Strong Truthfulness in Multi-Task Peer Prediction

Working paper

Victor Shnayder
edX, Harvard University
shnayder@eecs.harvard.edu

Rafael Frongillo
University of Colorado, Boulder
raf@colorado.edu

Arpit Agarwal
IISc, Bangalore
arpit.agarwal@csa.iisc.ernet.in

David C. Parkes
Paulson School of Engineering
Harvard University
parkes@eecs.harvard.edu

February 9, 2016

Abstract

The problem of peer prediction is to elicit information from agents in settings without any objective ground truth against which to score reports. Peer prediction mechanisms seek to exploit correlations between signals to align incentives with truthful reports. A long-standing concern has been the possibility of uninformative equilibria. For binary signals, the Dasgupta-Ghosh output agreement (OA) mechanism [2] leverages reports across multiple tasks to achieve strong truthfulness, so that the truthful equilibrium maximizes payoff. In this paper, we first characterize conditions on the signal distribution for which the OA mechanism remains strongly-truthful with non-binary signals. Our analysis also yields a greatly simplified proof of their binary-signal, strong truthfulness result.

We then introduce the 01 mechanism, which extends the OA mechanism to multiple signals, with a slightly weaker incentive property: no strategy provides more payoff in equilibrium than truthful reporting, and truthful reporting is strictly better than any uninformed strategy (where an agent avoids the effort of even obtaining a signal). In an analysis of peer-grading data from a large MOOC platform, we investigate how well student reports fit our model, and conclude that the 01 mechanism would be appropriate for use in this domain.

1 Introduction

We study the problem of information elicitation without verification (“peer prediction”). This is a challenging problem across a diverse range of multi-agent systems, in which participants are asked to provide an opinion about an information task, and where there is no objective ground truth available against which to score reports. Examples include completing surveys about the features of new products, providing feedback on the quality of food or the ambience in a restaurant, sharing emotions when watching video content, peer evaluation of assignments in Massive Open Online Courses (MOOCs), and identifying spam emails.

The challenge is to provide incentives so that participants will choose to invest effort in forming an informative opinion (a “signal”) about a task, and choose to make truthful reports about their signals. In the absence of a ground truth, peer-prediction mechanisms make payments to one agent based on the reports of others, and seek to align incentives by leveraging correlation between reports (i.e., peers are rewarded for making reports that are, in some sense, predictive of the reports of others).

Some domains have binary signals, for example “was a restaurant noisy or not?”, and “is an email spam or not?”. We are interested in domains with non-binary signals, for example:

- **Image labeling.** Signals could correspond to answers to questions such as “Is the animal in the picture a dog, a cat or a beaver”, or “Is the emotion expressed joyful, happy, sad or angry.” These signals are categorical, potentially with some structure: ‘joyful’ is closer to ‘happy’ than ‘sad’, for example.
• Counting objects. There could be many possible signals, representing answers to questions such as (“are there zero, 1-5, 6-10, 11-100, or > 100 people in the picture?”). The signals are ordered.

• Peer evaluation in MOOCs. Multiple students evaluate their peers’ submissions to an open-response question using a grading rubric. For example, an essay may be evaluated for clarity, reasoning, and relevance, with the grade for reasoning ranging from 1 (“wild flights of fancy throughout”), through 3 (“each argument is well motivated and logically defended”).

The design of peer-prediction mechanisms assumes the ability to make payments to agents, and that an agent’s utility is linear-increasing with payment and does not depend on signal reports other than through payment. Peer prediction precludes, for example, that an agent may prefer to misreport the quality of a restaurant because she is interested in driving more business to the restaurant. The payments need not be monetary; one could for example issue points to agents, these points conveying some value (e.g., redeemable for awards, or conveying status). On a MOOC platform, the payments could correspond to scores assigned as part of a student’s overall grade in the class. What is needed is a linear relationship between payment (of whatever form) and utility.

We are interested in minimal peer-prediction mechanisms which require only signal reports from participants. More complicated designs have been proposed (e.g. [14]), in which participants are also asked to report their beliefs about the signals that others will report. We believe that minimal peer-prediction mechanisms are considerably more likely to be adopted in practice. It is cumbersome to design user interfaces for reporting beliefs, and people are notoriously bad at reasoning about probabilities.

A basic, desirable property of minimal peer-prediction mechanisms is that of incentive compatibility. This means that truthful reporting of signals is a strict, correlated equilibrium of the game that is induced by the peer-prediction mechanism. Each agent receives a signal about a task, such as the label to associate with an image, and makes a report conditioned on this signal. Making truthful reports is a strict correlated equilibrium in an incentive-compatible mechanism.

Due to the existence of additional equilibria in peer-prediction mechanisms, however, it has long been recognized that incentive compatibility is likely insufficient for practical applications [7, 2, 10, 17]. For example, the equilibria of peer-prediction mechanisms must always include an uninformative, mixed Nash equilibrium [21]. An uninformative equilibrium is one in which the signals reported by agents do not depend on the signals received by agents.

