those human-powered tasks which require some domain knowledge. These tasks are common in the scientific process. For instance, scientists seek help from the crowds to recognize elephant calls from sound recordings of the rainforest, or to monitor the wasting disease of eelgrass [2].

Our key idea is to view the interaction between the annotation system and the crowd workers as an N-round two-stage *pedagogical game* between the "teacher" and the "student". In each round, the student first answers k sampled questions according to his current belief or policy (strategy) at the *practice stage*. At the *teaching stage*, the teacher infers the student's current belief, estimates his performance and sets an appropriate teaching target, then provides several teaching examples to help the student acquire the concept. If the student performance improves in the next round, then he will receive a bonus. The total bonus is awarded to the student after he finishes all the N-round annotation tasks. This paper considers the following binary data annotation settings.

Table 1: Frequently Used Notations in this Paper

Notation	Description	Notation	Description
$\mathcal{U}$	Instance Space	$\mathcal{Z}$	Finite Features Set
$\mathcal{X}$	Labeled Instance Set	(x, y)	Feature & Label
$\mathcal{G}$	Ground Truth Set	$\tilde{\mathcal{G}}$	All Instances & Labels
$\mathcal{H}$	$2^{\mathcal{Z}}$, Hypothesis Space	h^*	Target Concept
r	Total Bonus (Given by the end of the round N)	$\mathcal{R}^{\mathrm{S}}$	Student Immediate Reward (Bonus credits)
γ	Improvement Ratio	$\hat{\eta}_t$	Estimated Performance
$\tilde{\eta}_t$	Anticipated Performance	o_t	Revealed Examples
$P_\theta(\tilde{x}, \tilde{y})$	Observation Model	$\rho_\theta(h)$	Student's Belief

2.1 Binary Classification Settings

Suppose all instances in instance space $\mathcal{U}$ are independently drawn from the same distribution $\mathcal{D}$ over some finite feature space $\mathcal{Z}$, of which a small subset $\mathcal{X}$ is labeled by experts accordingly to the *target concept* $h^* \subseteq \mathcal{Z}$, but the remaining subset $\mathcal{U}\backslash\mathcal{X}$ is unlabeled. The ground truth $\mathcal{G} = \{(x, y) : x \in \mathcal{Z}|_{\mathcal{X}}\}$ consists of all the known pairs of feature and label $y = h^*(x) := \mathbb{1}_{x \in h^*} \in \{0, 1\}$, which is unrevealed to the crowd workers initially. The crowd workers are paid to annotate sampled instances $\tilde{x} \sim \mathcal{D}(\mathcal{Z})$ with binary labels $\tilde{y}$. The *hypothesis space* of the crowd workers, $\mathcal{H} \subseteq 2^{\mathcal{Z}}$, is a finite set containing possible hypotheses the workers might hold to label data. We assume *realizable settings* where the target concept $h^* \in \mathcal{H}$.

2.2 Formal Definition of Pedagogical Game

The pedagogical value-aligned crowdsourcing models the interactive teaching as a *pedagogical game*, an N-round two-stage Markov game between the annotation system and the crowd worker, where the system acts as the "teacher", **T**, who knows the ground truth $\mathcal{G}$ and the worker acts as the "student", **S**, who may not know. The student answers k questions (labels k sampled instances) each round according to their initial belief. The teacher uses examples in the ground truth set to help the student learn the target concept.

Definition 2.1 (Pedagogical Game). The pedagogical game is described as a tuple $\mathcal{M} = \langle \mathcal{S}, \{\mathcal{A}^{\mathrm{T}}, \mathcal{A}^{\mathrm{S}}\}, \mathcal{P}(\cdot|\cdot,\cdot,\cdot), \{\mathcal{R}^{\mathrm{T}}, \mathcal{R}^{\mathrm{S}}\}, t\rangle$ with following definitions:

- $\mathcal{A}^{\mathrm{T}} = 2^{\mathcal{G}}$ is an action space of the teacher. The teacher will give a variable number of teaching examples in each round. Let m_t be the number of examples in round t.
- $\mathcal{A}^{\mathrm{S}} = \mathcal{H}^k$ is an action space of the student. The student answers k questions each round by choosing hypotheses.
- $\mathcal{S} = 2^{\mathcal{G}} \times \tilde{\mathcal{G}}^k$ is a finite set of states. Each state $s_t = (o_t, g_t) \in \mathcal{S}$ represents revealed examples in $\mathcal{G}$ so far before round t

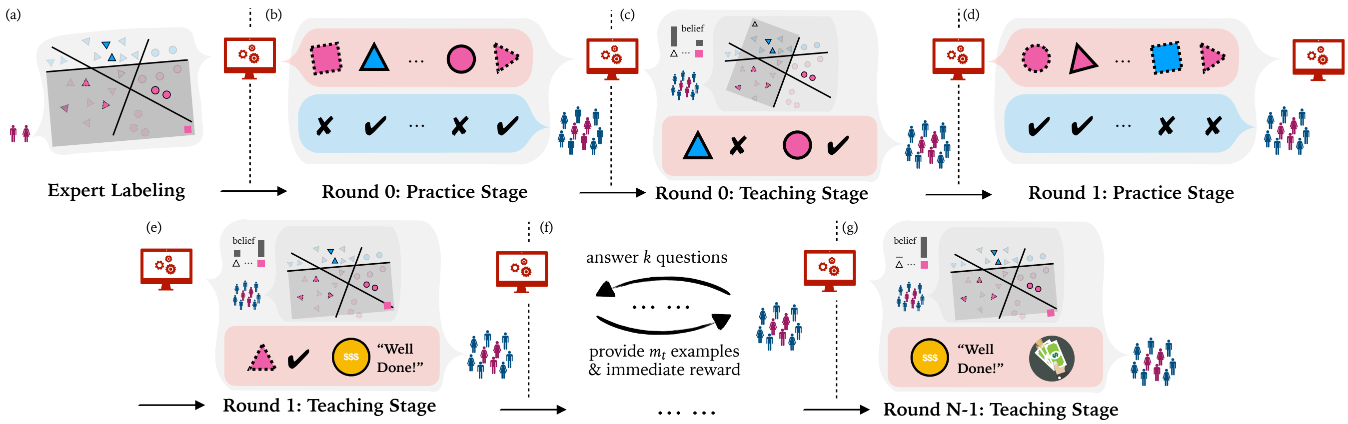


Figure 2: The basic process of pedagogical value-aligned crowdsourcing. A small ground truth set is labeled by experts, and the candidate features and hypotheses are elicited in advance (a). Each round, the annotation system random samples k instances for the crowd worker to label (b). By observing the worker's answers, the annotation infers the worker's belief and selects the most helpful examples (c). In the next round, the worker again labels k random sampled instances (d). If he improves, an immediate bonus credit will be given along with this round of new examples (e). Repeating practice and teaching stages until the (N-1)-th round (f), the worker will get paid by the end of the final round (g).

and k answers given by students in round t, where $\tilde{\mathcal{G}} = \{(\tilde{x}, \tilde{y}) : \tilde{x} \in \mathcal{Z} \text{ and } \tilde{y} \in \{0, 1\}\}$.

