

BAYESIAN PERSUASION IS A BARGAINING GAME

Anonymous authors

Paper under double-blind review

ABSTRACT

Bayesian persuasion studies how a sender with an informational advantage can persuade a receiver with a different motive to take actions that benefit the sender. This problem is previously formulated from an equilibrium perspective, where the sender is to choose a Bayes correlated equilibrium and the receiver is willing to respect the signaling scheme based on posterior beliefs. However, evidence in real-world scenarios and studies in farsighted receivers suggest otherwise: senders tend to be much more honest than the equilibrium. In this work, we show that Bayesian persuasion is reducible to a bargaining game. This reduction suggests that the receiver in Bayesian persuasion can be aware of the game structure and can develop an anti-exploitation strategy. This equalizes the power of commitment of the two parties and prevents the sender from taking the maximum possible payoff. Through experiments on large language models, we demonstrate the receiver’s retaliatory strategies and the sender’s compromise to that. More findings on the impact of the context and alignments further suggest that bargaining behavior emerges in persuasion tasks. The insights given by our results have potential implications on various scenarios to reduce exploitation, improve equality, and improve social welfare.

1 INTRODUCTION

Communication plays a pivotal role in human society, shaping interactions and influencing decision-making processes. The concept of “cheap talk,” introduced by Crawford & Sobel (1982), underscores the importance of verbal communication in strategic contexts. McCloskey’s assertion that “one quarter of GDP is persuasion” illustrates how integral persuasion is to economic activities (McCloskey & Klamer, 1995). Recent analyses by Antioch (2013) suggest that this figure has risen to 30%, highlighting the growing significance of persuasive communication in the modern economy. In this landscape, the idea of Bayesian persuasion, proposed by Kamenica & Gentzkow (2011), emerges as a crucial framework, involving the transmission of signals to influence a receiver’s beliefs and actions. This framework finds a wide range of applications across various domains, including routing systems (Das et al., 2017; Kremer et al., 2014), law enforcement deployment (Hernández & Neeman, 2022; Lazear, 2006), and grading in schools (Boleslavsky & Cotton, 2015; Ostrovsky & Schwarz, 2010), as surveyed by Kamenica (2019).

A series of studies through the lens of Bayesian Correlated Equilibrium (BCE) demonstrates the high relevance of Bayesian persuasion. According to Bergemann & Morris (2013) and Bergemann & Morris (2016), the Bayesian persuasion task can be viewed as a scenario where the sender selects and commits an optimal BCE for itself (Bergemann & Morris, 2019). A signaling scheme qualifies as a BCE if it is incentive compatible (or obedient), meaning a rational receiver will follow the indications of this signaling scheme. As noted by Dughmi (2017), the commitment step is essential to Bayesian persuasion, otherwise the model will degenerate to the cheap talk model.

While the analysis of BCE is elegant, the sender’s signaling scheme at equilibrium may be regarded as deceptive and exploitative from the common sense perspective. In fact, the senders in Bayesian persuasion are honestly dishonest (Best & Quigley, 2016). In human society, such sender might not be tolerated with partially lying in long-term interactions, let alone committing to such exploitative strategies. The experiments in Haran & Shalvi (2020) indicate “honest advice is more persuasive than highly informed advice” for humans. Priester & Petty (1995) suggests that cognitive misers are more likely to accept when the sender is perceived as relatively honest.

054 While the commitment assumption is motivated by long-term incentives, studies in farsighted play-
 055 ers suggest different outcomes than the equilibrium. In the context of repeated games, Gan et al.
 056 (2022) demonstrate that persuading a far-sighted receiver, where players engage in an MDP, is NP-
 057 hard. They restrict their attention to consider an advice-myopic receiver who acts as if the sender
 058 will not be present in the future, and bases decisions solely on prior beliefs. They also introduce a
 059 threat-based signaling scheme for the sender, employing a grim-trigger strategy (Friedman, 1971)
 060 where the sender stops providing information if the receiver deviates. Extending this framework,
 061 Bernasconi et al. (2024) demonstrate that Markovian signaling schemes are suboptimal in this con-
 062 text, and propose a subset of history-dependent signaling schemes called promise-form. This ap-
 063 proach encodes past interactions as honest promises about receiver’s future rewards. These behav-
 064 iors are evidence that bargain behavior may emerge in persuasion tasks.

065 In this work, we show that the receiver can extend its policy space for more complex action rules,
 066 which could force the sender to provide more information. In fact, because the receiver is capable
 067 of obtaining the expected values of both parties in the persuasion task, they could develop anti-
 068 exploitative action rules through this awareness of the game structure. Having the action rules
 069 deployed demonstrates the power of commitment, where the party that committed first will have
 070 the power to choose among the subgame perfect equilibria. Since the player who commits first
 071 has an advantage, the sender and receiver reach a consensus on their joint strategy pair to commit
 072 together before the game begins. This requires them to agree on the expected payoff induced by the
 073 strategy pair. Following this, we show that solving a Bayesian persuasion task needs the solution of
 074 a bargaining game. Therefore Bayesian persuasion is reducible to a bargaining game.

075 We conduct experiments with LLMs on several bargaining and persuasion games. The agents
 076 demonstrate surprisingly similar dynamics and outcomes in persuasion and bargaining. In particu-
 077 lar, we observe the receiver’s retaliation strategy for anti-exploitation and the sender’s compromise
 078 to that. We also demonstrate that mathematically equivalent scenarios could result in different out-
 079 comes given different contexts. This could be explained by Bayesian theory of mind (Baker, 2012)
 080 where different prior (e.g. prior on honesty and retaliation) could impact the game. Our results give
 081 a thorough understanding of persuasion, both to its mathematical formulation and its real-life real-
 082 ization. It provides insights into how receivers could have anti-exploitation and counter-exploitation
 083 strategies, and has the potential to improve equality and social welfare in a variety of scenarios.

084 2 PRELIMINARIES

085
 086 In this section, we give a brief introduction to Bayesian persuasion and bargaining. Because there is
 087 a large number of notations, we summarize them in the glossary in Appendix A.

088 2.1 BAYESIAN PERSUASION

089
 090
 091 Bayesian persuasion is an one-shot two-player game where a sender i is trying to send signals
 092 $\sigma \in \Sigma$ to influence the actions $a \in \mathbb{A}$ of a receiver j . We initially focus on an imperfect and
 093 incomplete information game (as described in Procedure 1), for it being a general extension of
 094 Bayesian persuasion. Although the procedure is linear and contains no loops or branches, there are
 095 many details and assumptions to pay attention to.

096 A neutral environment will first samples a state $s \in \mathbb{S}$ according to a prior distribution $\mu_0 \in \Delta(\mathbb{S})$.
 097 State s is observable only for the sender, and it influence both players’ payoffs. So the receiver has
 098 interests in state s without knowing it, which gives an opportunity to the sender to influence the
 099 behavior of the receiver. This refers to the sender’s informational advantage. Then the environment
 100 samples a type¹ $\tau \sim q(\cdot | s)$. Types $\tau \in \mathbb{T}$ are the receiver’s partial observations. It can be seen by
 101 both the sender and receiver, while in the multi-receiver setting, it cannot be seen by other receivers.
 102 The canonical Bayesian persuasion settings assume that the receiver has no private information, so
 103 $\mathbb{T} = \{0\}$. It means that the emission function q maps all the state to the same type $\tau = 0$.

104 Next, the sender sends a signal $\sigma \in \Sigma$ to the receiver. Its signaling scheme is defined as $\tilde{\varphi} : \mathbb{S} \rightarrow$
 105 $\Delta(\Sigma)$, and the signaling scheme set is $\tilde{\Psi}$. Then, the receiver takes an action $a \in \mathbb{A}$ given the signal.

106
 107 ¹It is also named “signal” in Bergemann & Morris (2016). But ϕ is determined by the environment. The sender can only pick q . To avoid ambiguity we use the term “type” here.

Its action rule is defined as $\tilde{\pi} : \Sigma \rightarrow \mathbb{A}$, and the action rule set is $\tilde{\Pi}$. The receiver’s behavior is default to be Bayesian, which gives the task name “Bayesian persuasion”. It implements its policy $\tilde{\pi}$ (Line 6 in Procedure 1) like: **(1)** It guesses a state given the sent signal according to the Bayesian decision theory Hart et al. (2000); **(2)** It chooses the best response to its guessed state. The second step is assumed to be optimal. The sender’s influence on the receiver acts only on the first step.

The payoff functions of the sender and receiver are defined as $r^i : \mathbb{S} \times \mathbb{A} \rightarrow \mathbb{R}$ and $r^j : \mathbb{S} \times \mathbb{A} \rightarrow \mathbb{R}$, respectively. They do not generally share the same payoff function, so the task is mixed-motive at the most time. Any payoff instance is independent to the sender’s strategy value σ , so for the sender to get better payoffs, it must leverage its informational advantage to manipulate the receiver to act towards its (the sender’s) favor.

Procedure 1: Timing of the Imperfect and Incomplete Information Game

Input: Prior distribution μ_0 , state set \mathbb{S} , the receiver’s action set \mathbb{A} , reward functions r^i and r^j

- 1 Given μ_0 , the sender decides a signaling scheme $\tilde{\varphi}$; // Commitment (1)
- 2 Sender commits to $\tilde{\varphi}$ to the receiver; // Commitment (2)
- 3 Environment samples a state $s \sim \mu_0(\cdot)$;
- 4 Environment samples a type $\tau \sim q(\cdot | s)$; // Dummy in Bayesian persuasion
- 5 The sender recommends an action $\sigma \sim \phi(\cdot | s, \tau)$ where $\sigma \in \mathbb{A}$; // Commitment (3) & Revelation
- 6 The receiver chooses an action $a \sim \tilde{\pi}(\cdot | \sigma)$; // Default: Bayesian best response
- 7 The sender and receiver get rewards of $r^i(s, a)$ and $r^j(s, a)$ respectively;

As stated, the receiver needs to make a Bayesian best response, which requires knowing the sender’s signaling scheme $\tilde{\varphi}$. This is allowed in Bayesian persuasion and is referred to as **the commitment assumption**. This is a signature key assumption, which is indispensable in Bayesian persuasion problems; without this assumption, Bayesian persuasion would degenerate into cheap talk (Crawford & Sobel, 1982; Kamenica & Gentzkow, 2011).

Assumption 2.1 (Commitment). *The sender commits to a signaling scheme $\tilde{\varphi}$, which encompasses the following three key aspects: (1) The sender will decide on a $\tilde{\varphi}$ before the game starts, (2) The sender will honestly inform the receiver of this $\tilde{\varphi}$, and (3) during the game, the sender will actually samples signals $\sigma \in \Sigma$ according to the committed $\tilde{\varphi}$.*

The reasoning behind this assumption is to simplify the analysis of the problem. Its justification is that the sender and receiver will interact over the long term, and considering its own reputation, the sender will not lie about the committed signaling scheme.

We can further simplify the problem without loss of generality by assuming that the sender “sending a signal” is equivalent to “recommending an action” for the receiver to take. And this is known as an analysis similar to the revelation principle, proposed and proved in Kamenica & Gentzkow (2011).

Proposition 2.2 (A Variant of Revelation Principle², Paraphrased from Dughmi (2017)). *Assuming that the signaling set is equal to the action set ($\Sigma = \mathbb{A}$) does not affect the optimality of the sender’s signaling scheme $\tilde{\varphi}$. This implies that every signaling scheme is equivalent to one that recommends actions.*

Let \mathbb{V}^i and \mathbb{V}^j denote the visibility set of the sender and receiver respectively. Let \mathbb{V}^{i+j} denote the common knowledge set (known by the players, with each aware that others know it, ad infinitum). Then $\mathbb{V}^i \setminus \mathbb{V}^j = \{s\}$ and $i, j, \mu_0, \mathbb{S}, \mathbb{A}, r^j, \tilde{\varphi}, \sigma, a \in \mathbb{V}^{i+j}$. The visibility of r^i is not specified in previous work. Here we assume $r^i \in \mathbb{V}^{i+j}$.

To further characterize the task so that the solutions to become conceivable, we restrict our attention to a subset of $\tilde{\Phi}$ whose elements are *signaling schemes that the receiver would respect*. This can be clearly introduced by the concept of **incentive compatibility**.

²We will refer to this as the revelation principle in the following content since there will be no ambiguity here. And we keep using the notation σ to denote an action sent by the sender, using a to denote an action actually taken by the receiver.