Sometimes there are non-truthful, perhaps uninformative, equilibria that payoff-dominate the truthful equilibrium. With binary signals, a single task, and two agents, Jurca and Faltings [6] show that an incentive-compatible, minimal peer-prediction mechanism will always have an uninformative equilibrium with a higher payoff than truthful reporting. Because of this, it becomes a valid concern that peer-prediction mechanisms may have the paradoxical effect of encouraging agents who would otherwise be truthful to seek higher payment through strategic behavior.

In this light, a result due to Dasgupta and Ghosh [2] is of interest: if agents are asked to provide information on overlapping sets of independent tasks, and for domains with binary signals, there is a mechanism that addresses the problem of multiple equilibria. The binary signal Dasgupta-Ghosh output agreement (OA) mechanism is strongly truthful, meaning that truthful reporting yields the maximum payoff across all equilibria (and is tied in payoff only with strategies that report permutations of true signals; e.g., $1 \rightarrow 2, 2 \rightarrow 3, 3 \rightarrow 1$).

The OA mechanism is constructed from the building block of a score matrix, with a score of ‘1’ for agreement and ‘0’ otherwise. When paired with another agent, the payment on a shared task is the score on this task minus the expected score for reports on a pair of disjoint tasks (perhaps just two other tasks sampled at random). Dasgupta and Ghosh [2] study the mechanism for binary signals, and remark that extending beyond two signals “is one of the most immediate and challenging directions for further work.”

Our main results are as follows:

• We study multi-task, multi-signal mechanisms and provide a tight characterization of when the OA mechanism generalizes to multi-signal domains. The multi-signal OA mechanism is strongly truthful in categorical domains, where receiving one signal reduces an agent’s belief that other agents will receive any other signal. In obtaining this result, we develop a new analysis framework that provides a dramatic simplification of the techniques adopted by Dasgupta and Ghosh [2].

\footnote{It has been more common to refer to the equilibrium concept in peer-prediction as a Bayes-Nash equilibrium in the literature. But more precisely, this is a correlated equilibrium, because there is no agent-specific, private information about payoffs. Rather, every agent is assumed to have a utility function that is linearly increasing in the payment that it receives.}
We use knowledge of the signal distribution to modify the score matrix of the OA mechanism to obtain general, positive results, for the multi-signal, multi-task peer-prediction problem. The 01 mechanism generalizes the OA mechanism and works for general domains, providing a slightly weaker incentive property of informed truthfulness: no strategy provides more payoff in equilibrium than truthful reporting, and truthful reporting is strictly better than any uninformed strategy (where an agent avoids the effort of even obtaining a signal). The informed truthfulness property only requires weak conditions on the signal distribution, and only requires the designer to understand the correlation structure of signals. The 01 mechanism reduces to the OA mechanism in categorical domains, and is informed truthful in other domains where OA is not informed- or strong- truthful.

We provide an empirical analysis of peer-evaluation data of student assignments on a large MOOC platform. We investigate how well signal reports fit our models. We find some surprises in the signal distributions, such as slight negative correlations between evaluations on some questions. The reported signal distribution is categorical on some questions, but there are many questions for which the 01 mechanism is necessary to achieve informed-truthfulness.

These results provide, to the best of our knowledge, the first results on strong truthfulness for minimal peer-prediction mechanisms in domains with non-binary signals that are not stated in the asymptotic limit of many agents performing each task (compare with [17, 9]). The positive results require the existence of multiple, overlapping tasks, each drawn from the same ‘super population’ of tasks. By allowing for some knowledge of the distribution— specifically, it suffices to know whether pairs of signals are positively or negatively correlated—we achieve positive results for very weak assumptions on the domain.

To motivate the informed-truthfulness property, consider the common setting in which it is costly to acquire a signal (e.g., it is costly to carefully examine an image, or costly to read through a student’s submission). In these kinds of domains, it is important that uninformed equilibria provide strictly less payoff than informed equilibria such as the truthful equilibrium. This is provided by the property of informed truthfulness, which also stipulates that the truthful equilibrium achieves at least as much payoff as any other informed equilibrium. In this way, informed truthfulness captures a primary, practical concern, which is to incentivize effort. The advantage for truthful reports over uninformed reports can be easily amplified by scaling all payments.

1.1 Related Work

The theory of peer prediction has developed rapidly in recent years. Beginning with the seminal work of Miller et al. [12], the field has adopted scoring-rule based approaches while relaxing the knowledge requirements on the part of the designer [22, 8], relaxing the requirement of a common prior [23] (see also [15, 18]), and explaining how to handle continuous signal domains [16]. Simple output-agreement, where a positive payment is received if and only if two agents make the same report, and as used in the ESP game [20], has also received some theoretical attention [21, 5].

Jurca and Faltings [7] show how to remove uninformative, pure-strategy Nash equilibria through a clever four-agent prediction game, and several recent papers have started to tackle the existence of uninformative mixed Nash equilibria. One is Dasgupta and Ghosh [2], already described above. Radanovic and Faltings [17] work in a participatory pollution sensing setting, where sensors in a local area are expected to agree on their local pollution level more closely than sensors that are far away. They prove strong truthfulness amongst symmetric strategies in the limit, where both the number of tasks and the number of agents assigned to each task grow without bound. Kamble et. al. [9] also provide a strong truthfulness amongst symmetric strategies result, in their case providing an asymptotic analysis as the number of tasks grows without bound (but only needing a finite number of agents per task.) With binary signals, Kamble et. al. [9] also handle heterogeneity in how participants perceive the world.