- $\mathcal{P}(s'|s, a_t^{\mathrm{T}}, a_t^{\mathrm{S}})$ is the transition model. In the pedagogical game it is partially deterministic, $o' \leftarrow o \cup_{i=1}^{k} a_{t,i}^{\mathrm{T}}$, whereas the transition of g is associated with sampling from $\mathcal{D}(\mathcal{Z})$.
- $\mathcal{R}^{\mathrm{T}}$ is the teacher's reward, which is equivalent to the student performance (see definition 2.2) in the final round.
- $\mathcal{R}^{\mathrm{S}} : \mathcal{S} \mapsto (-\infty, 1]$ is the student's immediate reward, where $\mathcal{R}^{\mathrm{S}}(s_t)$ indicates how many bonus credits should be given to student in round t. The total bonus student will gain in the final round is related to the cumulative reward $r = \left[\sum_t \mathcal{R}^{\mathrm{S}}(s_t)\right]^+$. If the cumulative reward is less than zero, the student will not earn any extra bonus.
- $t \in \{0, 1, \ldots, N-1\}$ is the round counter.

The basic process of the game is illustrated in figure 2. The game proceeds in N rounds. In each round t, there are two stages:

1) **Practice Stage** The student takes $a_t^{\mathrm{S}} = (h^{(1)}, \ldots, h^{(k)})$ sequentially to answer questions $(\tilde{x}_1, \ldots, \tilde{x}_k)$ independently sampled from data distribution $\mathcal{D}(\mathcal{Z})$. A partial state transition $g_t \leftarrow \{(\tilde{x}_i, h^{(i)}(\tilde{x}_i)) : i \in [k]\}$ will happen by the end of practice stage.
2) **Teaching Stage** The teacher observes answers g_t made by the student in the previous stage and infers the student's current belief. Then teacher selects m_t examples $a_t^{\mathrm{T}} = \{(x_1, y_1), \ldots, (x_{m_t}, y_{m_t})\}$ in the unrevealed ground truth set $\mathcal{G}/o_t$. One partial state transition $o_{t+1} \leftarrow o_t \cup_{i=1}^{k} a_{t,i}^{\mathrm{T}}$ will proceed by the end of teaching stage.

Two parties' behavior in this pedagogical game is defined by a pair of policies $(\pi^{\mathrm{T}}, \pi^{\mathrm{S}})$, that determine how teacher and student acts respectively. We assume the student independently picks k hypotheses following his policy, i.e. $\pi^{\mathrm{S}}(a_t^{\mathrm{S}}|o_t) = \prod_{i \in [k]} \rho_t(h^{(i)})$, where ρ_t is student's belief over hypothesis space. Since there exists a bijection between $\pi^{\mathrm{S}}(\cdot|o_t)$ and ρ_t, we use terms {student belief, student policy} interchangeably in the following article.

Definition 2.2 (Student Performance). The student performance η of a student's policy $\pi^{\mathrm{S}}(\cdot|o_t)$ is the expected label accuracy measured on the ground truth set

$$\eta(\pi^{\mathrm{S}}(\cdot|o_t)) := \mathbb{E}_{h \sim \rho_t}\left[\frac{1}{|\mathcal{G}|} \sum_{(x,y) \in \mathcal{G}} \mathbb{1}\{h(x) = y\}\right],$$

where $\mathbb{1}\{\cdot\} = 1$ if the condition in {} is true otherwise it is 0, and $\rho_t(\cdot)$ is the equivalent belief to $\pi^{\mathrm{S}}(\cdot|o_t)$. For convenience, we also define $\eta(\rho_t) := \eta(\pi^{\mathrm{S}}(\cdot|o_t))$.

In every teaching stage, the teacher will set a target on the student's next round performance $\tilde{\eta}_{t+1}$ and select examples to help the student reach the target, i.e. $\eta(\tilde{\pi}^{\mathrm{S}}(\cdot|o_{t+1})) \geq \tilde{\eta}_{t+1}$. If the estimated student performance in the next round, $\hat{\eta}_{t+1}$ surpasses the target performance, the student will receive full bonus credits as an immediate reward from the teacher. The teacher's reward is the final student performance, and the student's reward is

$$\mathcal{R}^{\mathrm{S}}(s_{t+1}) = \begin{cases} 1, & \hat{\eta}_{t+1} \geq \tilde{\eta}_{t+1} \\ \frac{\hat{\eta}_{t+1} - \hat{\eta}_t}{\tilde{\eta}_{t+1} - \hat{\eta}_t}, & \hat{\eta}_{t+1} < \tilde{\eta}_{t+1} \end{cases}$$

in the $(t+1)$-th round. It can be shown that this reward design for the crowd workers is fair and motivational.

In the next two sections, we will first discuss the teacher's strategy in the pedagogical game, which is associated with how the teacher estimates the student's current policy $\pi^{\mathrm{S}}(\cdot|s_t)$ (or belief $\rho_t(\cdot)$), sets target performance $\tilde{\eta}_{t+1}$ for student's next round performance and gives most helpful examples a_t^{T}, or all in one phrase, the *pedagogical reasoning*, and then analyze the student's *pragmatic behavior* to see how the crowd workers, no matter their honesty and proficiency, are incentivized to participate in the pedagogical game and to pursue the value aligned with the system.

3 ON TEACHER'S PEDAGOGICAL REASONING

In order to set appropriate teaching targets and teach the student with the most helpful examples, the teacher needs to assess the student's belief over hypotheses space as well as the student current performance. Different from previous work on teaching crowds [18] and machine teaching [24] which assume the student's prior (initial belief) is known, the teacher in the pedagogical value-aligned crowdsourcing estimates student's belief by observing answers each round in the interaction and then determines the best batch of examples for rational learners accordingly.

3.1 Belief Estimation

The parametric probabilistic model explaining the teacher's observation in pedagogical game is described below. In each round a student picks hypotheses $a^{\mathrm{S}} = (h^{(1)}, \dots, h^{(k)})$ in a generalized Bernoulli process, each step with a probability $\rho(h^{(i)})$, where ρ is a categorical distribution parameterized as $\rho_\theta(h_i) = e^{\theta_i} / \sum_{h_j \in \mathcal{H}} e^{\theta_j}$, $\theta \in \Theta = \mathbb{R}^{|H|}$, which guarantees a "grain of truth" for rational learning [9]. The teacher can only observe k answers $g_t = \{(\tilde{x}_1, \tilde{y}_1), \dots, (\tilde{x}_k, \tilde{y}_k)\}$ from the student, where $\tilde{y}_i = h^{(i)}(\tilde{x}_i)$ for $i \in [k]$. The goal of belief estimation is to find a point estimator of student's belief $\rho_\theta(\cdot)$ by samples drawn from the marginal distribution of observation

$$P_\theta(\tilde{x}, \tilde{y}) = \sum_{h \in \mathcal{H}} P_\theta(\tilde{x}, \tilde{y} | h) \cdot \rho_\theta(h),$$

where $P_\theta(\tilde{x}, \tilde{y}|h) = P_{\mathcal{D}}(\tilde{x}) \cdot \mathbb{1}\{h(\tilde{x}) = \tilde{y}\}$.