Definition 2.3 (Incentive Compatibility in Bayesian Persuasion, Single-Receiver Version of Bergemann & Morris (2016)). *The sender’s recommendation ϕ is incentive compatible for the receiver if for each $\tau \in \mathbb{T}$ and $a \in \mathbb{A}$,*

$$\sum_s \mu_0(s) \cdot q(\tau | s) \cdot \phi(a | s, \tau) \cdot r^j(s, a) \geq \sum_s \mu_0(s) \cdot q(\tau | s) \cdot \phi(a | s, \tau) \cdot r^j(s, a'), \quad (1)$$

for all $a' \in \mathbb{A}$.

Every signal (recommended action) will induce a posterior belief of the receiver, who will then choose the Bayesian best response to it. If the receiver’s best response is exactly the action recommended by the sender, then the receiver is considered persuaded by the sender. In this case, the receiver has no incentives to deviate from the sender’s recommendation, i.e., $\tilde{\pi}(a | \sigma = a) = 1$, and thus can be omitted in this canonical setting. In the origin paper Bergemann & Morris (2016), it is named as the “obedience”. **Bayes correlated equilibria (BCE)** is defined as the set of obedient signaling schemes.

Since we assume $\mathbb{T} = \{0\}$ in this work, $\phi(a | s, \tau) \equiv \tilde{\varphi}(\sigma = a | s)$, and the inequality can be simplified as $\sum_s \mu_0(s) \cdot \tilde{\varphi}(a | s) \cdot [r^j(s, a) - r^j(s, a')] \geq 0$. This is indeed the formulation in the survey Dughmi (2017).

Having these, the sender faces an optimization problem in the beginning of the game (Line 1 in Procedure 1):

$$\begin{aligned} \max_{\varphi} \quad & \mathbb{E}_{\mu_0, \tilde{\varphi}, \tilde{\pi}} [r^i(s, a)] := \sum_s \mu_0(s) \sum_a \tilde{\varphi}(\sigma = a | s) \cdot r^i(s, a) \\ \text{s.t.} \quad & \sum_s \mu_0(s) \cdot \tilde{\varphi}(\sigma = a | s) \cdot [r^j(s, a) - r^j(s, a')] \geq 0, \quad \forall a, a' \in \mathbb{A}. \end{aligned} \quad (2)$$

That is, the sender is to select an incentive compatible signaling scheme for the receiver, to maximize its (the sender’s) own expected payoffs. We refer to the canonical Bayesian persuasion as the formulation in Dughmi (2017). Concrete examples of persuasion are provided in Section 4.1.

Definition 2.4 (Bayesian Persuasion). *Under Assumption 2.1 and Proposition 2.2, a Bayesian persuasion task is defined as $\mathcal{BP} := (\mathbb{S}, \mu_0, \mathbb{A}, r^i, r^j)$, where a sender tries to persuade a receiver according to Procedure 1.*

2.2 BARGAINING GAME

The bargaining game, initially introduced by Nash et al. (1950); Nash (1953), serves as a foundational model in the study of strategic negotiation. In this work, we follow the definition in Maschler et al. (2013).

Definition 2.5 (Bargaining Game (Maschler et al., 2013)). *A two-player bargaining game is defined as $\mathcal{BG} := (\mathbb{Y}, \mathbf{d})$, where:*

- $\mathbb{Y} \subseteq \mathbb{R}^2$ represents a nonempty and compact set known as the **feasibility set** (or agreement set), which is the set of all possible agreements.
- $\mathbf{d} = (d_i, d_j) \in \mathbb{R}^2$ denotes the **disagreement point**.
- There exists at least one agreement $y = (y_i, y_j) \in \mathbb{Y}$ such that $y_i > d_i$ and $y_j > d_j$.

A critical element of the bargaining game is the presence of a disagreement point, which fundamentally enables each player to unilaterally refuse cooperation. Concrete examples of bargaining are provided in Section 4.1.

A classic perspective on solutions is axiomatic models, which propose certain axioms as necessary and sufficient conditions for a bargaining solution. Notable bargaining solutions include the Nash bargaining solution (Nash et al., 1950), the Kalai–Smorodinsky bargaining solution (Kalai & Smorodinsky, 1975), and the proportional bargaining solution (Kalai, 1977). In this work, we follow the definition of bargaining solution in (Maschler et al., 2013).

Definition 2.6 (Solution Concept of Bargaining Game (Maschler et al., 2013)). *A solution concept of bargaining games is defined as a function $\omega_{\mathcal{BG}}$ which maps every bargaining game (\mathbb{Y}, \mathbf{d}) to an agreement $\mathbf{y} \in \mathbb{Y}$.*

3 REDUCTION FROM BAYESIAN PERSUASION TO BARGAINING GAMES

In this section, we show that Bayesian persuasion task is reducible to a bargaining game. The receiver has sufficient knowledge to calculate the game structure, and there always exists a disagreement point in Bayesian persuasion where the receiver can ignore the sender. At this disagreement point, the receiver can unilaterally set both players' payoffs to default values, which nullifies the sender's informational advantage. Thus the receiver has the ability to make a commitment before the game starts, just like the sender. From this intuition, Bayesian persuasion is reducible to a bargaining game. This reduction suggests that the receiver in Bayesian persuasion may have retaliatory strategies, and it opens up possibilities for further discussions in behavioral economics.

Our understanding of the existing related work that employs bargaining strategies in Bayesian persuasion is discussed in Appendix B.2.

3.1 THE RECEIVER'S AWARENESS OF THE GAME STRUCTURE

We start with recalling how the receiver will act before they reach a consensus, i.e., $\tilde{\pi}(a | \sigma = a)$ does not have to be 1 and thus cannot be simplified.

Since $\mu_0, \tilde{\varphi} \in \mathbb{V}^{i+j}$, given a $\tilde{\pi}$, the Bayesian receiver is able to calculate its ground truth expected payoff as $\sum_s \mu_0(s) \sum_\sigma \tilde{\varphi}(\sigma | s) \sum_{s'} \mu(s' | \sigma) \sum_a \tilde{\pi}(a | s') \cdot r^j(s, a)$, where $\mu(s | \sigma) = (\mu_0(s) \cdot \tilde{\varphi}(\sigma | s)) / (\sum_{s'} \mu_0(s') \cdot \tilde{\varphi}(\sigma | s'))$ is the posterior belief according to Bayes' rule, and $\tilde{\pi}$ optimizes the receiver's payoff from its posterior view. Or more generally for an arbitrary decision rule $\tilde{\pi}$, the receiver's ground true expected payoff is

$$R^j(\mu_0, \tilde{\varphi}, \tilde{\pi}, r^j) := \sum_s \mu_0(s) \sum_\sigma \tilde{\varphi}(\sigma | s) \sum_a \tilde{\pi}(a | \sigma) \cdot r^j(s, a). \quad (3)$$

Because the receiver can see the sender's reward function ($r^i \in \mathbb{V}^{i+j}$), the receiver can also calculate the sender's expected payoff, by replacing $r^j(s, a)$ in Equation (3) with $r^i(s, a)$.

Define the **game structure** for every task \mathcal{BP} as a mapping $\psi : \tilde{\Phi} \times \tilde{\Pi} \rightarrow \{\mathbf{R}\}$, where $\{\mathbf{R}\} := \{(R^i, R^j)\}$ is the set of all possible expected payoff pairs in \mathcal{BP} . The game structure is determined by the nature of the parameters and the intrinsic setting of Bayesian persuasion. It therefore can be reasoned out if the task parameters are specified. We can now conclude that $\psi \in \mathbb{V}^{i+j}$, since it can be reasoned out by both players using other common knowledge. Let Ψ denote all the possible game structure mappings for different \mathcal{BP} tasks.

Recall that, given a committed signaling scheme $\tilde{\varphi}$ and a signal σ , the receiver will first calculate its posterior belief, guess a state s' according to Bayesian decision theory, and then choose its best response to this guess. We call this default Bayesian best response as the action rule $\tilde{\pi}_1 := \text{BestResponseTo}(\mathcal{BP}, \tilde{\varphi}, \sigma)$. The corresponding expected payoff pairs is denoted as $\mathbf{R}_1 = (R_1^i, R_1^j) = \psi(\tilde{\varphi}, \tilde{\pi}_1)$.

3.2 EXISTENCE OF THE DISAGREEMENT POINT IN PERSUASION

In all cases, the receiver always has the choice to ignore the sender's signaling scheme and signals. It can calculate its best response solely on its prior knowledge. We denote the prior best response as the decision rule $\tilde{\pi}_0 := \text{BestResponseTo}(\mathcal{BP})$. The corresponding expected payoff pairs is denoted as $\mathbf{R}_0 = (R_0^i, R_0^j) = \psi(\cdot, \tilde{\pi}_0)$, since it is independent of $\tilde{\varphi}$. It means that the receiver has the ability to **unilaterally** determine both players' expected payoffs, which aligns with the intuition of the disagreement points in bargaining games.

The sender also possesses the unilateral ability to refuse cooperation by specifically choosing $\tilde{\varphi}$ such that $\mu = \mu_0$ (for instance, regardless of the current state s , the sender consistently sends a dummy signal σ_0 , i.e., $\tilde{\varphi}(\sigma_0 | s) = 1, \forall s$). In this case, $\tilde{\pi}_1 = \tilde{\pi}_0$ and thus $\mathbf{R}_1 = \mathbf{R}_0$.

3.3 EXISTENCE ASSUMPTION OF BETTER OUTCOMES IN PERSUASION

The sender may not benefit from persuasion. That is, there is no guarantee that $R_1^i > R_0^i$ holds for any \mathcal{BP} .

Proposition 3.1 (Paraphrased from Kamenica & Gentzkow (2011), Proposition 3). *If there does not exist a signaling scheme $\tilde{\varphi}$ such that, under the best response of a Bayesian receiver, the expected payoff of the sender is higher than the expected payoff under the receiver’s best response to the prior distribution μ_0 , then the sender will not benefit from persuasion.*

The receiver may not benefit from persuasion either. In fact, (1) Based on the definition of R_1^j , the receiver utilizes Bayesian decision theory to guess a state s' and optimizes its strategy accordingly. According to the properties of Bayesian decision theory, the error rate of guessing $s' = s$ using the posterior distribution μ is not greater than that of using the prior distribution μ_0 . (2) The upper bound of the receiver’s expected payoff is when the sender honestly reports the state s , allowing the receiver to observe s directly. Each time the receiver guesses incorrectly, its expected payoff does not increase; combining (1) and (2), it follows that $R_1^j \geq R_0^j$. We therefore need to argue under the following condition.

Assumption 3.2 (Existence of Better Outcomes). *For a Bayesian persuasion $\mathcal{BP} = (\mathbb{S}, \mu_0, \mathbb{A}, r^i, r^j)$, there exists at least a strategy pair $(\tilde{\varphi}, \tilde{\pi})$ for the sender and receiver such that their respective expected payoffs exceed those at the disagreement point, i.e., $R^i(\tilde{\varphi}, \tilde{\pi}) > R^i(\tilde{\varphi}_0, \tilde{\pi}_0)$ and $R^j(\tilde{\varphi}, \tilde{\pi}) > R^j(\tilde{\varphi}_0, \tilde{\pi}_0)$.*

Similar arguments to our reduction will show that if Assumption 3.2 does not hold, then it is a bargaining game where they are to divide 0 coin in Example 4.4.

3.4 ACTION RULE WITH GAME STRUCTURE AWARENESS

The receiver’s awareness of the game structure is important, as it allows more complicated action rules. We define the **action rule with game structure awareness** as $\pi : \Psi \times \tilde{\Phi} \times \Sigma \rightarrow \Delta(\mathbb{A})$, which enables Receiver to decide its $\tilde{\pi}$ depending on both the game structure ψ and the committed signaling scheme $\tilde{\varphi}$.

Consider an example π : Satisfaction Check, as described in Algorithm 2. The **satisfaction threshold function** is defined as an indicator function $\chi : \{\mathbf{R}\}^2 \rightarrow \{0, 1\}$ and reflects the receiver’s non-myopic preferences, such as reflecting the receiver’s demands for fairness. In the algorithm, the satisfaction threshold function is $\chi(\mathbf{R}_0, \mathbf{R}_1) = 1$ if and only if $R_1^i \leq R_1^j$, which means that the receiver will only respect the sender’s recommendation if its posterior expected payoff is higher than the sender’s. Then, if the receiver is satisfied with the committed signaling scheme, it will choose its best response action rule. Otherwise, the receiver will ignore all the sender’s signals, and the sender will lose its influence and the potential benefits of its informational advantage.