A different direction is taken by Cai et al. [1], who elicit data for statistical estimation. A crucial difference is that they assume that, conditioned on effort, an agent will report truthfully. In particular, their mechanism is not informed truthful and appears vulnerable to collusion. In return for this stronger assumption, these authors achieves truthfulness as the unique dominant strategy by taking advantage of properties of the estimators their mechanism supports, and attain optimal statistical estimation properties despite the existence of costly effort.

Experimental evidence also suggests reasons for concern about the multiplicity of equilibria. Gao et al. [4] demonstrate that users in a crowdsourcing environment were able to coordinate on an uninformative equilibrium in many different peer-prediction designs, while behaving in an apparently confused way for a design inspired by Jurca and Faltings [7]. This study also suggested that truthfulness could be achieved by just making the same payment all the
time, without considering reports. We do not believe this is likely to be indicative of other, real-world domains where signals are costly to acquire (the participants in the experiment were provided with their signals).

A second experimental study shows results that go in the other direction, suggesting that in a setting with many possible reports, simple scoring mechanisms can encourage effort, with uninformative equilibria not proving to be a problem [3]. A possible reason for the discrepancy could be that Gao et al. [4] study a binary signal domain while the multiple-signal domain of Faltings et al. [3] may make it harder for agents to coordinate on a useful uninformative strategy.

Research into scaleable online peer evaluation has primarily focused on evaluating students’ skill at assessment and compensating for grader bias [13], as well as helping students self-adjust for bias and provide better feedback [11]. Other studies, such as the Mechanical TA [25], focus on reducing TA workload in high-stakes peer grading. A recent paper [26] outlines an approach to peer evaluation that relies on students flagging overly harsh feedback for instructor review. We are not aware of any systematic studies of peer prediction in the context of MOOCs.

2 Model

We consider two agents, 1 and 2, these agents perhaps members of a larger population. Let $k \in M = \{1, \ldots, m\}$ index a task from a universe of $m \geq 3$ tasks to which one or both of these agents are assigned. Each agent receives a signal when investing effort on an assigned task. The effort model is binary: either an agent invests no effort and does not receive an informed signal, or an agent invests effort and does receive an informed signal, or an agent invests effort and does not.

Let $S_1, S_2$ denote random variables for the signals to agents 1 and 2 on some task. The signals have a finite domain, with $i, j \in \{1, \ldots, n\}$ indexing a realized signal to agents 1 and 2, respectively.

Each task is ex ante identical, meaning that pairs of signals are i.i.d. sampled for each task. Let $P(S_1 = i, S_2 = j)$ denote the joint probability distribution on signals, with marginal probabilities $P(S_1 = i)$ and $P(S_2 = j)$ on the signals of agents 1 and 2, respectively. The signal distribution is common knowledge to agents.

We assume that the distribution satisfies stochastic relevance, so that for all $s' \neq s''$, there exists at least one signal $s$ such that

$$P(S_1 = s|S_2 = s') \neq P(S_1 = s|S_2 = s''),$$

(1)

and symmetrically, for agent 1’s signal affecting the posterior on agent 2’s signal. If two signals $s'$ and $s''$ are not stochastically relevant then they can be combined into one signal.

The OA mechanism is prior-free and does not depend on knowledge of the distribution. The design of the 01 mechanism relies on knowing which signals are positively correlated and which are negatively correlated. In particular, we define:

Definition 1 (Delta matrix). The Delta matrix $\Delta$ is an $n \times n$ matrix, with entry $(i, j)$ defined as

$$\Delta_{ij} = P(S_1 = i, S_2 = j) - P(S_1 = i)P(S_2 = j).$$

(2)

The Delta matrix describes the correlation (positive or negative) between different realized signal values. For example, if $\Delta_{1,2} > 0$ then it means that signal $S_1 = 1$ is positively correlated with signal $S_2 = 2$. In particular, if signals are positively correlated, so that $P(S_1 = i, S_2 = i) > P(S_1 = i)P(S_2 = i)$, then the Delta matrix has a positive diagonal. A row $i$ of $\Delta$ is $P(S_1 = i)$ times the difference between the conditional distribution on $S_2$ and the marginal distribution on $S_2$, and thus each row is a constant multiple of the difference between two distributions, and sums to zero. The same holds for columns.

Example 1. If the signal distribution is

$$P(S_1, S_2) = \begin{bmatrix} 0.4 & 0.15 \\ 0.15 & 0.3 \end{bmatrix}$$

with marginal distribution $P(S) = [0.55; 0.45]$, we get

$$\Delta = \begin{bmatrix} 0.4 & 0.15 \\ 0.15 & 0.3 \end{bmatrix} - \begin{bmatrix} 0.55 \\ 0.45 \end{bmatrix} \cdot \begin{bmatrix} 0.55 & 0.45 \end{bmatrix} \approx \begin{bmatrix} 0.1 & -0.1 \\ -0.1 & 0.1 \end{bmatrix}.$$
The designer needs the sign structure of the $\Delta$ matrix to design the score matrix in the 01 mechanism.