However, an accurate estimation of the student's belief is not always available since the observation model could be unidentifiable [15], which means there would be multiple student beliefs leading to the same observation distribution. In this case, the teacher may draw inconsistent conclusions from the same observed answers. To meaningfully discuss the belief estimation, we have to first verify the identifiability of student belief and performance.

Definition 3.1 (Identifiability [15]). In a probability space $(\Omega, \mathcal{E}, P)$, where F is the σ-algebra defined on sample space Ω and $P = \{P_\theta : \theta \in \Theta\}$ is a family of parameterized probability measure, two points θ_1 and $\theta_2 \in \Theta$ are said to be observationally equivalent (written as $\theta_1 \sim \theta_2$) if $P_{\theta_1}(E) = P_{\theta_2}(E), \forall E \in \mathcal{E}$.

(1) The point θ_1 is said identifiable if $\theta_1 \sim \theta_2 \Rightarrow \theta_1 = \theta_2$.
(2) The model P is said identifiable if the quotient set $\Theta / \sim$ is the finest possible partition.
(3) A function $\varphi(\theta)$ is identifiable if $\forall \theta_1, \theta_2 \in \Theta, \theta_1 \sim \theta_2 \Rightarrow \varphi(\theta_1) = \varphi(\theta_2)$.

Proposition 3.2. *In the pedagogical game, the student's belief can be unidentifiable, whereas the student performance is always identifiable.*

The basic idea to prove proposition 3.2 is to show that the linear transformation mapping student belief space to the observation space is commonly rank-deficient, while equivalent observation implies the same student performance. As illustrated in figure 3, the teacher can estimate the current student performance using any belief in the equivalent class of the real belief.

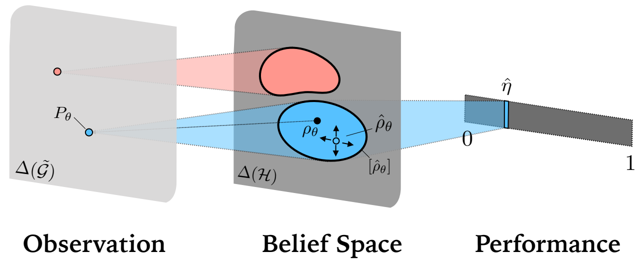


Figure 3: Pedagogical reasoning about student belief and performance. There might be multiple beliefs corresponding to the same observation distribution. But the teacher can use any of them to infer the student performance.

The following theorem provides a method to estimate the student performance unbiasedly.

Theorem 3.3. *If $\hat{\rho}_\theta$ is a maximum likelihood estimate (MLE) of the student's belief, the induced estimator of student performance, $\eta(\hat{\rho}_\theta)$, is unbiased.*

Proof. Let k i.i.d random variables $(\tilde{X}_1, \tilde{Y}_1), \dots, (\tilde{X}_k, \tilde{Y}_k) \sim P_\theta$ be the observation. Maximum likelihood estimates of θ is derived by maximizing $\mathcal{L}(\theta; \tilde{X}, \tilde{Y}) := \prod_{i=1}^k \prod_{(x,y) \in \tilde{\mathcal{G}}} P_\theta(x, y)^{\Lambda_i^{x,y}}$, where random variable indicators $\Lambda_i^{x,y} := \mathbb{1}\{\tilde{X}_i = x, \tilde{Y}_i = y\}$. Note that $\sum_{(x,y) \in \tilde{\mathcal{G}}} P_\theta(x, y) = 1$, the concave function $\mathcal{L}(\theta; \tilde{X}, \tilde{Y}) = \prod_{(x,y) \in \tilde{\mathcal{G}}} P_\theta(x, y)^{\sum_{i=1}^k \Lambda_i^{x,y}}$ of $P_\theta(x, y)$ has a global maximum $\mathcal{L}^*$. Consider the Lagrange function

$$\mathcal{F}(\theta) = \log \mathcal{L}(\theta; \tilde{X}, \tilde{Y}) - \lambda \left(\sum_{(x,y) \in \tilde{\mathcal{G}}} P_\theta(x, y) - 1 \right).$$

If $\hat{\theta}$ is a maximum estimate, then $\mathcal{F}(\hat{\theta})$ achieves its maximum $\log L^*$. Assume $P_\theta(x, y) > 0$ for all $(x, y) \in \tilde{\mathcal{G}}$ (otherwise $(x, y) \notin \tilde{\mathcal{G}}$ almost surely), we have

$$\frac{\partial \mathcal{F}(\theta)}{\partial P_\theta(x, y)} \Big|_{\theta = \hat{\theta}} = \left(\frac{\sum_{i=1}^k \Lambda_i^{x,y}}{P_\theta(x, y)} - \lambda \right) \Big|_{\theta = \hat{\theta}} = 0,$$

for all $(x, y) \in \tilde{\mathcal{G}}$. Therefore, when $P_{\hat{\theta}}(x, y) = \frac{\sum_{i=1}^k \Lambda_i^{x,y}}{\lambda}$, where $\lambda = \sum_{(x,y) \in \tilde{\mathcal{G}}} \sum_{i=1}^k \Lambda_i^{x,y} = k$, the likelihood function achieves its maximum. When the model is unidentifiable, there would be multiple MLEs $\hat{\theta}$, but all of them have the same induced estimator of student performance $\hat{\eta}$ as we argued in Proposition 3.2. Recall the definition 2.2 , the $\hat{\eta}$ is

$$\hat{\eta} = \eta(\rho_{\hat{\theta}}) = \frac{1}{|\mathcal{G}|} \sum_{(x,y) \in \mathcal{G}} \frac{1}{P_{\mathcal{D}}(x)} \cdot \frac{\sum_{i=1}^k \Lambda_i^{x,y}}{k}$$

Taking expectation over all observations results in $\mathbb{E}_{P_\theta}[\hat{\eta}] = \eta(\rho_\theta)$. Hence, the induced $\hat{\eta}$ is an unbiased estimator. □

Since finding a maximum likelihood estimate of the student belief is necessary for teaching and not difficult to compute, we can obtain the unbiased estimate of student performance as a byproduct, which reduces extra computations.

3.2 Optimal Teaching

The interactive teaching requires the teacher to set appropriate teaching targets based on students' current performance and give students the most helpful examples according to their current belief. We first introduce the setting of teaching targets and the learning model of honest students.

Definition 3.4 (Teaching Target). Given the estimate of the student current performance $\hat{\eta}_t$, the teaching target is the standard of the next round student performance

$$\tilde{\eta}_{t+1} = \gamma \cdot (1 - \hat{\eta}_t) + \hat{\eta}_t,$$

where $\gamma \in (0, 1)$ is called "improvement ratio".

Supposing a student is cooperative and rational, cognitive psychologists suggest the effects of teaching examples can be captured by the Bayesian model [16]. Formally, the student rational learning is defined as following a belief updating process.