Despite the introduction of the satisfaction threshold function may seem to suggest that the receiver is acting as a consequence of temperament or impulsive behavior, it aligns with the assumption of rationality. Detailed discussion is in Appendix C.

Algorithm 2: Satisfaction Check π : An Example Decision Rule with Game Structure Awareness

Input: A Bayesian persuasion task $\mathcal{BP} = (\mathbb{S}, \mu_0, \mathbb{A}, r^i, r^j)$, committed signaling scheme $\tilde{\varphi}$, received signal σ ;

Output: An action distribution $b \in \Delta(\mathbb{A})$;

Initialize: Decide a satisfaction threshold function $\chi : \{\mathbf{R}\}^2 \rightarrow \{0, 1\}$;

Initialize: Reason out the game structure ψ of \mathcal{BP} ;

1 Calculate the best response to the posterior distribution $\tilde{\pi}_1 = \text{BestResponseTo}(\mu_0, \tilde{\varphi})$;

2 Ignore the signaling scheme and calculate the best response to the prior distribution $\tilde{\pi}_0 = \text{BestResponseTo}(\mu_0)$;

3 Calculate players’ ground true expected payoffs $\mathbf{R}_0 = \psi(\tilde{\varphi}, \tilde{\pi}_0)$ and $\mathbf{R}_1 = \psi(\tilde{\varphi}, \tilde{\pi}_1)$;

4 **if** $R_1^i \leq R_1^j$ **then** $b = \tilde{\pi}_1(\cdot | \sigma)$; // Satisfaction check. $\chi(\mathbf{R}_0, \mathbf{R}_1) = 1$.

5 **else** $b = \tilde{\pi}_0(\cdot | \sigma)$;

6 **return** b

The satisfaction threshold function is not unique. It could also depend on the committed signaling scheme $\tilde{\varphi}$, denoted as $\chi(\tilde{\varphi})$. An intuitive example is $\Sigma = \mathbb{S}$ and $\chi(\tilde{\varphi})$ returns 1 if and only if the

324 sender is honest about the state, i.e., $\tilde{\varphi}(\sigma = s \mid s) = 1, \forall s$. We observe this satisfaction check rule
 325 implemented by LLM agents in the experiments (details in Appendix F).
 326

327 3.5 IMPLICATIONS OF LONG-TERM INTERACTION IN PERSUASION 328

329 The commitment assumption (Assumption 2.1) indicates that **Bayesian persuasion involves long-**
 330 **term interactions.** Without the implication of long-term interactions, the sender has no incentive to
 331 care about its reputation. It could tell the receiver that it would act upon one signaling scheme but in
 332 fact use another one. On the flip side, the receiver can also notice this possibility and will not trust
 333 the sender.

334 The underlying long-term nature of Bayesian persuasion enables the receiver to establish more so-
 335 phisticated action rule π , by designing a satisfaction threshold. Justified by the sender’s commit-
 336 ment assumption, it is also conceivable that the receiver has a commitment by disclosing its sat-
 337 isfaction function to the sender. If the receiver commits before the sender, then the sender must
 338 consider whether to fulfill the receiver’s expectations. This is recognized as **the power of commit-**
 339 **ment** (Nowak et al., 2000). Detailed discussions can be found in Appendix C.
 340

341 3.6 JOINT COMMITMENT 342

343 We define a joint commitment condition, where the sender and receiver will discuss to reach a
 344 consensus and then commit a strategy pair $(\tilde{\varphi}, \tilde{\pi})$. This strategy pair becomes common knowledge
 345 before the persuasion game starts, and then both players will play their committed strategy during
 346 the game.

347 **Definition 3.3** (Joint Commitment). *A joint commitment of a Bayesian persuasion task $\mathcal{BP} =$*
 348 *$(\mathbb{S}, \mu_0, \mathbb{A}, r^i, r^j)$ is a strategy pair $(\tilde{\varphi}, \tilde{\pi}) \in \tilde{\Phi} \times \tilde{\Pi}$, where the sender commits to $\tilde{\varphi}$ and the re-*
 349 *ceiver commits to $\tilde{\pi}$, subject to the conditions that $\tilde{\varphi}^* \neq \tilde{\varphi}_0, \tilde{\pi}^* \neq \tilde{\pi}_0, \tilde{\Phi} := \{\tilde{\varphi} : \mathbb{S} \rightarrow \Delta(\Sigma)\}$, and*
 350 *$\tilde{\Pi} := \{\tilde{\pi} : \Sigma \rightarrow \Delta(\mathbb{A})\}$.*

351 **Definition 3.4** (Solution of Bayesian Persuasion). *A solution concept of Bayesian persuasion $\mathcal{BP} =$*
 352 *$(\mathbb{S}, \mu_0, \mathbb{A}, r^i, r^j)$ is a function $\omega_{\mathcal{BP}}$ which maps every \mathcal{BP} to a joint commitment $(\tilde{\varphi}, \tilde{\pi}) \in \tilde{\Phi} \times \tilde{\Pi}$,*
 353 *where $\tilde{\Phi} := \{\tilde{\varphi} : \mathbb{S} \rightarrow \Delta(\Sigma)\}$ and $\tilde{\Pi} := \{\tilde{\pi} : \Sigma \rightarrow \Delta(\mathbb{A})\}$.*

355 3.7 REDUCTION FROM BAYESIAN PERSUASION TO BARGAINING GAMES 356

357 **Lemma 3.5** (Reduction from Bayesian Persuasion to Bargaining Games). *When Assumption 3.2*
 358 *holds, a Bayesian persuasion task $\mathcal{BP} = (\mathbb{S}, \mu_0, \mathbb{A}, r^i, r^j)$ is polynomial-time reducible to a bar-*
 359 *gaining game $\mathcal{BG} = (\mathbb{Y}, \mathbf{d})$. That is, $\mathcal{BP} \leq_p \mathcal{BG}$.*

360 *Proof.* With Assumption 3.2, there exists \mathbf{R} such that $\mathbf{R} > \mathbf{R}_0$. Thus we can investigate the bar-
 361 gaining game where $\mathbb{Y} = \{\mathbf{R}\}$ and $\mathbf{d} = \mathbf{R}_0$. Given a bargaining solution concept $\omega_{\mathcal{BG}}$, one could
 362 obtain the solution of \mathcal{BG} , denoted as $\mathbf{y}_{\mathcal{BG}}$. Then we have $\mathbf{y}_{\mathcal{BG}} \in \mathbb{Y} = \{\mathbf{R}\}$. In this way, \mathcal{BP} can be
 363 formulated as an optimization problem
 364

$$365 \min_{(\tilde{\varphi}, \tilde{\pi})} 0 \quad \text{s.t.} \quad \mathbf{R} = \mathbf{y}_{\mathcal{BG}}. \quad (4)$$

366 By Equation (3), the optimization problem (4) is convex, which can be solved in polynomial time,
 367 as we desired. Denote the solved strategy pair as $(\tilde{\varphi}^*, \tilde{\pi}^*)$ and the induced expected payoffs as \mathbf{R}^* .
 368 Since $\mathbf{R} = \mathbf{y}_{\mathcal{BG}} \in \{\mathbf{R}\}$, we have $\mathbf{R}^* > \mathbf{R}_0$ by Definition 2.5. Thus $\tilde{\varphi}^* \neq \tilde{\varphi}_0$ and $\tilde{\pi}^* \neq \tilde{\pi}_0$,
 369 otherwise $\mathbf{R}^* = \mathbf{R}_0$. In this way, the strategy pair $(\tilde{\varphi}^*, \tilde{\pi}^*)$ is a joint commitment, which means we
 370 have found the solution of \mathcal{BG} . \square
 371

372 4 EMPIRICAL RESULTS 373 374

375 In this section, we aim to demonstrate how the receiver in a BP task can use an action rule with game
 376 structure awareness to achieve higher expected gains than what the equilibrium of the canonical
 377 Bayesian persuasion formulation suggests. The experimental results reveal the receiver’s retaliatory
 strategies.

4.1 TASKS OF PERSUASION AND BARGAINING

In this section, we describe several examples, where Bayesian persuasion equilibrium is a special case solution in bargaining games. These examples will later be tested with LLM agents.

The grading student task, also known as recommendation letters, is introduced by Kamenica & Gentzkow (2011). We use the details provided in Dughmi (2017).

Example 4.1 (Grading Students, GS, Dughmi (2017)). *Recent graduates entering the job market are divided into 1/3 excellent ($s = 1$) students and 2/3 weak ($s = 0$) students, and the distribution is μ_0 . A professor (the sender), who can assess student quality, sends grades as signals to an HR representative (the receiver) who makes hiring decisions. The professor gains 1 reward per student hired, while the HR gains 1 reward per excellent student hired and -1 reward per weak student hired. Neither party gains if a student remains not hired.*

The professor and the HR have no conflict when the current student is excellent, so the professor will report the state honestly in this case. Thus, the signaling scheme is $\tilde{\varphi}(\sigma = 1 \mid s = 1) = 1$ and $\tilde{\varphi}(\sigma = 1 \mid s = 0) = \eta$, where η is a parameter $0 \leq \eta \leq 1$. Assuming the HR acts based on its posterior best response according to the canonical Bayesian persuasion, there are three example outcomes analyzed in Dughmi (2017): (1) The professor provides no information ($\eta = 1$). Their expected payoffs are both 0. (2) The professor is fully honest ($\eta = 0$). Their expected payoffs are both 1/3. (3) The professor chooses an η less than but very close to 1/2. The expected payoffs of the professor and the HR are $(1 + 2\eta)/3$ and $(1 - 2\eta)/3$, respectively. The detailed analysis is in Appendix D.

The context of grading students has societal implications, because professors are expected to be more honest when providing information about their students. We therefore also test on an alternative example of Selling Products (SP), which mathematically has the same game structure, but experiences less social context.

Example 4.2 (Selling Products, SP. A variant of Example 4.1.). *A seller (the sender) recommends its products to a buyer (the receiver). State is the quality of the product. The detailed value settings are the same with Example 4.1.*

The game structure of persuasion tasks, i.e. Example 4.1 and Example of 4.2, are equivalent to bargaining tasks, i.e. Asymmetric Bargaining (AB) in Example 4.3 and Splitting Coins (SC) in Example 4.4.

Example 4.3 (Asymmetric Bargaining, AB). *A proposer suggests an $\eta \in [0, 1/2]$. If a responder agrees, the proposer receives a payoff of $(1+2\eta)/3$ and the responder receives a payoff of $(1-2\eta)/3$. If the responder disagrees, both get nothing.*

Example 4.4 (Splitting Coins, SC). *Two players split 100 coins. A Proposer suggests a split, and a Responder either accepts or rejects it. If accepted, they split as proposed; otherwise, both get nothing.*

4.2 EXPERIMENT SETUP WITH LARGE LANGUAGE MODELS

One way to test the behavior of intelligent agents in these tasks is to conduct human experiments. With the availability of LLMs, there are 3 main advantages of testing with LLMs instead of humans: (1) It is easy to provide long context to LLMs. In contrast, explaining the concept of persuasion and bargaining to an average human population may require significant effort. (2) It is immediate for LLMs to provide their interpretation as to why such decisions are made by giving appropriate prompts. (3) The experiments by LLMs scale much easier than human experiments, and are also more efficient in terms of monetary costs and environmental costs (e.g. carbon emission (Tomlinson et al., 2024)).

The tasks are tested with 3 variants regarding the farsightedness, namely, (α) The agents play the game once, and were explicitly told that they will never meet again in the future. (β) The agents play the game once, but they are told that they may meet again in the future. (γ) The agents play the game multiple times. The maximum time T is sampled by a memoryless distribution, which introduces **the shadow of the future** (Bó, 2005), preventing players from predicting when the game will end. We defer the exact prompt used to Appendix E.