Each agent is assigned to two or more tasks, and the agents overlap on at least one task. Let $M_s \subseteq M$ denote a non-empty set of tasks to which both agents are assigned. Let $M_1 \subseteq M \setminus M_s$ and $M_2 \subseteq M \setminus M_s$, with $M_1 \cap M_2 = \emptyset$ denote non-empty sets of tasks to which agents 1 and 2 are assigned, respectively. For example, if both agents are assigned to each of three tasks, $A, B$ and $C$ then we might designate $M_s = \{A\}$, $M_1 = \{B\}$ and $M_2 = \{C\}$.

A multi-task peer-prediction mechanism takes reports from each agent on each assigned task and defines a total payment to each agent based on these reports. An agent’s strategy defines, for every signal it may receive, and each task it is assigned, the signal it will report. We allow for mixed strategies, so that an agent’s strategy defines a distribution over signals.

Given a set of tasks assigned to an agent, we assume that there is nothing to distinguish two tasks other than their signal, and that the tasks are presented in random order (so that the sequence number conveys no information). In particular, agents have no information about which tasks are shared and which are in set $M$. This is a reasonable assumption in largely anonymous settings, such as peer evaluation and crowdsourcing. It can also be achieved by choosing the designation for $M_s$, $M_1$ and $M_2$ after task assignment.

Our mechanisms’ payments will be computed separately for each shared task, and do not depend on the order of non-shared tasks. This independence, combined with the fact that the agents cannot condition on tasks based on non-signal information, means that it is without loss of generality for agents to adopt strategies that are invariant to the task. Changing a strategy from task to task is equivalent in terms of expected payment to adopting a linear combination over these strategies, given that tasks are presented in a random order, and given that tasks are equivalent, conditioned on signal.

Let $r_1$ and $r_2$ denote the reports by agents 1 and 2, respectively, on some task. Let matrices $F$ and $G$ denote the mixed strategies of agents 1 and 2, respectively, with $F_{ir} = P(r_1 = r | S_1 = i)$ and $G_{jr} = P(r_2 = r | S_2 = j)$ to denote the probability of making report $r$ given signal $i$ is observed (signal $j$ for agent 2). Let $r^k_1 \in \{1, \ldots, n\}$ and $r^k_2 \in \{1, \ldots, n\}$ refer to the report by agent 1 and 2, respectively, on some task $k$ (if assigned).

Definition 2 (Permutation strategy). A permutation strategy is a deterministic strategy in which an agent adopts a bijection between signals and reports, that is, $F$ (or $G$ for agent 2) is a permutation matrix.

Definition 3 (Informed and uninformed strategies). An informed strategy has $F_{ir} \neq G_{jr}$ for some $i \neq j$, some $r \in \{1, \ldots, n\}$ (and similarly for $G$ for agent 2). An uninformed strategy has the same report distribution for all signals.

Definition 4 (Uninformed equilibrium). An uninformative equilibrium is one that is comprised of only uninformed strategies.

The class of multi-task peer-prediction mechanisms that we study in this paper assign a total payment to each agent based on the sum of payments for each task in the shared set, $M_s$, where these payments depend on the reports on the shared task and the reports on non-shared tasks in sets $M_1$ and $M_2$. Moreover, each agent receives the same payment for each task in set $M_s$.

Let $E(F,G)$ denote the expected payment to an agent for any shared task $k \in M_s$. The expectation is taken with respect to the distribution on signals for each task, and any randomization in agent strategies. Let $F^*$ and $G^*$ denote a truthful reporting strategy for agents 1 and 2, respectively.

Definition 5 ((Strictly) Proper). A multi-task peer-prediction mechanism is (strictly) proper if truthful strategies $F^*$ and $G^*$ form a (strict) correlated equilibrium, so that

\[ E(F^*, G^*) \geq E(F, G^*), \]

for all strategies $F \neq F^*$, and similarly when reversing the roles of agents 1 and 2. For strict properness, the inequality must be strict.

This insists that the expected payment on a shared task is (strictly) higher when reporting truthfully than by deviating to some other strategy, given that the other agent is truthful.

Definition 6 (Strong-truthful). For truthful strategies $F^*$ and $G^*$,

\[ E(F^*, G^*) \geq E(F, G), \]

for all $F, G$, and equality only occurs when $F$ and $G$ are both the same permutation strategy.
In words, a multi-task peer-prediction mechanism is strong truthful if all agents being truthful yields strictly greater expected payment than any other strategy profile, excepting for both agents playing the same permutation strategy (which can provide the same expected payment).

Permutation strategies seems unlikely to be a practical concern, since permutation strategies require coordination and provide no benefit over being truthful (truthful reports are more focal, in this sense, among informed strategies than permutation strategies).

**Lemma 1.** Any strong-truthful mechanism is strictly proper.

*Proof.* We need to show

\[ E(F^*, G^*) > E(F, G^*), \]

for truthful strategies \( F^*, G^* \), and non-truthful strategy \( F \). This follows directly from the condition for strong truthful, since \( F \neq G^* \).

**Definition 7** (Informed-truthful). For truthful strategies \( F^* \) and \( G^* \),

\[ E(F^*, G^*) \geq E(F, G), \tag{5} \]

for all \( F, G \), and equality may only occur when \( F \) and \( G \) are informed strategies.