Definition 3.5 (Rational Learning). Given a set of examples o_t, the current belief ρ_{t-1}, the rational learning is to update belief as its Bayesian posterior,

$$\rho_t(h) \leftarrow \frac{P_{\text{learner}}(o_t|h)\rho_{t-1}(h)}{\sum_{h_j \in \mathcal{H}} P_{\text{learner}}(o_t|h_j)\rho_{t-1}(h_j)}$$

where $P_{\text{learner}}(o_t|h) = \prod_{(x,y)\in o_t} \sigma_\alpha(h(x), y)$, and σ_α is a noise-tolerant likelihood function,

$$\sigma_\alpha(h(x), y) = \frac{1}{1 + e^{-\alpha\cdot(1-2|h(x)-y|)}}.$$

We use $Z(\boldsymbol{\rho}) = \sum_{h_j\in\mathcal{H}} \prod_{(x,y)\in o_t} \sigma_\alpha(h(x), y)\rho(h_j)$ and the operator $\boldsymbol{\psi}(\boldsymbol{\rho}_{t-1}, o_t) := \boldsymbol{\rho}_t$ to denote the partition function and the rational learning in short form.

The scaling parameter α controls the impact of shown examples on student's belief, which can also be interpreted as student learnability assumed by the teacher. The larger the α is, the stronger the impact of counterexamples will be on eliminating student inconsistent hypotheses.

To select the best examples, the teacher estimates the effect of new examples a_t^{T} to a model student with displayed learnability α and current belief $\hat{\rho}_\theta$. The anticipated student performance is

$$\hat{\eta}_{t+1}(a_t^{\mathrm{T}}) = \eta(\boldsymbol{\psi}(\hat{\boldsymbol{\rho}}_\theta, o_t \cup a_t^{\mathrm{T}})) = \boldsymbol{u}^T \boldsymbol{\Psi} \hat{\boldsymbol{\rho}}_\theta,$$

where each entry of $|\mathcal{H}|$-vector $\boldsymbol{u}$ is the accuracy of hypothesis on the ground truth, $u(h) = \frac{1}{|G|}\sum_{(x,y)\in\mathcal{G}} \mathbb{1}\{h(x) = y\}$; and $\boldsymbol{\Psi}$ is an $|\mathcal{H}| \times |\mathcal{H}|$-diagonal matrix, whose diagonal entries are $\Psi_{h,h} = \frac{1}{Z(\hat{\boldsymbol{\rho}}_\theta)} \prod_{(x,y)\in o_t \cup a_t^{\mathrm{T}}} \sigma_\alpha(h(x), y)$.

Finding the minimal a_t^{T} such that the student next round performance $\hat{\eta}_{t+1}(a_t^{\mathrm{T}})$ reaches the teaching target $\tilde{\eta}_{t+1}$ is an NP-hard combinatorial optimization problem [18]. Instead of directly solving this difficult problem, we try to improve a surrogate student performance $\eta'_{t+1} = 1 - (1 - \boldsymbol{u})^T \boldsymbol{\Psi}' \hat{\boldsymbol{\rho}}_\theta$, where $\boldsymbol{\Psi}' = Z(\hat{\boldsymbol{\rho}}_\theta)\boldsymbol{\Psi}$ discarding the partition function, which makes η'_{t+1} easier to improve.

To see the effectiveness of this approach, we first introduce two lemmas. In Lemma 3.6 we derive upper and lower bounds of real performance by using a surrogate student performance. Lemma 3.7 shows the surrogate loss is a monotonic submodular function. Using these two lemmas, we show the effectiveness and efficiency of greedy improving the surrogate performance in Theorems 3.8 & 3.9 , even when the student belief is unidentifiable.

Algorithm 1 Pedagogical Reasoning - Teaching

procedure TEACHING(o_t, $[\hat{\boldsymbol{\rho}}_{\theta_t}]$, $\mathcal{G}$, $\{\boldsymbol{\psi}_{(x,y)}\}$)
  $a_t^{\mathrm{T}}, \boldsymbol{\psi} \leftarrow \{\}, \bigotimes_{(x,y)\in o_t} \boldsymbol{\psi}_{(x,y)}$
  $\Delta_\beta \leftarrow (1 - (1 - \gamma))(1 - \hat{\eta}_t)$
  while $\tilde{\eta}'_{t+1} - \hat{\eta}_t < \Delta_\beta$ **do**
    for $(x, y) \in \mathcal{G}\backslash(o_t \cup a_t^{\mathrm{T}})$ **do**
      $E_{(x,y)} \leftarrow \min_{\boldsymbol{\rho}\in[\hat{\boldsymbol{\rho}}_{\theta_t}]} 1 - \boldsymbol{\psi}^T_{(x,y)}(\boldsymbol{\psi} \otimes (1 - \boldsymbol{u}) \otimes \boldsymbol{\rho})$
    end for
    $a \leftarrow \arg\max_{(x,y)\in\mathcal{G}} E_{(x,y)}$
    $\tilde{\eta}'_{t+1} \leftarrow E_a$
    $a_t^{\mathrm{T}}, \boldsymbol{\psi} \leftarrow a_t^{\mathrm{T}} \cup \{a\}, \boldsymbol{\psi}_a \otimes \boldsymbol{\psi}$
    $\beta \leftarrow \min_{\boldsymbol{\rho}\in[\hat{\boldsymbol{\rho}}_{\theta_t}]} \boldsymbol{\psi}^T \boldsymbol{\rho}$
    $\Delta_\beta \leftarrow (1 - \beta(1 - \gamma))(1 - \hat{\eta}_t)$
  end while
  return a_t^{T}
end procedure

LEMMA 3.6. *Let $\boldsymbol{\Psi}' = Z(\hat{\boldsymbol{\rho}}_\theta) \cdot \boldsymbol{\Psi}$, and $\tilde{\eta}' = 1 - (\mathbf{1} - \boldsymbol{u})^T \boldsymbol{\Psi}' \hat{\boldsymbol{\rho}}_\theta$, we have $\tilde{\eta}' \geq \tilde{\eta} \geq \frac{1}{\beta} \cdot \tilde{\eta}' - \frac{1-\beta}{\beta}$, where $\beta \in (0, Z(\hat{\boldsymbol{\rho}}_\theta)]$ controls the scaling ratio.*

LEMMA 3.7. *The surrogate performance $\tilde{\eta}'$ is a monotonic submodular function of the teaching examples o_t.*

We provide a proof sketch. First, write $\eta' = \sum_{h\in\mathcal{H}} (1 - u(h)) \cdot \hat{\rho}_\theta(h) \cdot P(o_t|h) + \hat{\eta}$, where $P(o_t|h) = 1 - \prod_{(x,y)\in o_t} \sigma_\alpha(h(x), y)$. Then we verify for any $o_a \subseteq o_b \subseteq \mathcal{G}$, the set function $P(o_b|h) \geq P(o_a|h)$ and $P(o_b \cup (x, y)|h) - P(o_b|h) \leq P(o_a \cup (x, y)|h) - P(o_a|h)$. □

THEOREM 3.8. *When the model is unidentifiable, i.e., there is an equivalent class $[\hat{\boldsymbol{\rho}}_\theta] \neq \{\hat{\boldsymbol{\rho}}_\theta\}$, greedily selecting examples $a_t^{\mathrm{T}} \subseteq \mathcal{G}\backslash o_t$ to increase the worst improvement $E_{a_t^{\mathrm{T}}} = \min_{\boldsymbol{\rho}_\theta\in[\hat{\boldsymbol{\rho}}_\theta]} (\mathbf{1} - \boldsymbol{u})^T (\boldsymbol{I} - \boldsymbol{\Psi}'_{a_t^{\mathrm{T}}})\boldsymbol{\rho}_\theta$ until $E_{a_t^{\mathrm{T}}}$ is no less than $(1 - \beta(1 - \gamma))(1 - \hat{\eta})$, ensures $\tilde{\eta}$ achieves the teaching target $(\gamma \cdot (1 - \hat{\eta}) + \hat{\eta})$.*