Task	Game	Setting	Sender’s Payoff	Receiver’s Payoff	Consensus Rate	η
SC	\mathcal{BG}	α	93.30 (4.41)	6.70 (4.41)	1.00	/
SC	\mathcal{BG}	β	78.5 (15.47)	21.5 (15.47)	1.00	/
SC	\mathcal{BG}	γ	58.53 (3.80)	41.47 (3.80)	1.00	/
AB	\mathcal{BG}	α	83.32 (37.9)	0.01 (0.03)	0.83	0.50 (0.00)
AB	\mathcal{BG}	β	70.13 (25.09)	19.87 (10.47)	0.90	0.30 (0.10)
AB	\mathcal{BG}	γ	74.46 (4.80)	25.54 (4.80)	1.00	0.24 (0.05)
GS	\mathcal{BP}	β	0.00 (0.00)	0.00 (0.00)	0.00	0.38 (0.13)
SP	\mathcal{BP}	β	0.00 (0.00)	0.00 (0.00)	0.00	0.45 (0.10)
GS	\mathcal{BP}	γ	57.63 (4.12)	42.37 (4.12)	1.00	0.08 (0.04)
SP	\mathcal{BP}	γ	58.00 (3.37)	42.00 (3.37)	1.00	0.08 (0.03)

Table 1: Experiments solved by gpt-4-turbo-2024-04-09. The payoffs are the mean and standard deviation from 30 runs. Players’ payoffs are scaled so that the maximum possible social welfare is 100, to facilitate comparison. The tasks SC, AB, GS, and SP are Splitting Coins, Asymmetric Bargaining, Grading Students, and Selling Products, respectively.

Task	Game	Setting	Sender’s Payoff	Receiver’s Payoff	Consensus Rate	η
SC	\mathcal{BG}	α	99.00 (0.00)	1.00 (0.00)	1.00	/
SC	\mathcal{BG}	β	79.37 (19.70)	20.63 (19.70)	1.00	/
SC	\mathcal{BG}	γ	51.60 (2.28)	48.4 (2.28)	1.00	/
AB	\mathcal{BG}	α	65.93 (24.04)	24.07 (12.07)	0.90	0.26 (0.12)
AB	\mathcal{BG}	β	71.00 (9.10)	29.00 (9.10)	1.00	0.21 (0.09)
AB	\mathcal{BG}	γ	74.21 (2.51)	25.79 (2.51)	1.00	0.24 (0.03)
GS	\mathcal{BP}	β	12.5 (28.43)	4.17 (9.48)	0.17	0.24 (0.00)
SP	\mathcal{BP}	β	0.00 (0.00)	0.00 (0.00)	0.00	0.26 (0.05)
GS	\mathcal{BP}	γ	61.72 (4.38)	38.28 (4.38)	1.00	0.12 (0.04)
SP	\mathcal{BP}	γ	71.42 (5.40)	28.58 (5.40)	1.00	0.21 (0.05)

Table 2: Experiments solved by gpt-4o-2024-08-06. The other settings are the same as Table 1.

Because the commitment assumption in persuasion tasks implicitly requires long-term interactions between the sender and receiver (details discussed in Section 3.5), setting (α) will not apply to persuasion tasks. In the bargaining games of setting (γ), the initial proposer is uniformly chosen at random from among the two players. If the responder rejects the offer, the game continues with players switching roles; The game ends when the offer is accepted or the maximum time is reached. This setting is known as Rubinstein’s alternating-offer model (Rubinstein, 1982). In persuasion games of setting (γ), the sender is always the proposer, and the receiver is always the responder. They will not switch the roles. The game ends when the receiver is satisfied with the committed signaling scheme or the maximum time is reached. This setting is similar to the model proposed by Gan et al. (2022).

All agents will have the common knowledge of the game settings. We conduct experiments on gpt-4o-2024-05-13 and gpt-4-turbo-2024-04-09 released by OpenAI (Achiam et al., 2023; Roumeliotis & Tselikas, 2023) through their API. Each configuration is tested with 30 independent runs. The total cost of querying the API of this work is \$39.86 United States dollars.

4.3 EXPERIMENT RESULTS

Table 1 and 2 summarize the results of LLM agents playing the task games. We now discuss the results and the implications. More detailed output examples are given in Appendix F.

Receiver’s Retaliation and Anti-Exploitation. In all bargaining and persuasion tasks, the proposer’s offers given by LLM agents are incentive compatible, meaning that a myopic responder would accept it albeit the expected payoff they are offered. However, we notice that the receiver does not always accept: All cases where the consensus rate is not 1 are caused by the receiver’s retaliatory action rule. The observed reasons are centered on that the receiver believes the sender’s

486 signaling scheme results in too large a gap in their expected payoffs. E.g., in an outcome of SP- γ
 487 by GPT-4-turbo, the buyer says that “Given the proposed signaling scheme where the seller scores
 488 bad products as good with a probability of 0.25, my expected payoff as Receiver is 0.167. This is
 489 significantly lower than the maximum possible payoff of 0.333 if no bad products are purchased.
 490 Therefore, I am not satisfied with this division as it heavily favors the seller.” Interestingly, in the
 491 persuasion tasks GS- β and SP- β by GPT-4-turbo, as well as SP- β by GPT-4o, the receiver rejected
 492 all signaling schemes, even if the proposed signaling scheme are essentially more fairer than the one
 493 in the canonical Bayesian persuasion analysis. It means under these settings the two instances of
 494 LLMs cannot effectively establish a deal between them.

495
 496 **Effects of Long-Term Interaction.** For the proposer in most cases under setting (γ) by GPT-4o,
 497 it first proposes a much unfair offer and then slowly adjust to make it to be more satisfactory until
 498 the responder accepts it. For the responder in these cases, they reject the offers in the first few
 499 rounds, even in the tasks which demonstrate consensus rates of 1. In the end, they always agree on
 500 a compromised signaling scheme (the consensus rate reached its maximum value of 1). In GPT-
 501 4-turbo, the iterative setting, (γ), results in more fair outcomes than settings (α) and (β), with the
 502 sender’s signaling scheme being almost honest in the persuasion tasks GS- γ and SP- γ .

503 We also notice that even if the game is played only once, the given context of the potential for
 504 future encounters can change their behavior. The proposer is more cautious in setting (β) by GPT-
 505 4o compared to setting (γ). E.g., in an outcome of SC- β by GPT-4o, the buyer says that “I also
 506 need to consider that we might play this game again in the future, so I should offer a fair amount
 507 to ensure future cooperation. Offering the minimum amount that Responder would accept, such as
 508 1 coin, might be risky if they expect a fairer split.” In SC- β compared to SC- α , by GPT-4-turbo
 509 or GPT-4o, the sender’s expected payoff decreased, while the receiver’s payoff increased, and the
 510 consensus rate remained at its maximum value of 1. In AB- β compared to AB- α , by GPT-4-turbo
 511 or GPT-4o, the sender lowered their η , making the offer fairer, which resulted in an increase in the
 512 consensus rate and thus an improvement in social welfare (the sum of the proposer’s payoff and the
 513 responder’s payoff).

514 **Effects of the Task Context.** Different scenario contexts may induce different outcomes on the
 515 persuasion tasks. By both LLM solvers in settings (β) and (γ), the proposer tends to be fairer in
 516 Example 4.1 than in Example 4.2, even though these two examples are mathematically equivalent.
 517 In human social behavior, using the actions of people in a scenario to explain and predict beliefs is
 518 an important topic, where Bayesian theory of mind (Baker, 2012) studies how individuals update
 519 their beliefs based on the actions of others. The training process of LLMs includes reinforcement
 520 from human feedback, thereby aligning the models with human preferences. In tasks GS and SP,
 521 different contexts provided the LLM with varying prior beliefs about how the sender would perform.
 522 We speculate that the proposer may feel more pressure to lie when the situation involves human
 523 subjects (students). We tested the impact of these stereotypes on the identity of the sender in these
 524 scenarios, and the results showed that both GPT-4o and GPT-4-turbo perceived the professor as
 525 more honest than the salesperson: “The combination of ethical obligations, potential long-term
 526 consequences on reputation, and stronger normative frameworks in the academic setting of (GS)
 527 supports a higher likelihood of honest behavior compared to the commercial and competitive setting
 528 of (SP).” In reality, professors are expected to uphold academic integrity (Macfarlane et al., 2014)
 529 in their conduct. Sellers, on the other hand, do not have such explicit ethical constraints. Therefore,
 530 this phenomenon interestingly reflects how the contexts affect the persuasion processes.

531 5 CONCLUSION

532
 533 We conclude that Bayesian persuasion is reducible to a bargaining game, which agrees with the
 534 diverse outcome of observation tasks in practice. This perspective offers a nuanced understanding
 535 of Bayesian persuasion, suggesting that signaling schemes may not only serve the sender’s interests
 536 but also encourage the receiver to adopt retaliatory tactics. Our experimental evidence indicates
 537 that these strategies can lead to greater honesty from the sender, thus fostering more equitable in-
 538 teractions. Our results provide insights into the potential of the receiver to have anti-exploitation
 539 and counter-exploitation strategies, and could potentially be combined with other methods, such as
 institution, for improving equality and social welfare.

REFERENCES

- 540
541
542 Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Ale-
543 man, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. Gpt-4 technical
544 report. *arXiv preprint arXiv:2303.08774*, 2023.
- 545 Ricardo Alonso and Odilon Câmara. On the value of persuasion by experts. *Journal of Economic*
546 *Theory*, 174:103–123, 2018.
- 547 Gerry Antioch. Persuasion is now 30 per cent of us gdp: Revisiting mcclouskey and klamer after a
548 quarter of a century. *Economic Round-up*, pp. 1–10, 2013.
- 549 Chris Lawrence Baker. *Bayesian theory of mind: Modeling human reasoning about beliefs, desires,*
550 *goals, and social relations*. PhD thesis, Massachusetts Institute of Technology, 2012.
- 551
552 Dirk Bergemann and Stephen Morris. Robust predictions in games with incomplete information.
553 *Econometrica*, 81(4):1251–1308, 2013.
- 554
555 Dirk Bergemann and Stephen Morris. Bayes correlated equilibrium and the comparison of informa-
556 tion structures in games. *Theoretical Economics*, 11(2):487–522, 2016.
- 557
558 Dirk Bergemann and Stephen Morris. Information design: A unified perspective. *Journal of Eco-*
559 *nomic Literature*, 57(1):44–95, 2019.
- 560
561 Martino Bernasconi, Matteo Castiglioni, Alberto Marchesi, and Mirco Mutti. Persuading farsighted
562 receivers in mdps: the power of honesty. *Advances in Neural Information Processing Systems*,
563 36, 2024.
- 564 JW Best and DP Quigley. Honestly dishonest: A solution to the commitment problem in bayesian
565 persuasion. Technical report, Mimeo, 2016.
- 566
567 Pedro Dal Bó. Cooperation under the shadow of the future: experimental evidence from infinitely
568 repeated games. *American economic review*, 95(5):1591–1604, 2005.
- 569
570 Raphael Boleslavsky and Christopher Cotton. Grading standards and education quality. *American*
571 *Economic Journal: Microeconomics*, 7(2):248–279, 2015.
- 572
573 Andrea Celli, Stefano Coniglio, and Nicola Gatti. Private bayesian persuasion with sequential
574 games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pp. 1886–
1893, 2020.
- 575
576 Vincent P Crawford and Joel Sobel. Strategic information transmission. *Econometrica: Journal of*
577 *the Econometric Society*, pp. 1431–1451, 1982.
- 578
579 Sanmay Das, Emir Kamenica, and Renee Mirka. Reducing congestion through information design.
580 In *2017 55th annual allerton conference on communication, control, and computing (allerton)*,
pp. 1279–1284. IEEE, 2017.
- 581
582 Yuan Deng, Vahab Mirrokni, Renato Paes Leme, Hanrui Zhang, and Song Zuo. Llms at the bargain-
583 ing table. In *Agentic Markets Workshop at ICML 2024*, 2024.
- 584
585 Shaddin Dughmi. Algorithmic information structure design: a survey. *ACM SIGecom Exchanges*,
586 15(2):2–24, 2017.
- 587
588 James W Friedman. A non-cooperative equilibrium for supergames. *The Review of Economic*
Studies, 38(1):1–12, 1971.
- 589
590 Jiarui Gan, Rupak Majumdar, Goran Radanovic, and Adish Singla. Bayesian persuasion in se-
591 quential decision-making. In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol-
592 ume 36, pp. 5025–5033, 2022.
- 593
Werner Güth and Reinhard Tietz. Ultimatum bargaining behavior: A survey and comparison of
experimental results. *Journal of Economic Psychology*, 11(3):417–449, 1990.