In words, a multi-task peer-prediction mechanism is informed-truthful if and only if the truthful strategy profile has strictly higher expected payment than any profile in which one or both agents play an uninformed strategy, and weakly greater expected payment than all other strategy profiles.

**Lemma 2.** Any informed-truthful mechanism is proper.

*Proof.* We need to show

\[ E(F^*, G^*) \geq E(F, G^*), \]

for truthful strategies \( F^*, G^* \), and non-truthful strategy \( F \). This follows directly from the condition for informed truthful.

Although weaker than the property of strong-truthfulness, this property is responsive to what we consider to be the primary, practical concern in peer-prediction applications— namely, equilibria where agents can avoid the effort of forming a careful opinion about an information task but still achieve the same (or greater) expected payment than is possible from following an informed strategy. For example, it would be undesirable for agents to be able to do just as well (or better) by reporting the same signal all the time. Informed-truthful mechanisms motivate effort without introducing additional, undesirable and even collusive behaviors.

Once agents exert effort and observe a signal, it seems reasonable to expect them to make truthful reports as long as this is an equilibrium and as long as there is no other equilibrium that provides a higher expected payment. This is what is enabled through an informed-truthful peer-prediction mechanism.

It is a straightforward extension of the present model to model the cost of effort explicitly, as in Dasgupta and Ghosh [2]. An agent that does not exert effort receives an expected payment of zero in our mechanisms, while the payment for agents that exert effort and play the truthful equilibrium is positive. With knowledge of the maximum possible cost of effort, scaling the payments so that expected payoff is greater than this cost would provide incentives for agents to exert effort.

### 3 Generalized-Score, Multi-Task Peer-Prediction Mechanisms

We define a class of multi-task peer-prediction mechanisms that is parametrized by a score matrix, \( S : \{1, \ldots, n\} \times \{1, \ldots, n\} \to \mathbb{R} \), that map a pair of reports into a score, the same score for both agents. This mechanism class extends the binary-signal, Dasgupta-Ghosh (OA) mechanism [2] in a natural way.

**Definition 8** (The Generalized-Score, Multi-signal mechanism (GS)).
1. Assign the agents to three or more tasks, with each agent to two or more tasks, including at least one overlapping task. Let $M_s$, $M_1$ and $M_2$ denote the shared, agent-1 and agent-2 tasks.

2. Let $r^k_i$ denote the report received from agent 1 on task $k$ (and similarly for agent 2). The payment to each of agent 1 and agent 2 for a shared task $k \in M_s$ is

$$S(r^k_1, r^k_2) - \sum_{i=1}^{n} \sum_{j=1}^{n} S(i, j) \cdot h_{1,i} \cdot h_{2,j},$$

where $S: \{1, \ldots, n\} \times \{1, \ldots, n\} \rightarrow \mathbb{R}$ is the score matrix, $h_{1,i} = \frac{|\{\ell \in M_1 | r^\ell_1 = i\}|}{|M_1|}$ is the empirical frequency with which agent 1 reports signal $i$ in tasks in set $M_1$, and $h_{2,j} = \frac{|\{\ell \in M_2 | r^\ell_2 = j\}|}{|M_2|}$ is the empirical frequency with which agent 2 reports signal $j$ in tasks in set $M_2$.

3. The total payment to an agent is the sum total payment across all shared tasks.

As explained earlier, the assignment to tasks could be the result of assigning all agents to all tasks and then designating some tasks as shared, some as agent-1 tasks and some as agent-2 tasks. Alternatively, this could be the result of partitioning the tasks into shared, 1-only and 2-only tasks.

The GS mechanism computes the payment to an agent on a shared task based on the score for their reports on the shared task, subtracting the expected payment from reports on non-shared tasks. An alternate mechanism with the same expected payoff and the same incentive analysis is to pick a random $1 \in M_1$ and $n \in M_2$, and pay $S(r^k_1, r^k_2) - S(r^k_1, r^k_m)$ for shared task $k$. This is somewhat simpler, but has higher variance.

In this way, the mechanism rewards only the excess agreement that arises from both agents observing the same task and not the default agreement that they could get by, for example, both making the same report all the time. The effect is that expected payment is zero for example if agents make random reports, or always make the same report, or when one agent is truthful and the other reports randomly.

We can now define the obvious extension of the OA mechanism to multiple signals, as a special case of GS with the identity score matrix (‘1’ for agreement, ‘0’ for disagreement.)

**Definition 9** (The Multi-Signal Dasgupta-Ghosh mechanism (OA)). Use the GS mechanism, with

$$S(i, j) = \begin{cases} 
1 & \text{if } i = j \\
0 & \text{if } i \neq j
\end{cases}$$

In this special case, the expected payment on a shared task is equal to the difference between the probability that the reports made by each agent on this task will agree on some signal and the probability that the reports made by each agent on disjoint tasks will agree.