PROOF. Theorem 3.8 is a direct consequence of Lamma 3.6. □

THEOREM 3.9. *Given the current student performance $\hat{\eta}$ and the target improvement ratio γ, by greedily providing $\mathtt{OPT}(\tilde{\eta}'_\xi) \cdot \log \frac{1}{\xi\beta(1-\gamma)}$ examples it is guaranteed to improve the student performance to the teaching target $\gamma(1-\hat{\eta})+\hat{\eta}$, where $\tilde{\eta}'_\xi = \hat{\eta}+[1-\beta(1-\xi)(1-\gamma)]\cdot(1-\hat{\eta})$ and $\mathtt{OPT}(\cdot)$ is the minimal number of examples to increase surrogate performance to a certain value.*

PROOF. From lemma 3.7, we know that $\tilde{\eta}'$ is a nonnegative monotone submodular function. Using the result of greedy maximization of submodular function that for any ℓ and k,

$$f(S_\ell) \geq \left(1 - e^{-\ell/k}\right) \max_{S:|S|=k} f(S),$$

where S_ℓ is the set picked after ℓ steps.

Let $k^* = \mathtt{OPT}(\tilde{\eta}'_\xi)$, when $\ell \geq k^* \log \frac{1}{\xi\beta(1-\gamma)}$, we have

$$\begin{aligned} \tilde{\eta}'(S_\ell) - \hat{\eta} &\geq (1 - \xi\beta(1 - \gamma))(\tilde{\eta}'_\xi - \hat{\eta}) \\ &\geq (1 - \beta(1 - \gamma))(1 - \hat{\eta}). \end{aligned}$$

This indicates greedily providing $\texttt{OPT}(\tilde{\eta}'_{\xi}) \cdot \log \frac{1}{\xi\beta(1-\gamma)}$ more examples will absolutely improve the student performance to the teaching target. □

Jointly, theorems 3.8 and 3.9 depict the effectiveness and approximate optimality of greedy algorithm 1 (where the vector $\boldsymbol{\psi}_{(x,y)} = [\sigma_\alpha(h(x), y)]^T_{h\in\mathcal{H}}$) , to improve the workers' proficiency.

4 ON STUDENT'S PRAGMATIC BEHAVIOR

In this section, we will investigate the crowd worker's pragmatic behavior in the pedagogical crowdsourcing. If a worker is dishonest and behaves badly, he is supposed to earn no extra bonus from the task. However, those hardworking but initial non-professional workers are supposed to earn more money if they learn from provided examples. Our aim in this section is to show: 1) The worker feels fair about the bonus no matter what prior he initially holds; 2) Workers will pursue the aligned value with the system, i.e., the more the student earns, the higher his final performance should be.

4.1 Fairness: Prior Does not Matter

We state the fairness property as the following theorem:

Theorem 4.1. *Two workers who have different initial beliefs (priors) but the same displayed effort will earn the same amount of bonuses.*

Here, the displayed effort is student's relative improvement to the teaching target each round. It is easy to verify that the final bonuses only depends on how well they achieve the adaptive teaching targets which eliminate the influence of prior.

This fair design has many benefits for the crowd workers. First, the skilled crowd workers who are good at certain annotation tasks can happily get the full bonus, as long as they keep their performance during the task, since the required improvement is negligible. Second, a student who has almost no background knowledge can earn the full bonus as long as he can reach the teaching target every round, since our bonus mechanism is prior-free and encourages improvement. If the system gives bonus according to his absolute performance every round, he will receive almost the lowest payoff in the task and soon he will fail to contribute. Moreover, this fair design will attract more workers, because it really relaxes the rules for entry. Our results in the next section show that, in order to earn more bonuses, his performance must finally improve to required levels with the help of teaching examples. Therefore, these newly recruited workers indeed can make contributions to the task.

4.2 Value-Alignment: Rational Learning as Student's Optimal Policy

From the organizer's perspective, this teaching incentive design actually facilitates the value alignment between the teacher and the student. The more the teacher pays, the higher the student performance is guaranteed for any individual worker. Our main results of value alignment are shown below.

Lemma 4.2. *If a student* S *earns* $\omega \in (0, 1)$ *in each round, then his his overall improvement* $\Delta = \eta_{N-1} - \eta_0$ *should be no less than* $\omega\bar{\Delta}$*, where* $\bar{\Delta} = \bar{\eta}_{N-1} - \bar{\eta}_0$ *is the overall improvement of a "model" student who has the same initial performance and exactly achieves each round target.*

Proof. Let $\Delta(\eta) = \gamma(1-\eta)$ be the required improvement for achieving the teaching target of η. It is easy to verify two properties: if $\eta^{(1)} \le \eta^{(2)}$, then (a) $\Delta(\eta^{(1)}) \ge \Delta(\eta^{(2)})$, and (b) $\eta^{(1)} + \Delta(\eta^{(1)}) \le \eta^{(2)} + \Delta(\eta^{(2)})$. Let $\bar{\eta}_0, \bar{\eta}_1, \dots, \bar{\eta}_{n-1}$ be each round of performance of the model student, and $\eta_0, \eta_1, \dots, \eta_{n-1}$ be each round of performance of student S. $\eta_0 = \bar{\eta}_0$ Suppose the teacher estimates his performance accurately. Using the property (b), we know that $\eta_t \le \bar{\eta}_t$, for all $t \in \{0, \dots, N-1\}$, since $\eta_1 < \bar{\eta}_1$ and $\eta_{t-1} \le \bar{\eta}_{t-1} \Rightarrow \eta_t \le \bar{\eta}_t$. Furthermore, according to property (a) we know each round of S's improvement $\omega\Delta(\eta_{t-1})$ should be no less than $\omega\Delta(\bar{\eta}_{t-1})$. Therefore,

$$\Delta = \sum_{t=1}^{N-1} \omega\Delta(\eta_{t-1}) \ge \omega \sum_{t=1}^{N-1} \Delta(\bar{\eta}_{t-1}) = \omega\bar{\Delta}.$$

Hence, a student earning constant $\omega \in (0, 1)$ each round has overall improvement at least $\omega\bar{\Delta}$, where $\bar{\Delta}$ is the overall improvement of the model student. □

Theorem 4.3. *If a student* S *earns* $\omega \in (0, 1)$ **on average** *each round, then his overall improvement should be no less than that of the "model" student in lemma 4.2.*

Proof. The overall improvement of the model student can be written as

$$\bar{\Delta} = \sum_{t=1}^{N-1} \gamma(1-\gamma)^{t-1}(1-\bar{\eta}_0) = \left(1-(1-\gamma)^{N-1}\right)(1-\eta_0).$$