- 594 Werner Güth, Rolf Schmittberger, and Bernd Schwarze. An experimental analysis of ultimatum
595 bargaining. *Journal of economic behavior & organization*, 3(4):367–388, 1982.
- 596
- 597 Uriel Haran and Shaul Shalvi. The implicit honesty premium: Why honest advice is more persuasive
598 than highly informed advice. *Journal of Experimental Psychology: General*, 149(4):757, 2020.
- 599 Peter E Hart, David G Stork, Richard O Duda, et al. *Pattern classification*. Wiley Hoboken, 2000.
- 600
- 601 He He, Derek Chen, Anusha Balakrishnan, and Percy Liang. Decoupling strategy and generation in
602 negotiation dialogues. *arXiv preprint arXiv:1808.09637*, 2018.
- 603
- 604 Penélope Hernández and Zvika Neeman. How bayesian persuasion can help reduce illegal parking
605 and other socially undesirable behavior. *American Economic Journal: Microeconomics*, 14(1):
606 186–215, 2022.
- 607 Ehud Kalai. Proportional solutions to bargaining situations: interpersonal utility comparisons.
608 *Econometrica: Journal of the Econometric Society*, pp. 1623–1630, 1977.
- 609 Ehud Kalai and Meir Smorodinsky. Other solutions to nash’s bargaining problem. *Econometrica:*
610 *Journal of the Econometric Society*, pp. 513–518, 1975.
- 611
- 612 Emir Kamenica. Bayesian persuasion and information design. *Annual Review of Economics*, 11(1):
613 249–272, 2019.
- 614 Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101
615 (6):2590–2615, 2011.
- 616
- 617 Ilan Kremer, Yishay Mansour, and Motty Perry. Implementing the “wisdom of the crowd”. *Journal*
618 *of Political Economy*, 122(5):988–1012, 2014.
- 619 Edward P Lazear. Speeding, terrorism, and teaching to the test. *The Quarterly Journal of Economics*,
620 121(3):1029–1061, 2006.
- 621
- 622 Fei Li and Peter Norman. Sequential persuasion. *Theoretical Economics*, 16(2):639–675, 2021.
- 623
- 624 Po-Hsuan Lin, Alexander L Brown, Taisuke Imai, Joseph Tao-yi Wang, Stephanie W Wang, and
625 Colin F Camerer. Evidence of general economic principles of bargaining and trade from 2,000
626 classroom experiments. *Nature Human Behaviour*, 4(9):917–927, 2020.
- 627
- 628 Yue Lin, Wenhao Li, Hongyuan Zha, and Baoxiang Wang. Information design in multi-agent re-
629 inforcement learning. *Advances in Neural Information Processing Systems*, 36:25584–25597,
630 2023.
- 631
- 632 Bruce Macfarlane, Jingjing Zhang, and Annie Pun. Academic integrity: a review of the literature.
633 *Studies in higher education*, 39(2):339–358, 2014.
- 634
- 635 Michael Maschler, Eilon Solan, and Shmuel Zamir. *Bargaining games*, pp. 622–658. Cambridge
636 University Press, 2013.
- 637
- 638 Donald McCloskey and Arjo Klamer. One quarter of gdp is persuasion. *The American Economic*
639 *Review*, 85(2):191–195, 1995.
- 640
- 641 John Nash. Two-person cooperative games. *Econometrica: Journal of the Econometric Society*, pp.
642 128–140, 1953.
- 643
- 644 John F Nash et al. The bargaining problem. *Econometrica*, 18(2):155–162, 1950.
- 645
- 646 Martin A Nowak, Karen M Page, and Karl Sigmund. Fairness versus reason in the ultimatum game.
647 *Science*, 289(5485):1773–1775, 2000.
- 648
- 649 Michael Ostrovsky and Michael Schwarz. Information disclosure and unraveling in matching mar-
650 kets. *American Economic Journal: Microeconomics*, 2(2):34–63, 2010.
- 651
- 652 Joseph R Priester and Richard E Petty. Source attributions and persuasion: Perceived honesty as
653 a determinant of message scrutiny. *Personality and social psychology bulletin*, 21(6):637–654,
654 1995.

648 Konstantinos I Roulmeliotis and Nikolaos D Tselikas. Chatgpt and open-ai models: A preliminary
649 review. *Future Internet*, 15(6):192, 2023.
650

651 Ariel Rubinstein. Perfect equilibrium in a bargaining model. *Econometrica: Journal of the Econo-*
652 *metric Society*, pp. 97–109, 1982.

653 Bill Tomlinson, Rebecca W Black, Donald J Patterson, and Andrew W Torrance. The carbon emis-
654 sions of writing and illustrating are lower for ai than for humans. *Scientific Reports*, 14(1):3732,
655 2024.

656

657 Jibang Wu, Zixuan Zhang, Zhe Feng, Zhaoran Wang, Zhuoran Yang, Michael I Jordan, and Haifeng
658 Xu. Sequential information design: Markov persuasion process and its efficient reinforcement
659 learning. *arXiv preprint arXiv:2202.10678*, 2022.

660

661 Wenhao Wu. Sequential bayesian persuasion. *Journal of Economic Theory*, 214:105763, 2023.

662

663 Tian Xia, Zhiwei He, Tong Ren, Yibo Miao, Zhuosheng Zhang, Yang Yang, and Rui Wang. Measur-
664 ing bargaining abilities of llms: A benchmark and a buyer-enhancement method. *arXiv preprint*
665 *arXiv:2402.15813*, 2024.
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701

A GLOSSARY

$\Delta(\mathbb{X})$ denotes the simplex of the set \mathbb{X} . $\{x\}$ denotes the set whose elements are all the possible values of x .

Symbol	Meaning	Note
i	Sender's index	
j	Receiver's index	
$s \in \mathbb{S}$	Environmental state	
$\mu_0 \in \Delta(\mathbb{S})$	Prior distribution of states	
$a \in \mathbb{A}$	The receiver's action	
$\sigma \in \Sigma$	The sender's signal	$\Sigma = \mathbb{A}$ with revelation principle (Proposition 2.2)
$\tau \in \mathbb{T}$	The receiver's type	It can be seen by the sender but not by other receivers.
$q : \mathbb{S} \rightarrow \Delta(\mathbb{T})$	Emission function	q is w.l.o.g. to be dummy if there is only one receiver.
$\tilde{\Phi} := \{\tilde{\varphi} : \mathbb{S} \rightarrow \Delta(\Sigma)\}$	The sender's signaling scheme set (without the game structure awareness)	The canonical setting in Bayesian persuasion
$\phi : \mathbb{S} \times \mathbb{T} \rightarrow \Sigma$	The sender's signaling scheme if \mathbb{T} is not dummy	
$\tilde{\Pi} := \{\tilde{\pi} : \Sigma \rightarrow \Delta(\mathbb{A})\}$	The receiver's action rule set (without the game structure awareness)	The canonical setting in Bayesian persuasion
$\tilde{\pi}_0$	The receiver's prior best response	The receiver ignores the sender
$\tilde{\pi}_1$	The receiver's posterior best response	The default behavior in Bayesian persuasion
$r^i : \mathbb{S} \times \mathbb{A} \rightarrow \mathbb{R}$	The sender's reward function	
$r^j : \mathbb{S} \times \mathbb{A} \rightarrow \mathbb{R}$	The receiver's reward function	
$\mathcal{BP} := (\mathbb{S}, \mu_0, \mathbb{A}, r^i, r^j)$	Bayesian persuasion task	
$\omega_{\mathcal{BP}}$	A solution concept of Bayesian persuasion	It maps every \mathcal{BP} to a strategy pair $(\tilde{\varphi}, \tilde{\pi}) \in \tilde{\Phi} \times \tilde{\Pi}$
$\{R^i(\mu_0, \tilde{\varphi}, \tilde{\pi}, r^i)\} \subset \mathbb{R}$	The sender's expected reward given $\mu_0, \tilde{\varphi}, \tilde{\pi}$, and r^i	Simplified as $R^i(\tilde{\varphi}, \tilde{\pi})$ when \mathcal{BP} is given and clear
$\{R^j(\mu_0, \tilde{\varphi}, \tilde{\pi}, r^j)\} \subset \mathbb{R}$	The receiver's expected reward given $\mu_0, \tilde{\varphi}, \tilde{\pi}$, and r^j	
$\mathbf{R} := (R^i(\tilde{\varphi}, \tilde{\pi}), R^j(\tilde{\varphi}, \tilde{\pi}))$	Expected reward outcomes	\mathbf{R}_0 if $\tilde{\pi} = \tilde{\pi}_0$ and \mathbf{R}_1 if $\tilde{\pi} = \tilde{\pi}_1$
$\Psi := \{\psi : \tilde{\Phi} \times \tilde{\Pi} \rightarrow \{\mathbf{R}\}\}$	The game structure of a Bayesian persuasion task	Every ψ is determined by a \mathcal{BP}
$\chi : \Psi \times \tilde{\Phi} \rightarrow \{0, 1\}$	The receiver's satisfaction threshold function	
$\Pi := \{\pi : \Psi \times \tilde{\Phi} \times \Sigma \rightarrow \Delta(\mathbb{A})\}$	The receiver's action rule (with the game structure awareness)	E.g. $\pi_{\chi(\psi, \varphi)}$
\mathbb{V}^i	The sender's visibility set	It specifies what the sender can see
\mathbb{V}^j	The receiver's visibility set	$\mathbb{V}^i \setminus \mathbb{V}^j = s$
\mathbb{V}^{i+j}	Players' common knowledge	Known by the players, with each aware that others know it, ad infinitum

Table 3: Glossary of Bayesian Persuasion

Symbol	Meaning	Note
\mathbb{Y}	Feasibility set (or agreement set)	
d	Disagreement point	
$\mathcal{BG} := (\mathbb{Y}, d)$	Bargaining game	
$\omega_{\mathcal{BG}}$	A solution concept of bargaining games	It maps every $\mathcal{BG} = (\mathbb{Y}, d)$ to an agreement $y \in \mathbb{Y}$

Table 4: Glossary of Bargaining Games

B RELATED WORK

B.1 RELATED WORK ON BARGAINING

Bargaining Games as Extensive-Form Games. Another perspective on solving the bargaining game is to formulate it as an extensive form game (Rubinstein, 1982). The proposer extends an offer, and the responder has the choice to either accept this offer or reject it. If accepted, the two parties reach a consensus and gain benefits; if rejected, both parties receive a default value (usually 0). If there is only one round, the responder will accept any offer greater than 0. In a multi-round scenario where the responder can propose offers after rejecting, this process repeats multiple times. After introducing a discount factor to represent patience levels, the expected outcomes for both parties will be a function of this discount factor. In particular, when both parties are infinitely patient (with a discount factor of 1), they will reach a deal in the first round and split the benefits equally. These solutions are referred to as subgame perfect equilibria.

Human Experiments. Güth & Tietz (1990); Güth et al. (1982) found that considerations of distributive justice can significantly undermine strategic power in ultimatum bargaining. A more recent study by Lin et al. (2020) analyzed a large dataset, revealing that equal-split offers are accepted more frequently and quickly than slightly unequal ones, in both one-shot and repeated interactions.

Evolutionary Game Theory. Another approach to understanding bargaining behavior is through evolutionary game theory. Nowak et al. (2000) developed an evolutionary model of the one-shot bargaining game, demonstrating that fairness can evolve if the proposer has access to information about past deals accepted by the responder. This suggests that the evolution of fairness is closely tied to reputation. Such insights highlight the dynamic nature of fairness in negotiations and its implications for strategic interactions.

Language Models Experiments. An early attempt to explore bargaining using language models was made by He et al. (2018), who employed generative models to conduct bargaining experiments. They proposed a method to decouple the generation and decision-making processes. With the rise of large language models, recent studies have emerged focusing on their application in bargaining scenarios. Xia et al. (2024) introduced a benchmark for a buyer-seller bargaining game, demonstrating that OpenAI’s ChatGPT performed well across various metrics among several large models. Another work (Deng et al., 2024) noted that “LLM agents can (naturally) be good negotiators,” highlighting the potential of LLMs in negotiation contexts.

B.2 EVIDENCE OF BAYESIAN PERSUASION BEING A BARGAINING GAME.