**Example 2.** Suppose agent 1 is assigned to tasks $\{A, B\}$ and agent 2 to tasks $\{B, C, D\}$, so that $M_s = \{B\}$, $M_1 = \{A\}$ and $M_2 = \{C, D\}$. Now, if the reports on $B$ are both 1, and the reports on $A, C$, and $D$ were 0, 0, and 1, respectively, the payment to each agent for shared task $B$ is $1 - (1 \cdot 0.5 + 0 \cdot 0.5) = 0.5$. In contrast, if both agents use an uninformed coordinating strategy and always report 1, the score for both is $1 - (1 \cdot 0.5 + 1 \cdot 0.5) = 0$.

Leveraging the Delta matrix, the expected payment on a shared task for the identity score matrix is

$$E(F, G) = \sum_{i=1}^{n} \sum_{j=1}^{n} P(S_1 = i, S_2 = j) \sum_{r=1}^{n} P(r_1 = r | S_1 = i)P(r_2 = r | S_2 = j)$$

$$- \sum_{i=1}^{n} \sum_{j=1}^{n} P(S_1 = i)P(S_2 = j) \sum_{r=1}^{n} P(r_1 = r | S_1 = i)P(r_2 = r | S_2 = j)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \Delta_{ij} \sum_{r=1}^{n} F_{rr}G_{jr}.$$  

The expected payment for a shared task can also be written as $\text{tr}(F^T \Delta G)$. The expected payment is the same for both agents.
In the general case, the score becomes the difference in expected payment between shared tasks and disjoint tasks, rather than the probability of agreement on the shared tasks minus the probability of agreement on disjoint tasks. The expression for the expected payment becomes

$$E(F, G) = \sum_{i=1}^{n} \sum_{j=1}^{n} \Delta_{ij} \sum_{r_1=1}^{n} \sum_{r_2=1}^{n} S(r_1, r_2) F_{ir_1} G_{jr_2}. \quad (10)$$

This generalized expression for the expected payment can also be written as

$$E(F, G) = \text{tr}(F^\top \Delta G S G^\top). \quad (11)$$

In words, the expected payment on a shared task is the sum, over all pairs of possible signals, of the product of the correlation (negative or positive) between those signals, and the expected payment given this signal pair and given the agent strategies. For the special case of the identity score matrix, the expected payment for a given signal pair is just the “excess” probability that the agents will make the same report on a shared task over an unshared task.

The matrix $\Delta$ is a property of the domain, and does not depend on agent strategies. This will make it a convenient object for reasoning about the incentive properties of the GS mechanism. The structure of (10) shows that agents are trying, through the use of strategies $F$ and $G$, to “pick out” a high score when $\Delta_{ij}$ is positive and a low score when $\Delta_{ij}$ is negative.

**Lemma 3.** For any world model, there exists a deterministic, optimal joint strategy for GS.

**Proof.** We argue based on convex optimization. The game value can be written $V = \max_F \max_G h(F, G)$, where

$$h(F, G) = \sum_{i=1}^{n} \sum_{j=1}^{n} \Delta_{ij} \sum_{r_1=1}^{n} \sum_{r_2=1}^{n} S(r_1, r_2) F_{ir_1} G_{jr_2}. \quad (12)$$

Note that $h$ is linear in both $F$ and $G$ separately. Now letting $V(F) = \max_G h(F, G)$ be the value for the $G$ player for a fixed $F$, we have $V = \max_F V(F)$ by definition. As $h(F, \cdot)$ is linear, and the strategy space for $G$, all binary row-stochastic matrices, is convex, there exists a maximizer at an extreme point. These extreme points are exactly the deterministic strategies, and thus for all $F$ there exists an optimal $G = G^{\text{opt}}$ which is deterministic.

Now considering the maximization over $F$, we see that $V(F) = \max_G h(F, G)$ is a pointwise supremum over a set of linear functions, and is thus convex. $V$ is therefore optimized by an extreme point, some deterministic $F = F^{\text{opt}}$, and for that $F^{\text{opt}}$ there exists a corresponding deterministic $G^{\text{opt}}$ by the above.

**Lemma 3** has several consequences:

- It is without loss of generality to focus on deterministic strategies when establishing strong truthful or informed truthful properties of a mechanism.

- There is a deterministic, perhaps asymmetric equilibrium, because the optimal solution that maximizes $E(F, G)$ is also an equilibrium.

- It is without loss of generality to consider deterministic deviations when checking whether or not truthful play is an equilibrium.

We will henceforth assume deterministic strategies. By a slight abuse of notation, let $F_i \in \{1, \ldots, n\}$ and $G_j \in \{1, \ldots, n\}$ denote the reported signals by agent 1 for signal $i$ and agent 2 for signal $j$, respectively. The expected score then simplifies to

$$E(F, G) = \sum_{i=1}^{n} \sum_{j=1}^{n} \Delta_{ij} S(F_i, G_j). \quad (12)$$
4 The Multi-signal OA Mechanism

This section explores the equilibria of the OA mechanism. We will show that if the signal distribution is categorical then the identity score matrix makes GS strongly truthful in multi-signal domains.

**Definition 10 (Categorical model).** A world model is categorical if, when an agent sees a signal, all other signals become less likely than their prior probability; i.e., \( P(S_2 = j|S_1 = i) < P(S_2 = j) \), for all \( i \), for all \( j \neq i \) (and analogously for agent 2).