Let $\omega_1, \omega_2, \dots, \omega_{N-1} \in (-\infty, 1)$ be each round of bonus credits received by the student S, whose average is $\omega = \frac{1}{N-1}\sum_{t=1}^{N-1}\omega_t$. And let $\kappa_t = \frac{\eta_t - \eta_{t-1}}{\bar{\eta}_{t-1} - \eta_{t-1}} \in (-\infty, 1/\gamma), t \in [N-1]$ be the student real improvement ratio, whose average $\kappa \ge \omega$ since by definition $\omega_t = \begin{cases} 1, & \kappa_t > 1 \\ \kappa_t, & \kappa_t \le 1 \end{cases}$. The student S's overall improvement is

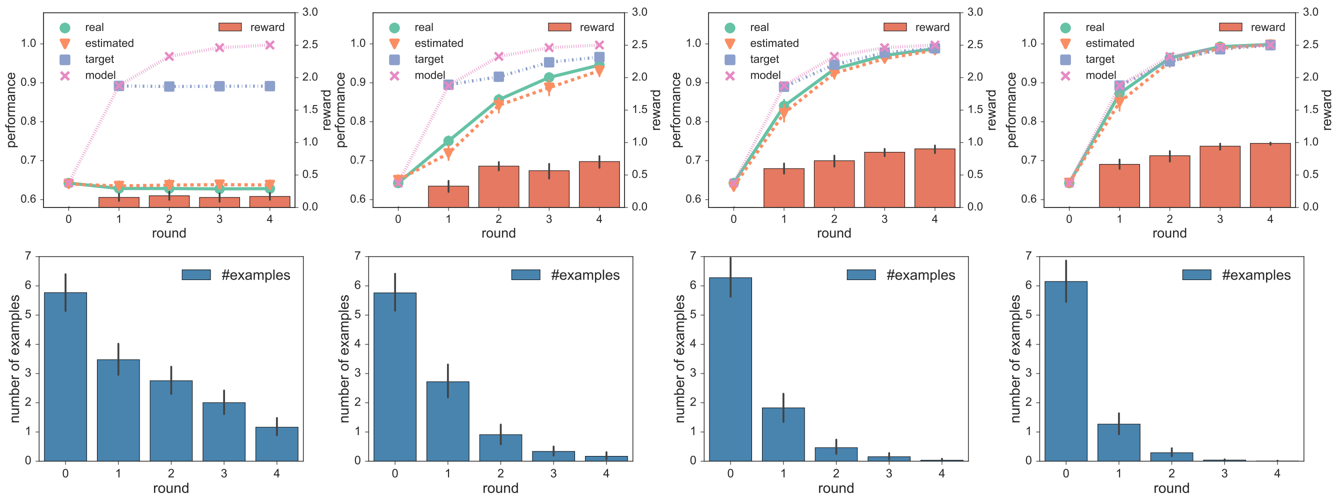
$$\begin{aligned} \Delta &= \sum_{t=1}^{N-1} \kappa_t\gamma\left(\prod_{i=1}^{t-1}(1-\kappa_i\gamma)\right)(1-\eta_0) \\ &= \left(1-\prod_{t=1}^{N-1}(1-\kappa_t\gamma)\right)\cdot(1-\eta_0) \\ &\ge \left(1-(1-\kappa\gamma)^{N-1}\right)\cdot(1-\eta_0) \\ &\ge \left(1-(1-\omega\gamma)^{N-1}\right)\cdot(1-\eta_0) \end{aligned}$$

The second last step is because of the AM-GM inequality,

$$\left(\prod_{t=1}^{N-1}(1-\kappa_t\gamma)\right)^{\frac{1}{(N-1)}} \le \frac{\sum_{t=1}^{N-1}(1-\kappa_t\gamma)}{N-1} = 1-\kappa\gamma$$

where the equality holds when $\kappa_1 = \cdots = \kappa_{N-1} = \omega_1 = \cdots = \omega_{N-1} = \omega$. According to lemma 4.2, $\Delta \ge \omega\bar{\Delta}$ in this case. Therefore, we conclude that S's overall improvement is at least $\omega\bar{\Delta}$. □

In other words, theorem 4.3 indicates a student earning final bonus $r = (N-1)\cdot\omega$ guarantees a teacher's reward $\mathcal{R}^{\mathbf{T}} \ge \eta_0 + \omega\bar{\Delta}$. It also indicates the fact that the only way to earn the full bonus ($\omega = 1$) is to learn at least as well as the model student. If a worker



(a) Spammer **(b) Weak Learner** **(c) Median Learner** **(d) Strong Learner**

Figure 4: The dynamics of learning and teaching for different worker behavior when $\gamma = 0.7$.

is dishonest, his expected overall improvement is zero. Thus he will earn no bonus other than basic payment from the tasks. That being said, rationally learning becomes the student's optimal policy to achieve the Stackelberg equilibrium in the pedagogical game.

5 EXPERIMENTS

We conduct pedagogical crowdsourcing experiments on various simulated crowd workers for an interesting cognitive task, "ROSA"[1] game, in order to observe the dynamics of teaching and learning, explore the effects of different improvement ratios, and verify fairness and value-alignment properties of the pedagogical crowdsourcing.

5.1 Experimental Setup

The goal of our cognitive task is to make crowd workers label sampled graphics to identify whether they are "ROSA". This binary classification is not trivial for the crowd workers because the target concept might be unclear at the beginning. Therefore, proficiency issue would be critical.

Task Description ("ROSA") There is a large dataset containing graphics of different *colors*, *shapes* and *border styles*. The crowd workers are required to label a total of 75 instances to indicate whether they are "ROSA" or not in $N = 5$ rounds, i.e. $k = 15$ questions each round.

The feature space $\mathcal{Z}$ includes all the combinations of:

- *colors* (5): *blue, red, yellow, green, pink.*
- *shapes* (3): *triangle, square, circle.*
- *border stype* (2): *real line, dotted line.*

Therefore, the feature space size is $|\mathcal{Z}| = 5\times3\times2 = 30$. The hypothesis space contains all the hypotheses which associate with single attributes of the graphics, i.e. $\mathcal{H} = \{h_{\text{blue}}, h_{\text{red}}, h_{\text{yellow}}, h_{\text{green}}, h_{\text{pink}}, h_{\text{tri}}, h_{\text{sqr}}, h_{\text{cir}}, h_{\text{real}}, h_{\text{dot}}\}$. For example, $h_{\text{dot}}(\cdot)$ assign +1 to all the instances with dotted border, and assign 0 to all the instances without dotted border (they should be real lined border in our settings).

[1] In Spanish, word "rosa" means the color pink. We teach the workers who may not know this concept.

The target concept "ROSA" means "pink" ($h^* = h_{\text{pink}}$), which is not necessarily clear to all the crowd workers. A small subset $\mathcal{X}$ containing 50 instances is randomly sampled and labeled by $h_{\text{pink}}(\cdot)$ as the ground truth set $\mathcal{G}$.

Worker Behavior. We observe workers with different learnability, as well as the dishonest "spammer". The honest worker will Beyasianly learn from teaching examples as we defined in definition 3.5. A stronger learner has larger α in the noise-tolerant likelihood function. In our experimental setting, $\alpha_{\text{strong}} = 3, \alpha_{\text{median}} = 1.2, \alpha_{\text{weak}} = 0.4$. The dishonest worker updates his belief randomly, which is for illustrating how the *integrity issue* is addressed in the pedagogical game.