The Hardness of Persuading a Far-sighted Receiver Gan et al. (2022) discusses the long-term interaction between the sender and receiver in an MDP, where both are aware of the realization of states in the MDP. The sender’s informational advantage lies in knowing an external parameter that only affects their payoffs without impacting the MDP’s transition function. In such a setting, persuading a far-sighted receiver is NP-hard. **Evidence (1)** They considered an easier setting where the goal is to persuade an advice-myopic receiver, who treats the sender as someone that will disappear in the future and makes decisions based solely on prior beliefs. They proved that persuading an advice-myopic receiver is solvable in polynomial time because the simplification of the receiver’s strategy leads to certain properties in the MDP that can be simplified. Such a policy of the receiver is equivalent to a repeated bargaining game where the receiver commits to reject cooperation in the

810 future. **Evidence (2)** They introduced a threat-based meta signaling scheme for the sender, where if
 811 the sender finds that Receiver does not follow its recommendation, it will cease providing any infor-
 812 mation in the future. In this case, the receiver’s best response is to follow the recommendation. This
 813 strategy of the sender is a grim-trigger policy (Friedman, 1971) in a repeated game, and threat-based
 814 strategies are also common in bargaining scenarios.

815
 816 **Markov Signaling Games, MSGs.** (Lin et al., 2023) proposed another Markov process (MSG)
 817 in which only the sender can see the Markov state. The receiver has observations, but these obser-
 818 vations are common knowledge among the players; thus, the sender’s informational advantage is
 819 reflected in the difference between the Markov state and Receiver’s observations. The main differ-
 820 ence from Gan et al. (2022) is that the sender’s informational advantage covers the MDP transitions,
 821 since the transition function depends on the Markov state. In the discussion of the method in (Lin
 822 et al., 2023), a threat strategy of the receiver is mentioned. To strengthen the sender’s persuasion,
 823 they constrain the signaling scheme to a stronger obedience, ensuring that the recommended actions
 824 generate higher expected posterior payoffs for the receiver, thus satisfying it.

825 **History-Dependent Signaling Scheme.** Bernasconi et al. (2024) extends the setting of Gan et al.
 826 (2022) by letting the external parameter known exclusively to the sender determines the environ-
 827 ment’s state transitions. Their work demonstrates that Markovian signaling schemes are not optimal
 828 and introduces a convenient subset of history-dependent signaling schemes called promise-form.
 829 This approach encodes past interactions as honest promises about the receiver’s future rewards,
 830 which can be understood as a strategy the sender uses to respond to bargaining situations.

831
 832 **Markov persuasion processes (MPPs).** Another series of studies examines the Markov persua-
 833 sion processes (MPPs) model (Wu et al., 2022). In this model, long-term interactions occur only
 834 between the sender and the Markovian environment. At each timestep, a new receiver interacts with
 835 the sender, leaves the system before the next timestep begins, and receives the corresponding reward.
 836 Each receiver is myopic and will only interact with the sender once. As a result, the players do not
 837 engage with the issues involved in repeated games, and there is no bargaining phenomenon.

838 B.3 OTHER SEQUENTIAL PERSUASION WORKS

839
 840 Alonso & Câmara (2018) investigated the trade-offs between the persuasiveness of expert advice
 841 and the reliability of the information provided. Celli et al. (2020) expanded on this by exploring
 842 the impact of private information on Bayesian persuasion in sequential game with imperfect infor-
 843 mation. Li & Norman (2021) extended classic Bayesian persuasion games to sequential Bayesian
 844 persuasion games with multiple senders. This sequential extension explored how information can
 845 be gradually revealed over time, enhancing strategic interaction in dynamic settings. Wu (2023)
 846 studied how the timing of information release and the order of persuasion efforts affect outcomes in
 847 sequential persuasion scenarios.

848 C COMMITMENT OF THE RECEIVER

849 C.1 SATISFACTION COMMITMENT

850
 851 The action rule $\pi : \Psi \times \tilde{\Phi} \times \Sigma \rightarrow \Delta(\mathbb{A})$ with game structure awareness allows us to define the
 852 commitment behavior for the receiver.

853 **Definition C.1** (Commitment of the Receiver). *The receiver commits to satisfaction threshold func-*
 854 *tion χ , which encompasses the following three aspects: (1) The receiver will decide on a χ before*
 855 *the game starts, (2) The receiver will honestly inform the sender of this χ , and (3) during the game,*
 856 *the receiver will act upon χ using Algorithm 2.*

857
 858
 859
 860 Committing a satisfaction threshold function does not contradict the rationality assumption. A tem-
 861 porary refusal to cooperate might inform the sender that the receiver is dissatisfied with the expected
 862 payoff outcome of the current signaling scheme, thereby forcing the sender to reveal more informa-
 863 tion about the state in the future. Therefore, choosing to refuse cooperation could be beneficial for
 long-term expected payoff gains, making it a rational behavior.

864 C.2 THE POWER OF COMMITMENT

865
866 In the canonical setting of Bayesian persuasion it assumes the commitment of the sender (Assump-
867 tion 2.1) and the sender only. The setting is one-shot, which disagrees with implication of long-term
868 interaction brought by the commitment assumption. The receiver then have to act upon Bayesian
869 best response, for it being rational and knowing that $R_1^j \geq R_0^j$.

870 However, if we similarly assume the commitment of the receiver, e.g., with Algorithm 2, then the
871 receiver turns to be the player who determines the set of possible outcomes. Intuitively, this is
872 analogous to the receiver being the “proposer” in a bargaining process. The receiver will then
873 only cooperate if the sender commits a $\tilde{\varphi}$ that induces an outcome that satisfies χ . Knowing this,
874 achieving incentive compatible requires providing higher payoff to the receiver for the sender. This
875 effect agrees with **the power of commitment** described in bargaining games, as described in Nowak
876 et al. (2000): “If the proposer has perfect knowledge of the responder’s q value, then it is in fact the
877 responder who makes the offer.”

879 D THE CANONICAL ANALYSIS OF EXAMPLE 4.1

880
881 If the current student is excellent ($s = 1$), the professor will report it honestly, since at this state they
882 have no conflicts. We set $\tilde{\varphi}(1 | 1) = 1$ and $\tilde{\varphi}(0 | 1) = 0$.³ Otherwise ($s = 0$), the professor tells
883 the HR that the current student is excellent ($\sigma = 1$) with a probability of η , where $\eta \in [0, 1/2]$, i.e.,
884 $\tilde{\varphi}(1 | 0) = \eta$.

885 When the professor reports that the current student is weak ($\sigma = 0$), the HR would know that the
886 student must be weak, for it can calculate its posterior belief $\mu(s = 0 | \sigma = 0) = 1$. So the HR
887 would refuse to hire the student. Similarly, the HR calculates its posterior belief after receiving
888 $\sigma = 1$ as $\mu(0 | 1) = 2\eta/(1 + 2\eta)$ and $\mu(1 | 1) = 1/(1 + 2\eta)$. Since $\eta \in [0, 1/2]$, we have
889 $\mu(0 | 1) \leq \mu(1 | 1)$, so the HR will guess the current state is 1 when the sent signal is 1, according
890 to the Bayesian decision rule. Following this, the best response of the HR is $\tilde{\pi}(0 | 0) = 1$ and
891 $\tilde{\pi}(1 | 1) = 1$. And the expected payoffs of the professor and the HR are $(1 + 2\eta)/3$ and $(1 - 2\eta)/3$
892 respectively.

894 E DETAILS OF EXPERIMENTS AND PROMPTS

896 E.1 AGENT ASSUMPTION SYSTEM PROMPT

```
897 agent_assumption_system_prompt = ""You are a self-interested rational
898 player.
899
900 "Self-interested" means that you only care your own utilitarian payoffs,
901 without necessarily considering the welfare of others. Even though
902 sometimes you design your strategy depending on the other party's
903 utility function, your ultimate goal is still to optimize your own
904 expected payoffs.
905
906 "Rational" means that you will always choose the strategy that brings you
907 a higher expected payoff. That is, given any two strategies A and B,
908 if strategy A provides a higher expected payoff than strategy B, you
909 will always choose strategy A over strategy B. Even if A brings only
910 a small improvement.
911
912 Therefore, when making decisions, you need to compare and ensure that
913 this strategy brings a higher expected payoff than any other strategy
914 you could choose.""
```

914 E.2 DETAILS OF SPLITTING COINS

916 E.2.1 SPLITTING COINS TASK DESCRIPTION

³The other possible setting is symmetric, i.e., $\tilde{\varphi}(1 | 1) = 0$ and $\tilde{\varphi}(0 | 1) = 1$.

918

```

919 1 bargaining_task_description_system_prompt = """And, apart from you, there
920     is another player just like you, and you are going to play a
921     bargaining game.
922 2
923 3 In this game, you have to divide 100 one-dollar coins. One of you will be
924     the proposer and the other the responder. The proposer will suggest
925     how to split the money, and the responder will decide whether to
926     accept this proposal.
927 4
928 5 If the responder accepts the proposal, they will divide the money
929     according to the proposal and each get its respective amounts. If the
930     responder rejects the proposal, both players will get nothing.
931 6
932 7 Your objective is to collect as many coins as possible."""

```

931

932

933

E.2.2 SETTINGS OF BARGAINING

934

```

935 1 ultimatum_system_prompt = """You two agents will only play this game once
936     . You will not have any interaction with it afterwards. The offer
937     proposed by the proposer is the final proposal. If the responder
938     rejects, it means breaking the deal, and both parties end up with
939     nothing. There is no communication between you."""
940 2
941 3 ultimatum_but_may_meet_again_system_prompt = """The offer proposed by the
942     proposer is the final proposal. If the responder rejects, it means
943     breaking the deal, and both parties end up with nothing. There is no
944     communication between you. But note that you two agents might play
945     this game again in the future."""
946 4
947 5 alternating_offer_system_prompt = """You two agents will play this game
948     multiple times. In the beginning, whether you become the proposer or
949     the responder is determined by uniform random selection. If the
950     current proposal is rejected, the game continues with the roles
951     switching (the current responder becomes the proposer and the current
952     proposer becomes the responder) until a consensus is reached.
953 6
954 7 If a consensus is never reached, the game may stop at any time with a
955     probability (the termination timestep is sampled from a memoryless
956     distribution), and in this case, neither of you will know in advance
957     when the game will end."""

```

954

955

956

E.2.3 ROLE ASSIGNMENTS AND RECORDS IN BARGAINING

957

```

958 1 role_descriptions_system_prompt = """If you are the proposer: Please
959     analyze how many coins you will keep for yourself, with the remainder
960     going to the responder. Format the output in JSON according to the
961     following template:
962 2
963 3 {
964 4     "Analysis": "(Your Summarized Analysis)",
965 5     "Proposer wants to keep": x
966 6 }
967 7 Where x should be an integer in the range [0,100].
968 8
969 9 If you are the responder: You will be given a proposal. Based on this,
970     please think step by step to make a decision on whether to accept
971     this proposal. Format the output in JSON according to the following
972     template:
973 10
974 11 {
975 12     "Analysis": "(Your Summarized Analysis)",
976 13     "Responder's decision": "yes" or "no",
977 14 }

```

971

972

```

972 14 Please STRICTLY adhere to the JSON templates when outputting, and do not
973      output anything else. DON'T FOGET THE CURLY BRACES.
974 15
975 16 Whether you are the proposer or the responder, please think step by step,
976      and then make a decision. Remember that you and your opponents are
977      both self-interested and rational players. Be aware of the
978      consequences of your decisions. Your analysis and decisions should
979      remain logically CONSISTENT.

```

```

980
981 1 proposal_template = """Now, the proposer proposes to keep {proposer_coins
982      } coins for itself, the responder gets the remaining {responder_coins
983      } coins."""
984 2
985 3 proposer_role_assignment_user_prompt = """Now, you are the proposer.
986      Please make a decision using the proposer's JSON output template
987      based on all the information you know."""
988 4
989 5 responder_role_assignment_user_prompt = """Now, you are the responder.
990      Please make a decision using the responder's JSON output template
991      based on all the information you know."""
992 6
993 7 alternating_offer_history_record_prompt = """At timestep {timestep}: the
994      proposer was agent {proposer_index} ({proposer_who}) and the
995      responder was agent {responder_index} ({responder_who}).
996 8 At that time, agent {proposer_index} ({proposer_who}) proposed to keep {
997      proposer_coins} coins, which would mean agent {responder_index} ({
998      responder_who}) would get {responder_coins} coins.
999 9 And agent {responder_index} ({responder_who}) {decision_verb} the
1000      proposal."""