Categorical is equivalent to the Delta matrix having positive diagonal and negative off-diagonal elements.

**Theorem 1.** If the world is categorical, then mechanism OA is strong truthful and strictly proper. Conversely, if the Delta matrix \( \Delta \) is symmetric and the world is not categorical, then OA is not strong truthful.

**Proof.** First, we show that truthfulness maximizes the expected payment. Recall the formula for the expected payment, which, given Lemma 3, we now write for deterministic strategies only:

\[
E(F, G) = \sum_{i,j} \Delta_{ij} [F_i = G_j]
\]

The truthful strategy corresponds to the identity matrix \( I \), and results in a payment equal to the trace of \( \Delta \): \( E(I, I) = \text{tr}(\Delta) = \sum_{i,j} \Delta_{ii} \). By the categorical assumption, \( \Delta \) has positive diagonal and negative off-diagonal elements, so this is the sum of all the positive elements of \( \Delta \). Because \( I[F_i = G_j] \leq 1 \), this is the maximum possible payment for any pair of strategies.

To show strong truthfulness, consider asymmetric joint strategy, with \( F \neq G \). Then there exists \( i \) s.t. \( F_i \neq G_i \), reducing the expected payment by at least \( \Delta_{ii} > 0 \). Now consider symmetric, non-permutation strategies \( F = G \). Then there exist \( i \neq j \) with \( F_i = F_j \). The expected payment will then include \( \Delta_{ij} < 0 \). This shows that truthfulness and symmetric permutation strategies are the only optimal strategy profiles. Strict properness follows by Lemma 1.

For the tightness of the categorical assumption, first consider a symmetric \( \Delta \) with positive off-diagonal elements \( \Delta_{ij} \) and \( \Delta_{ji} \). Then agents can benefit by both “merging” signals \( i \) and \( j \). Let \( F \) be the strategy that is truthful on all signals other than \( j \), and reports \( i \) when the signal is \( j \). Then \( E(F, F) = \Delta_{ij} + \Delta_{ji} + \text{tr}(\Delta) > E(I, I) \), so OA is not strong truthful. Now consider a \( \Delta \) where one of the on-diagonal entries is negative, say \( \Delta_{ii} < 0 \). Then, because rows and columns add to 0, there must be a \( j \) such that \( \Delta_{ij} > 0 \), and this reduces to the previous case where “merging” \( i \) and \( j \) is useful.

For binary signals, any positively correlated model, such that \( \Delta_{1,1} > 0 \) and \( \Delta_{2,2} > 0 \) (for signals ‘1’ and ‘2’), is categorical and we obtain much simpler proof for the main result in Dasgupta and Ghosh [2].

4.1 Discussion

Which world models are categorical? One example is a noisy observation model, where each agent observes the “correct” signal \( t \) with with probability \( q \) greater than \( 1/n \), and otherwise makes a mistake uniformly at random, receiving any signal \( s \neq t \) with probability \((1-q)/(n-1)\). We might expect this property to hold for example for classification tasks in which the classes are fairly distinct. For example, this categorical property would likely hold if the question is “Does the animal in this photo swim, fly, or walk?” On the other hand, a classification problem such as the ImageNet challenge [19], with 1000 nuanced and often similar image labels, is unlikely to be categorical. For example, if “Ape” and “Monkey” are possible signals, one agent seeing “Ape” is likely to increase the probability that another says “Monkey”, when compared to the prior for “Monkey” in a generic set of photos. The categorical property is also unlikely to hold in settings where signals have a natural order; e.g., when grading a homework on a 1 to 100 scale, if one agent’s signal is 74, this increases the likelihood that another agent’s signal would be 73 or 75.

The categorical property is stronger than other properties that have been assumed for signal distributions in the literature, such as the domain properties assumed in the analysis of the Jurca-Faltings [8] “1/prior” mechanism and the Witkowski-Parkes [22] shadowing mechanism. The 1/prior mechanism requires condition \( P(S_2 = j|S_1 = i) < P(S_2 = j) \) on the signal distribution, whereas the categorical property insists on an upper bound of \( P(S_2 = j) \), which is smaller than \( P(S_2 = j|S_1 = i) \) by positive correlation. The shadowing mechanism requires condition \( P(S_2 = i|S_1 = j) - P(S_2 = i) < P(S_2 = j|S_1 = j) - P(S_2 = j) \), which says that the likelihood of signal \( S_2 = i \) cannot go up “too much” given signal \( S_1 = j \), whereas the categorical property requires that \( P(S_2 = i|S_1 = j) - P(S_2 = i) < 0 \).
5 Handling the General Case

We have established that the OA mechanism is strongly truthful for a categorical domain but not for non-categorical domains. In this section, we present an extension that achieves informed-truthfulness for general domains.

5.1 The 01 Mechanism

Definition 11 (01 mechanism). The 01 mechanism is a GS mechanism with score matrix

\[
S(i, j) = \begin{cases} 
1 & \text{if } \Delta_{i,j} > 0 \\
0 & \text{else}
\end{cases}
\]

Theorem 2. The 01 mechanism is informed-truthful and proper for all models.