Worker Population. We also conduct the experiment on four worker populations who have different initial belief. Two know what "ROSA" means and have relatively high initial performance levels 0.808 and 0.904 , and the other two are the typical inexpert workers who perform only at levels 0.546 and 0.714 initially.

Teaching Settings. We test different teaching targets of improvement ratios $\gamma \in \{0.3, 0.5, 0.7, 0.9\}$. We set $\alpha = 3$ for our teaching algorithm. To avoid the numerical issue when $\tilde{\eta}_{t+1}$ and $\hat{\eta}_t$ are close, we set the immediate reward $\mathcal{R}^{\text{S}} = \frac{\hat{\eta}_{t+1}-\hat{\eta}_t+\epsilon}{\tilde{\eta}_{t+1}-\hat{\eta}_t+\epsilon}$ when $\hat{\eta}_{t+1} < \hat{\eta}_t$, where $\epsilon = 0.07$. We simulate 200 workers with certain initial belief for each experimental configuration to obtain the following results.

5.2 Results on Simulated Workers

General Analysis on Teaching and Learning Dynamics. Figure 4 reports the dynamics of workers' performance, teaching targets, and the system's estimation as well as the reward gain, and the number of elicited examples when we set the teaching target as 70% of the maximum improvement. The curve of the estimated student performance (the orange dashed line) almost accurately tracks the temporal changes in the real performance for different workers with acceptable standard errors. As the teaching progresses, the performance of honest workers goes higher, and they require fewer

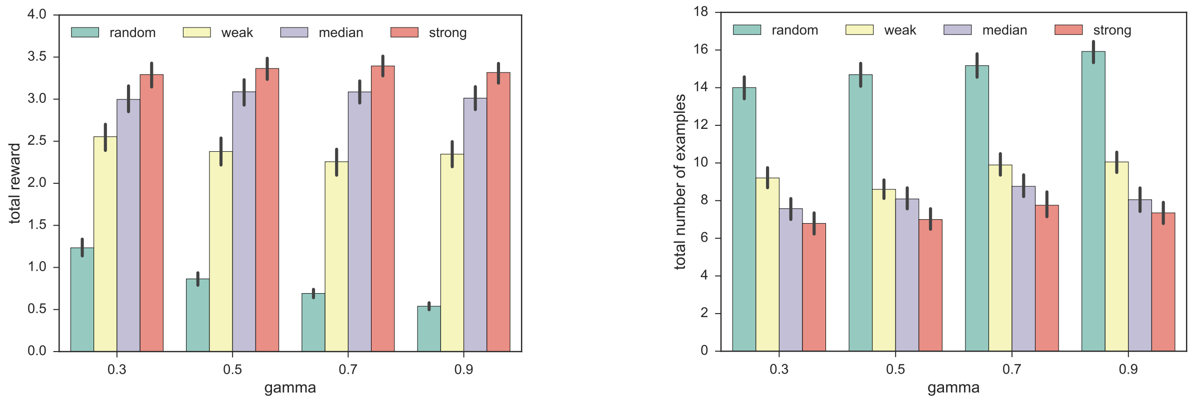
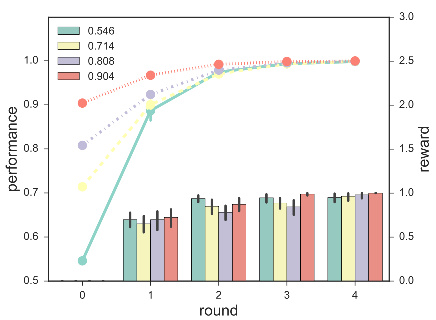
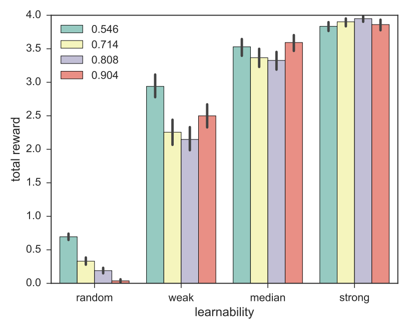


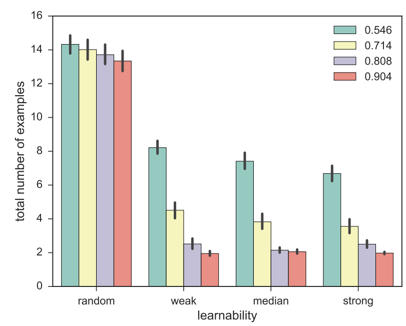
Figure 5: The final bonus and number of examples for different worker behavior when $\gamma \in \{0.3, 0.5, 0.7, 0.9\}$.



(a) Median learners



(b) Reward



(c) Examples

Figure 6: The dynamics of learning and teaching, as well as final bonus and number of examples for workers with different initial performance $\eta \in \{0.546, 0.714, 0.808, 0.904\}$, when the target improvement ratio is $\gamma = 0.7$.

and fewer examples. Even for weak learners, their final performance is close to that of the model student. The dishonest spammers, who require more examples, gain almost no reward each round.

Effects of Different Teaching Targets. Our simulation on various levels of teaching targets shows more robust results. As Figure 5 reflects, in any level of teaching targets, the bonus and the number of examples are both ranked by the worker's learnability. The stronger the learnability of the worker, the more bonus he will gain and the fewer teaching examples are required. The results also show the bonus disparity between honest workers and dishonest workers becomes larger as the level of teaching target rises.

Effects of Different Priors. We observe the framework for crowd workers who have different priors. Figure 6 (a) depicts the learning dynamics of median learnability workers with different initial proficiency to the task. They all earn competitive bonuses in this fair crowdsourcing and reach very good final performance. As Figure 6 (b) shows, there is no obvious divergence on bonus among honest workers with different initial expertise in general, as long as they put in similar effort to learn in the pedagogical crowdsourcing. For weak learners, having low initial expertise elicits more teaching from the system, therefore they displayed effort is a bit higher. However, for those dishonest workers who initially show strong expertise, the final bonus will be very low since they don't make contributions compatible with their competency. Fewer examples are needed for workers who have higher initial performance.

6 CONCLUSION & FUTURE WORK

In this paper, we propose a novel interactive teaching framework of crowdsourcing to mitigate the integrity and proficiency issues to improve the crowd workers' reliability. We formulate the interactions between the annotation system and the crowd workers as an incentivized pedagogical process between the teacher and the students. The system is able to infer the workers' belief from their current answers, incentive them with improvement-contingent bonuses, and select near-optimal examples for instruction. An efficient algorithm is developed for the annotation system to select examples, even when a worker's belief is unidentifiable. We further show that our framework ensures fair payments to crowd workers regardless of their initial priors and facilitates value-alignment between the task requester and the crowd workers.

Our work hopes to inspire future research to continue exploring *pedagogical crowdsourcing* with combining various machine teaching approaches and incentive mechanisms to improve crowd workers reliability. Current experiments on simulated workers demonstrate the effectiveness and robustness of our approach and suggest its applicability to real-world crowdsourcing tasks. We expect to conduct real deployment and human evaluation in the future.