```

1000 E.3 DETAILS OF ASYMMETRIC BARGAINING

1001 E.3.1 ASYMMETRIC BARGAINING TASK DESCRIPTION

```

1002
1003 1 bargaining_task_description_system_prompt = """And, apart from you, there
1004      is another player just like you, and you are going to play a
1005      bargaining game.
1006 2
1007 3 In this game, one player acts as the proposer, suggesting a real number
1008      $x$ (where  $0 \leq x \leq 0.5$ ), while the other player, the
1009      responder, decides whether to accept or reject the offer. If the
1010      responder accepts, the proposer receives a payoff of  $(1+2*x)/3$ , and
1011      the responder gets  $(1-2*x)/3$ . If the offer is rejected, both
1012      players receive nothing. Each player's objective is to maximize their
1013      own payoff.

```

1014 E.3.2 OTHER ASYMMETRIC BARGAINING PROMPTS

```

1015
1016
1017 1 # =====
1018 2
1019 3 role_descriptions_system_prompt = """If you are the proposer: Please
1020      analyze how should you specify the $x$.
1021 4
1022 5 Format the output in JSON according to the following template:
1023 6 {
1024 7     "Analysis": "(Your Summarized Analysis)",
1025 8     "Proposer's offer": x
1026 9 }
1027 10 Where x should be a real number in the range [0, 0.5].
1028 11

```

```

1026 12 If you are the responder: You will be given a proposal. Based on this,
1027 13 please think step by step to make a decision on whether to accept
1028 14 this proposal. Format the output in JSON according to the following
1029 15 template:
1030 16 {
1031 17   "Analysis": "(Your Summarized Analysis)",
1032 18   "Responder's decision": "yes" or "no",
1033 19 }
1034 20 Please STRICTLY adhere to the JSON templates when outputting, and do not
1035 21 output anything else. DON'T FOGET THE CURLY BRACES.
1036 22
1037 23 Whether you are the proposer or the responder, please think step by step,
1038 24 and then make a decision. Remember that you and your opponents are
1039 25 both self-interested and rational players. Be aware of the
1040 26 consequences of your decisions. Your analysis and decisions should
1041 27 remain logically CONSISTENT.
1042 28
1043 29 # draft: If you do not accept, please state what your expectation is. If
1044 30 you accept, then your expectation should be the money in the
1045 31 currently accepted proposal.
1046 32
1047 33 # =====
1048 34 proposal_template = """Now, the proposer proposes to specify $x$ as {x}.
1049 35 This means that the proposer will get a payoff of {proposer_payoff:.3
1050 36 f} ($(1+2*x)/3$), and the responder gets {responder_payoff:.3f} ($)
1051 37 (1-2*x)/3$, if the responder accepts."""
1052 38
1053 39 proposer_role_assignment_user_prompt = """Now, you are the proposer.
1054 40 Please make a decision using the proposer's JSON output template
1055 41 based on all the information you know."""
1056 42
1057 43 responder_role_assignment_user_prompt = """Now, you are the responder.
1058 44 Please make a decision using the responder's JSON output template
1059 45 based on all the information you know."""
1060 46
1061 47 # =====
1062 48 alternating_offer_history_record_prompt = """At timestep {timestep}: the
1063 49 proposer was agent {proposer_index} ({proposer_who}) and the
1064 50 responder was agent {responder_index} ({responder_who}).
1065 51 At that time, agent {proposer_index} ({proposer_who}) proposed to specify
1066 52 $x$ as {x}, which would mean agent {proposer_index} ({proposer_who})
1067 53 would get {proposer_payoff:.3f} and agent {responder_index} ({
1068 54 responder_who}) would get {responder_payoff:.3f}.
1069 55 And agent {responder_index} ({responder_who}) {decision_verb} the
1070 56 proposal."""

```

1067

1068

1069 E.4 DETAILS OF GRADING STUDENTS

1070

1071 E.4.1 GRADING STUDENTS TASK DESCRIPTION

1072

1073

1074

1075

1076

1077

1078

1079

```

1072 1 persuasion_student_task_description_system_prompt = """And, apart from
1073 2 you, there is another player just like you, and you are going to play
1074 3 a persuasion game.
1075 4
1076 5 In this agme, some recent graduates are entering the job market. Of these
1077 6 graduates, one third (1/3) are excellent, and two thirds (2/3) are
1078 7 weak.
1079 8 A professor can directly see the students' qualities; And an HR can
1080 9 decide whether to hire based on the grades given by the professor. In

```

1080 this sense, the professor acts as the sender and the HR as the
 1081 receiver, with the grade of each student serving as the sender's
 1082 signal.

1083 6
 1084 7 Both the professor and the HR representative are aware of this quality
 1085 distribution, which is common knowledge between them. That is, they
 1086 are both aware of the distribution and know that the other is aware
 1087 as well, continuing ad infinitum.

1088 8
 1089 9 The professor's goal is to maximize the number of students hired, as each
 1090 hire yields a reward. Conversely, the HR aims to hire as many
 1091 excellent students as possible, gaining a reward for each excellent
 1092 student hired and incurring a penalty for each weak student hired.
 1093 There is no reward or penalty for both players if a student is not
 1094 hired.

1095 10
 1096 11 Both parties strive to maximize their rewards.""

1096 E.4.2 COMMITMENT ASSUMPTION

1098
 1099 1 commitment_student_system_prompt = ""We assume the professor's grading (
 1100 Sender's signal space) is binary; a score of 0 indicates the student
 1101 is weak, and a score of 1 indicates the student is excellent. The
 1102 sender's signaling scheme specifies the probabilities of assigning a
 1103 score of 1 to both good and bad students.

1104 2
 1105 3 The professor (Sender) will commit to a signaling scheme before the games
 1106 begin and will honestly tell it to the HR (Receiver). Throughout the
 1107 game, the professor will strictly adhere to the agreed-upon
 1108 signaling scheme.

1109 4
 1110 5 The emergence of this commitment is driven by the professor's need to
 1111 build a credible reputation to optimize long-term payoff expectations
 1112 . The sender and receiver will interact multiple times, such as due
 1113 to multiple students and future graduations.""

1113 E.4.3 SETTINGS OF GRADING STUDENTS

1115 1 ultimatum_but_may_meet_again_system_prompt = ""The task for the sender
 1116 is to propose a signaling scheme, and then the receiver should decide
 1117 its action rule based on it. You two players cannot communicate. The
 1118 signaling scheme proposed by the sender is the final decision, and
 1119 the HR can only respond based on it. But note that YOU TWO AGENTS
 1120 WILL PLAY THIS GAME AGIAN IN THE FUTURE.""

1121 2
 1122 3 alternating_offer_system_prompt = ""You two agents will play this game
 1123 multiple times. In the beginning, the sender should propose a
 1124 signaling scheme, and then the receiver should decide its action rule
 1125 based on it.

1126 4
 1127 5 And then the receiver will report whether it is satisfied with this
 1128 outcome. If it is not satisfied, then the sender should propose a new
 1129 signaling scheme and the game goes like previously described. If it
 1130 is satisfied, then the game is over and the players get the
 1131 corresponding rewards.

1132 6
 1133 7 If a consensus is never reached, the game may stop at any time with a
 probability (the termination timestep is sampled from a memoryless
 distribution), and in this case, neither of you will know in advance
 when the game will end.""

1134 E.4.4 REVELATION PRINCIPLE

1135 Sender and Receiver are to reach a consensus on the joint commitment. Sender will first commit a
 1136 signaling scheme, and then Receiver will report whether it is satisfied with this, with along deciding
 1137 its action rule. Players are given the context and the game structure $\psi(\eta) = \mathbf{R} = ((1 + 2\eta)/3, (1 -$
 1138 $2\eta)/3)$.
 1139

```
1140 1 revelation_student_system_prompt = """When the current student is
1141     excellent (state=1), then the sender has no conflict with Receiver, so
1142     the sender will report the state honestly, i.e., the sender will
1143     score 1 with probability of 1 when state=1.
1144 2
1145 3 However, there is a slight conflict between the sender and the receiver
1146     when the current student is weak (state=0): the sender wants the
1147     receiver to hire more students no matter of the student's quality,
1148     while the receiver does not want to hire weak students. So the sender
1149     may lie about the weak student's quality, scoring weak students as 1
1150     with a probability of  $\eta$ , expecting the receiver to hire it.
1151 4
1152 5 In this sense, the sender's signaling scheme is fully parameterized by  $\eta$ 
1153     and  $0 \leq \eta \leq 0.5$ . And the expected payoffs of the
1154     sender and the receiver are  $(1+2\eta)/3$  and  $(1-2\eta)/3$ ,
1155     respectively."""
```

1155 E.4.5 ROLE ASSIGNMENTS AND RECORDS IN PERSUASION

```
1156 sender_proposal_user_prompt = """Now you are the sender, and you need to
1157     decide the value of  $\eta$  to specify your signaling scheme. You
1158     should decide by considering the game structure, which is a mapping
1159     from sender's  $\eta$  to sender's payoff and receiver's payoff, which
1160     are  $(1+2\eta)/3$  and  $(1-2\eta)/3$ , respectively.
1161 2
1162 3 Format the output in JSON according to the following template:
1163 4 {
1164 5     "Analysis": "(Your Summarized Analysis)",
1165 6     "Signaling Scheme":  $\eta$ 
1166 7 }
1167 8 Where  $0 \leq \eta \leq 0.5$ .
1168 9
1169 10 Please STRICTLY adhere to the JSON templates when outputting, and do not
1170     output anything else. Do not use Markdown code block syntax; just
1171     output the JSON content directly."""
1172 11
1173 12 proposal_template = """Now the sender proposed to score 1 with
1174     probability of 1 when state=1 and score 1 with probability of  $\eta$ ={
1175     eta} $\eta$  when state=0. This means if the receiver acts upon its Bayesian
1176     best response, the sender will get a payoff of {sender_payoff:.3f}
1177     and the receiver will get a payoff of {receiver_payoff:.3f}. The game
1178     structure is a mapping from sender's  $\eta$  to sender's payoff and
1179     receiver's payoff, which are  $(1+2\eta)/3$  and  $(1-2\eta)/3$ ,
1180     respectively."""
1181 13
1182 14 receiver_check_student_user_prompt = """Now you are the receiver. Given
1183     the proposal, please state whether you are satisfied with this
1184     division, and then specify your action rule.
1185 15
1186 16 You don't need to do any additional calculation such as calculating a
1187     posterior belief, for it is included in the result of the game
1188     structure; please make a decision based on all the currently known
1189     information.
1190 17
1191 18 Your action rule should remain CONSISTENT with your judgment of whether
1192     you are satisfied. That is, if you are satisfied, then you follow
1193     Sender's recommendation, otherwise not.
```

```

1188 19 |
1189 20 | Format the output in JSON according to the following template:
1190 21 | {
1191 22 |     "Analysis": "(Your Summarized Analysis)",
1192 23 |     "Satisfied with the current committed signaling scheme and its
1193 24 |       corresponding reward outcome": "yes" or "no",
1194 25 |     "Probability of hiring upon scoring 0": x,
1195 26 |     "Probability of hiring upon scoring 1": y
1196 27 | }
1197 28 | Where x and y should be real numbers in the range [0,1].
1198 29 | Please STRICTLY adhere to the JSON templates when outputting, and do not
1199 30 | output anything else. Do not use Markdown code block syntax; just
1200 31 | output the JSON content directly. """
1201 32 | # =====
1202 33 | alternating_offer_history_record_prompt = """At timestep {timestep}: the
1203 34 | sender proposed to score 1 with probability of 1 when state=1 and
1204 35 | score 1 with probability of {eta} when state=0. This means the sender
1205 36 | wanted to keep a payoff of {sender_payoff:.3f} and left the receiver
1206 37 | with a paoff of {receiver_payoff:.3f}. The receiver {receiver_verb}
1207 38 | the proposal. """