Proof. The truthful strategy \(F^*, G^*\) has higher payment than any other pair \(F, G\):

\[
E(F^*, G^*) = \sum_{i,j} \Delta_{i,j} S_{i,j} \geq \sum_{i,j} \Delta_{i,j} S_{F^i,G^j} = E(F, G),
\]

where the inequality follows from the fact that \(S_{i,j} \in \{0, 1\}\).

The truthful score is positive, while any uninformed strategy has score zero. Consider an uninformed strategy \(F\), with \(F_i = r\) for all \(i\). Then, for any \(G\),

\[
E(F, G) = \sum_i \sum_j \Delta_{i,j} S_{r,G_j} = \sum_j S_{r,G_j} \sum_i \Delta_{i,j} = \sum_j S_{r,G_j} \cdot 0 = 0,
\]

where the next-to-last equality follows because rows and columns of \(\Delta\) sum to zero, as described in Section 2.

Example 3. Let’s look at a setting where the OA mechanism is manipulable, and see how 01 fixes this. If two evaluators grade essays on a scale from one to one hundred, they are more likely to agree on a range (e.g. 61-70) than a particular number (e.g. 62 or 63). In OA, they would be better off agreeing on such ranges, and always reporting a specific number (e.g. 55, 65, 75, etc) per range. With the 01 mechanism, they will get a reward without this manipulation, including when their two signals fall in adjacent ranges (e.g. one graded as 58, one as 61).

The 01 mechanism considers signals that are positively correlated on shared tasks (and thus have a positive entry in \(\Delta\)) to be matching, so there is no need to agents to misreport to ensure matching. In simple cases, e.g. a three signal model where signals 1 and 2 are positively correlated, this results in them being “merged,” and reports of one treated equivalently to the other.

6 An Empirical Analysis of Peer-assessment Data

We now turn to a dataset from a large MOOC provider, focusing on 104 questions with over 100 submissions each, for a total of 325,523 assessments from 17 courses. Each assessment consists of a numerical store, which we examine, and an optional comment, which we do not study here. As an example, one assessment for a writing assignment asks how well the student presented their ideas, with options “Not much of a style at all”, “Communicative style”, and “Strong, flowing writing style”, and a paragraph of detailed explanation for each. These correspond to 0, 1, and 2 points on this rubric element.

While we only see student reports, we believe that they reasonably approximate the true world model. Students are not graded on their evaluations, but do need to complete them to see the feedback on their own submission. Although we estimate signal distributions for these reported assessments, and consider them representative of true models, our motivation for designing appropriate incentives remains. We believe that as MOOCs develop along with valuable credentials based on their peer-assessed work, it will become increasingly important to provide explicit credit mechanisms for peer assessment.

We estimate \(\Delta\) matrices from the evaluations on shared and non-shared tasks. The questions in our data set had five or fewer rubric options (signals), with three being most common (Figure 1). We found that our positive correlation assumption holds in all but two cases (Figure 1L).
Figure 1: Left: models with positive correlation, by size. All but two satisfy the positive correlation condition. Right: MOOC peer evaluation is an ordinal scoring setting, so most models with 3 or more signals are not categorical.

![Average Δ matrices](image)

Figure 2: Averaged Δs, by size. The positive diagonals show that users tend to agree on their evaluations. For models of size 4 and 5, the ordinal nature of peer assessment is clear: e.g. an evaluation of 2/5 is positively correlated with another agent reporting 3/5.

Given that peer evaluation is an ordinal task, we were not surprised that the categorical condition only holds for about one third of our three-signal models, and not for larger models (Figure 1R). We computed the average Δ matrix for each model size, as visualized in Figure 2. The bands of positive correlation around the diagonal is typical of an ordinal domain. This suggests that the 01 mechanism would be useful in providing better incentive properties than the OA mechanism — they are equivalent for the categorical models, and provide strong truthfulness in those cases, with the 01 mechanism additionally non-manipulable in the non-categorical cases.

7 Conclusion

We study the design of peer-prediction mechanisms that leverage signal reports on multiple tasks to ensure that truthful reporting is the joint strategy with highest payoff across all joint strategies, and strictly higher payoff than all uninformed strategies (i.e., those that do not involve effort and do not depend on signals). Attaining our positive result, we extend an earlier mechanism due to Dasgupta and Ghosh [2] to multi-signal settings. The immediate extension using an identity score matrix works in domains with a categorical structure, providing strong truthfulness. We generalize this approach to the 01 mechanism, which handles general domains while still providing a useful claim in regard to the optimality of truthful reporting (i.e., the informed truthful property). The mechanism reduces to the OA mechanism and strong truthfulness in categorical domains.

One potential shortcoming of our analysis is that informed-truthfulness depends on a binary notion of effort— if it is easier to get an approximate signal (e.g. “this is either a 1 or a 2”) than a precise one, and user know which distinctions matter to the mechanism, they would not need to invest effort in getting fully precise signals.

Directions for future work include characterizing domains where the signal distribution is categorical, as well as focusing on specific non-categorical domains such as ordinal signals. Another question is whether these model-based approaches can be made model-free via learning by the center, while still achieving strong alignment of incentives (c.f., Witkowski and Parkes [24] in the classical setting for peer prediction). Finally, we would like to evaluate these mechanisms in real peer prediction applications.
References