REFERENCES

[1] Azad Abad, Moin Nabi, and Alessandro Moschitti. 2017. Self-Crowdsourcing Training for Relation Extraction. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017, Vancouver, Canada, July 30 - August 4, Volume 2: Short Papers*. 518–523. https://doi.org/10.18653/v1/P17-2082

[2] David M Burdick, Frederick T Short, and Jaimie Wolf. 1993. An index to assess and monitor the progression of wasting disease in eelgrass Zostera marina. *Marine Ecology Progress Series* (1993), 83–90.

[3] Nilesh Dalvi, Anirban Dasgupta, Ravi Kumar, and Vibhor Rastogi. 2013. Aggregating Crowdsourced Binary Ratings. In *Proceedings of the 22Nd International Conference on World Wide Web (WWW '13)*. ACM, New York, NY, USA, 285–294. https://doi.org/10.1145/2488388.2488414

[4] Jia Deng, Jonathan Krause, and Fei-Fei Li. 2013. Fine-Grained Crowdsourcing for Fine-Grained Recognition. In *2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, June 23-28, 2013*. 580–587. https://doi.org/10.1109/CVPR.2013.81

[5] Shayan Doroudi, Ece Kamar, Emma Brunskill, and Eric Horvitz. 2016. Toward a Learning Science for Complex Crowdsourcing Tasks. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 2623–2634. https://doi.org/10.1145/2858036.2858268

[6] X. Gan, X. Wang, W. Niu, G. Hang, X. Tian, X. Wang, and J. Xu. 2017. Incentivize Multi-Class Crowd Labeling Under Budget Constraint. *IEEE Journal on Selected Areas in Communications* 35, 4 (April 2017), 893–905. https://doi.org/10.1109/JSAC.2017.2680838

[7] Dion Hoe-Lian Goh, Ei Pa Pa Pe-Than, and Chei Sian Lee. 2017. Perceptions of virtual reward systems in crowdsourcing games. *Computers in Human Behavior* 70 (2017), 365–374. https://doi.org/10.1016/j.chb.2017.01.006

[8] Rebecca A. Hutchinson, Liqiang He, and Sarah C. Emerson. 2017. Species Distribution Modeling of Citizen Science Data as a Classification Problem with Class-Conditional Noise. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017, San Francisco, California, USA*. 4516–4523. http://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14859

[9] Ehud Kalai and Ehud Lehrer. 1993. Rational Learning Leads to Nash Equilibrium. *Econometrica* 61, 5 (1993), 1019–1045. http://www.jstor.org/stable/2951492

[10] David R. Karger, Sewoong Oh, and Devavrat Shah. 2011. Iterative Learning for Reliable Crowdsourcing Systems. In *Advances in Neural Information Processing Systems 24*, J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger (Eds.). Curran Associates, Inc., 1953–1961. http://papers.nips.cc/paper/4396-iterative-learning-for-reliable-crowdsourcing-systems.pdf

[11] Adriana Kovashka, Olga Russakovsky, Li Fei-Fei, and Kristen Grauman. 2016. Crowdsourcing in Computer Vision. *Foundations and Trends in Computer Graphics and Vision* 10, 3 (2016), 177–243. https://doi.org/10.1561/0600000071

[12] Edith Law, Ming Yin, Joslin Goh, Kevin Chen, Michael A. Terry, and Krzysztof Z. Gajos. 2016. Curiosity Killed the Cat, but Makes Crowdwork Better. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, May 7-12, 2016*. 4098–4110. https://doi.org/10.1145/2858036.2858144

[13] Mengchen Liu, Liu Jiang, Junlin Liu, Xiting Wang, Jun Zhu, and Shixia Liu. 2017. Improving Learning-from-Crowds through Expert Validation. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*. 2329–2336. https://doi.org/10.24963/ijcai.2017/324

[14] Shengying Pan, Kate Larson, Josh Bradshaw, and Edith Law. 2016. Dynamic Task Allocation Algorithm for Hiring Workers that Learn. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9-15 July 2016*. 3825–3831. http://www.ijcai.org/Abstract/16/538

[15] Carlos Daniel Mimoso Paulino and Carlos Alberto de Bragança Pereira. 1994. On identifiability of parametric statistical models. *Journal of the Italian Statistical Society* 3, 1 (01 Feb 1994), 125–151. https://doi.org/10.1007/BF02589044

[16] Patrick Shafto, Noah D. Goodman, and Thomas L. Griffiths. 2014. A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology* 71, Supplement C (2014), 55 – 89. https://doi.org/10.1016/j.cogpsych.2013.12.004

[17] Nihar Shah and Dengyong Zhou. 2016. No Oops, You Won't Do It Again: Mechanisms for Self-correction in Crowdsourcing. In *Proceedings of The 33rd International Conference on Machine Learning (Proceedings of Machine Learning Research)*, Maria Florina Balcan and Kilian Q. Weinberger (Eds.), Vol. 48. PMLR, New York, New York, USA, 1–10. http://proceedings.mlr.press/v48/shaha16.html

[18] Adish Singla, Ilija Bogunovic, Gábor Bartók, Amin Karbasi, and Andreas Krause. 2014. Near-Optimally Teaching the Crowd to Classify. In *Proceedings of the 31th International Conference on Machine Learning, ICML 2014, Beijing, China, 21-26 June 2014*. 154–162. http://jmlr.org/proceedings/papers/v32/singla14.html

[19] Tian Tian and Jun Zhu. 2015. Max-Margin Majority Voting for Learning from Crowds. In *Advances in Neural Information Processing Systems 28*, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett (Eds.). Curran Associates, Inc., 1621–1629. http://papers.nips.cc/paper/5976-max-margin-majority-voting-for-learning-from-crowds.pdf

[20] Yexiang Xue, Ian Davies, Daniel Fink, Christopher Wood, and Carla P. Gomes. 2016. Avicaching: A Two Stage Game for Bias Reduction in Citizen Science. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems (AAMAS '16)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 776–785. http://dl.acm.org/citation.cfm?id=2937029.2937038

[21] Ming Yin and Yiling Chen. 2015. Bonus or Not? Learn to Reward in Crowdsourcing. In *Proceedings of the 24th International Conference on Artificial Intelligence (IJCAI'15)*. AAAI Press, 201–207. http://dl.acm.org/citation.cfm?id=2832249.2832277

[22] Han Yu, Zhiqi Shen, Chunyan Miao, and Bo An. 2013. A Reputation-aware Decision-making Approach for Improving the Efficiency of Crowdsourcing Systems. In *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems (AAMAS '13)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1315–1316. http://dl.acm.org/citation.cfm?id=2484920.2485201

[23] Omar F. Zaidan and Chris Callison-Burch. 2011. Crowdsourcing Translation: Professional Quality from Non-professionals. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1 (HLT '11)*. Association for Computational Linguistics, Stroudsburg, PA, USA, 1220–1229. http://dl.acm.org/citation.cfm?id=2002472.2002626

[24] Xiaojin Zhu. 2013. Machine Teaching for Bayesian Learners in the Exponential Family. In *Proceedings of 27th Annual Conference on Neural Information Processing Systems, NIPS 2013. Lake Tahoe, Nevada, United States, December 5-8, 2013*. 1905–1913. http://papers.nips.cc/paper/5042-machine-teaching-for-bayesian-learners-in-the-exponential-family