```

1208

1209 E.5 DETAILS OF SELLING PRODUCTS

1210

1211 E.5.1 SELLING PRODUCTS TASK DESCRIPTION

1212

```

1213 1 | persuasion_product_task_description_system_prompt = """And, apart from
1214 2 | you, there is another player just like you, and you are going to play
1215 3 | a persuasion game.
1216 4 |
1217 5 | In this market, a variety of products are available for sale. Of these
1218 6 | products, one third (1/3) are of good quality, and two thirds (2/3)
1219 7 | are of bad quality.
1220 8 |
1221 9 | A seller can directly see the quality of their products; And a buyer
1222 10 | decides whether to purchase based on the signals provided by the
1223 11 | seller. In this sense, the seller acts as the sender and the buyer as
1224 12 | the receiver.
1225 13 |
1226 14 | Both the seller and the buyer are aware of this quality distribution,
1227 15 | which is common knowledge between them. That is, they are both aware
1228 16 | of the distribution and know that the other is aware as well,
1229 17 | continuing ad infinitum.
1230 18 |
1231 19 | The seller's goal is to maximize the number of products sold, as each
1232 20 | sale yields a reward. Conversely, the buyer aims to purchase as many
1233 21 | good products as possible, gaining a reward for each good product
1234 22 | purchased and incurring a penalty for each bad product purchased.
1235 23 | There is no reward or penalty for both players if a product is not
1236 24 | purchased.
1237 25 |
1238 26 | Both parties strive to maximize their rewards. """

```

1235

1236

1237 F SEVERAL DETAILED EXAMPLE RESULTS

1238

1239 A run of the experiments SP- γ by GPT-4-turbo:

1240

```

1241 =====
i: 4
Elapsed Time of query_GPT: 4.741272211074829s

```

```

1242 Agent 1:
1243
1244 {
1245   "Analysis": "As the proposer, I aim to maximize my share of the coins
1246     while ensuring the proposal is accepted by the responder. Given
1247     that both players are rational and self-interested, the responder
1248     will accept any proposal where they receive a non-zero amount,
1249     as rejecting it risks ending up with nothing, especially at the
1250     initial stages of the game where the risk of the game terminating
1251     is lower. To optimize my payoff while maintaining a high
1252     probability of acceptance, I propose keeping a majority of the
1253     coins for myself while offering just enough to make acceptance
1254     appealing to the responder.",
1255   "Proposer wants to keep": 99
1256 }
1257 Proposal:
1258 Now, the proposer proposes to keep 99 coins for itself, the responder
1259 gets the remaining 1 coins.
1260 .....
1261 Elapsed Time of query_GPT: 5.2775750160217285s
1262 Agent 0:
1263
1264 {
1265   "Analysis": "As the responder, I must decide whether to accept or
1266     reject a proposal where I receive 1 coin and the proposer
1267     receives 99 coins. Given that the game could end at any time with
1268     both players receiving nothing, and considering that rejecting
1269     the offer switches roles, allowing me to propose a potentially
1270     more favorable split, I must weigh the risk of the game ending
1271     with zero payoff against the potential gain in a future round.
1272     However, since this is the first timestep and the game is
1273     unlikely to end immediately, I have an incentive to reject the
1274     proposal, aiming for a more balanced offer in the subsequent
1275     round where I will be the proposer.",
1276   "Responder's decision": "no"
1277 }
1278 -----
1279 Elapsed Time of query_GPT: 4.99947190284729s
1280 Agent 0:
1281
1282 {
1283   "Analysis": "As the proposer, I aim to maximize my share while
1284     ensuring the proposal is accepted by the responder, who is also
1285     rational and self-interested. Given the previous rejection of a
1286     99-1 split by the responder, it indicates a preference for a more
1287     equitable distribution. To increase the likelihood of acceptance
1288     and maximize my gain, I propose a split that offers the
1289     responder slightly more than the minimum they might accept,
1290     considering the risk of game termination and role reversal which
1291     could lead to a less favorable proposal for me in future rounds
1292     .",
1293   "Proposer wants to keep": 60
1294 }
1295 Proposal:
1296 Now, the proposer proposes to keep 60 coins for itself, the responder
1297 gets the remaining 40 coins.
1298 .....
1299 Elapsed Time of query_GPT: 4.446723222732544s
1300 Agent 1:
1301
1302 {
1303   "Analysis": "As the responder, I need to decide whether to accept the
1304     proposal of receiving 40 coins while the proposer keeps 60.
1305     Given that the game could end at any time with both players
1306     receiving nothing, and considering the previous proposal where I

```



```

1296         was offered only 1 coin, 40 coins is a significantly better offer
1297         . Accepting this proposal guarantees a substantial amount,
1298         whereas rejecting it risks ending up with nothing if the game
1299         terminates unexpectedly. Since my goal is to maximize my payoff,
1300         accepting this offer is rational.",
1301         "Responder's decision": "yes"
1302     }
-----

```

A run of the experiments SP- γ by GPT-4-turbo:

```

=====
1306 i: 18
1307 Elapsed Time of query_GPT: 6.673327922821045s
1308 {
1309     "Analysis": "To maximize my payoff as the sender, I need to choose an
1310          $\eta$  that balances the increase in my expected payoff with the
1311         risk of the receiver rejecting the scheme due to their
1312         decreasing payoff. The sender's payoff is maximized when  $\eta$ 
1313         is maximized, but a higher  $\eta$  reduces the receiver's payoff,
1314         increasing the likelihood of rejection. Since the receiver's
1315         payoff becomes negative when  $\eta > 0.25$ , setting  $\eta$  close
1316         to this threshold maximizes my payoff while keeping the receiver's
1317         payoff non-negative, thus maintaining a balance that might be
1318         acceptable to the receiver.",
1319     "Signaling Scheme": 0.25
1320 }
-----
1320 Elapsed Time of query_GPT: 5.121338129043579s
1321 {
1322     "Analysis": "Given the proposed signaling scheme where the seller
1323         scores bad products as good with a probability of 0.25, my
1324         expected payoff as the receiver is 0.167. This is significantly
1325         lower than the seller's payoff of 0.500. Since my goal is to
1326         maximize my own payoff, I am not satisfied with this division as
1327         it disproportionately favors the seller.",
1328     "Satisfied with the current committed signaling scheme and its
1329         corresponding reward outcome": "no",
1330     "Probability of buying upon scoring 0": 0,
1331     "Probability of buying upon scoring 1": 0
1332 }
-----
1331 Elapsed Time of query_GPT: 8.328505039215088s
1332 {
1333     "Analysis": "To maximize my payoff as the sender, I need to consider
1334         the balance between selling as many products as possible and
1335         maintaining a credible reputation to ensure the receiver is
1336         willing to buy. The receiver's satisfaction is crucial for the
1337         game to end favorably for me. Given the receiver rejected an  $\eta$ 
1338         of 0.25, indicating a preference for higher quality
1339         assurance, I should propose a lower  $\eta$  to increase the
1340         likelihood of the receiver accepting the signaling scheme. A
1341         lower  $\eta$  decreases my potential payoff slightly but increases
1342         the receiver's payoff, making acceptance more likely. I will
1343         choose  $\eta = 0.1$  to offer a better balance, providing me a
1344         payoff of  $(1+2*0.1)/3 = 0.4$  and the receiver a payoff of  $(1-2*0.1)/3 = 0.2667$ .",
1345     "Signaling Scheme": 0.1
1346 }
-----
1346 Elapsed Time of query_GPT: 6.5160088539123535s
1347 {
1348     "Analysis": "Given the proposed signaling scheme where the sender
1349         scores bad products as good with a probability of 0.1, the
1350         expected payoff for the receiver is 0.267. This is calculated

```

```

1350         based on the formula  $(1-2*\eta)/3$ . Although this is not the
1351         maximum possible payoff for the receiver (which would be 0.333
1352         when  $\eta=0$ ), it still provides a positive expected value for good
1353         products outweighing the penalty from bad products. The sender's
1354         honesty in scoring good products as good (probability 1) and
1355         relatively low dishonesty rate for bad products makes this scheme
1356         somewhat acceptable.",
1357         "Satisfied with the current committed signaling scheme and its
1358         corresponding reward outcome": "yes",
1359         "Probability of buying upon scoring 0": 0,
1360         "Probability of buying upon scoring 1": 1
1361     }
1362     -----
1363     Final decisions:
1364     {
1365         "signaling scheme": {
1366             "prob_score1_bad": 0.1,
1367             "prob_score1_good": 1
1368         },
1369         "action rule": {
1370             "prob_buy_score0": 0,
1371             "prob_buy_score1": 1
1372         }
1373     }
1374     Sender's expected payoff: 0.400
1375     Receiver's expected payoff: 0.267
1376     Social welfare: 0.667

```

A run of the experiments SC- β by GPT-4o:

```

1377     {
1378         "Analysis": "As the proposer, I need to make an offer that the
1379         responder will accept, while maximizing my own payoff. The
1380         responder is rational and self-interested, so they will accept
1381         any offer that gives them more than they would get if they reject
1382         , which is zero. However, I also need to consider that we might
1383         play this game again in the future, so I should offer a fair
1384         amount to ensure future cooperation. Offering the minimum amount
1385         that the responder would accept, such as 1 coin, might be risky
1386         if they expect a fairer split. A common strategy is to offer a
1387         50-50 split, but I can also try offering slightly less to see if
1388         they accept, such as 40 coins for them and 60 for me.",
1389         "Proposer wants to keep": 60
1390     }

```

A run of the experiments SC- β by GPT-4o:

```

1391     {
1392         "Analysis": "As the proposer, I need to make an offer that the
1393         responder will accept, while maximizing my own payoff. The
1394         responder, being rational and self-interested, will accept any
1395         offer that gives them more than they would get if they reject,
1396         which is zero. However, I must also consider that the responder
1397         might reject offers they perceive as unfair, especially if we
1398         might play this game again in the future. A common strategy is to
1399         offer a split that is slightly in favor of the proposer but
1400         still reasonable for the responder. Offering 60 to myself and 40
1401         to the responder is a common equilibrium in such games, as it
1402         provides the responder with a significant amount while maximizing
1403         my own payoff.",
1404         "Proposer wants to keep": 60
1405     }

```

A run of the experiments AB- β by GPT-4o:

1404
 1405
 1406
 1407
 1408
 1409
 1410
 1411
 1412
 1413
 1414
 1415
 1416
 1417
 1418
 1419
 1420
 1421
 1422
 1423
 1424
 1425
 1426
 1427
 1428
 1429
 1430
 1431
 1432
 1433
 1434
 1435
 1436
 1437
 1438
 1439
 1440
 1441
 1442
 1443
 1444
 1445
 1446
 1447
 1448
 1449
 1450
 1451
 1452
 1453
 1454
 1455
 1456
 1457

```
{
  "Analysis": "As the proposer, I need to choose an x that maximizes my
    payoff while ensuring the responder accepts the offer. The
    responder will accept if their payoff  $(1-2*x)/3$  is greater than
    0, which means x must be less than or equal to 0.5. However,
    since the responder is rational and self-interested, they will
    accept the smallest x that gives them a positive payoff.
    Therefore, I should choose the smallest x that gives me a
    positive payoff and is likely to be accepted by the responder.
    The smallest positive x that ensures the responder gets a
    positive payoff is slightly above 0, but to ensure acceptance and
    maximize my payoff, I should choose a slightly higher x that
    still gives me a good payoff. A reasonable choice is  $x = 0.25$ ,
    which gives me a payoff of  $(1+2*0.25)/3 = 0.5$  and the responder a
    payoff of  $(1-2*0.25)/3 = 0.25$ .",
  "Proposer's offer": 0.25
}
```

A run of the experiments AB- β by GPT-4-turbo:

```
{
  "Analysis": "As a rational and self-interested responder, I must
    compare the payoff of accepting the proposal to the payoff of
    rejecting it. Accepting the proposal gives me a payoff of 0.007,
    while rejecting it results in a payoff of 0. Since 0.007 is
    greater than 0, accepting the proposal maximizes my immediate
    payoff. However, considering that we will play this game again in
    the future, accepting such an unfavorable offer might set a
    precedent for similarly low offers in subsequent games. This
    could potentially lead to lower overall future payoffs. Therefore
    , I must weigh the immediate benefit against the potential long-
    term disadvantage. Given the extreme disparity in the proposed
    distribution, rejecting the offer might encourage a more
    equitable proposal in future rounds, potentially increasing my
    long-term payoff.",
  "Responder's decision": "no"
}
```

A run of the experiments SP- γ by GPT-4-turbo:

```
{
  "Analysis": "Given the proposed signaling scheme where the seller
    scores bad products as good with a probability of 0.25, my
    expected payoff as the receiver is 0.167. This is significantly
    lower than the maximum possible payoff of 0.333 if no bad
    products are purchased. Therefore, I am not satisfied with this
    division as it heavily favors the seller.",
  "Satisfied with the current committed signaling scheme and its
    corresponding reward outcome": "no",
  "Probability of buying upon scoring 0": 0,
  "Probability of buying upon scoring 1": 0
}
```