3.1 Thrust 1: Designing Information in Machine-in-the-Loop Decision Making

In this thrust, we plan to develop algorithmic frameworks for decision making with machines in the loop. As the main theme of this research proposal, we will explore the usage of *information design*, i.e., what types of information disclosing policy should ML choose to provide to human decision makers. For example, consider a ML-assisted navigation system, in which ML provides information about traffic to humans and humans can then decide what route to take based on both their own prior (previous experience and knowledge) and ML-provided information. The goal of ML is to optimize some pre-specified objective (e.g., minimizing the total transit time) while considering the human driver's response to the information (e.g., drivers might rely on their prior experience on the traffic more than the ML-provide information). Our focus will be on designing the information structure to present to humans. The information can be structured either as a recommendation of action (e.g., which route to take) or a distribution of signals conditional on the realization of the world state (e.g., the estimated time for the particular route given the traffic).[1]

To formulate this information design framework, we will start with the full information setting in which we assume human behavior models are known. This enables us to develop a game theoretical framework and formulate the information design as a bi-level optimization problem, where ML is choosing the optimal information design while considering humans optimizing their actions conditional on the provided information (Task 1.1). We will then relax the full-knowledge assumption and investigate the associate learning (Task 1.2) and robust design problems (Task 1.3).

Prior work. The proposed activities in this thrust will be built on my prior works. In particular, the information design problem can be formulated as a *Stackelberg game*, in which ML first decides on the information disclosing policy, and humans decide on what decision to take based on the provided information. My prior works have explored Stackelberg games in a range of different domains. I have studied problems of contract design [45], in which the firm posts a contract and then the worker decides on the amount of effort in response to the contract, learning with strategic responses [92], in which the learner posts a decision rule and then the agent responds with the goal of receiving a favorably treatment, and Bayesian persuasion [24, 87], in which the sender decides an information disclosure strategy to persuade the receiver to take certain actions. In particular, my most relevant work in the persuasion setting aligns well with this research if we consider the sender as ML and the receiver as human decision makers.

3.1.1 Task 1.1: Develop an optimization framework for information design

In this task, we aim to develop an information design framework that accounts for different human models.

Setting up the optimization framework. We consider the setting in which ML decides on a information policy and then the human decision maker take action based on both her prior information and the ML-provided information. Below we use ML and human to denote the two roles respectively. We extend the classical information design framework by Kamenica and Gentzkow [52] (that only consider cases when humans are Bayesian rational) to account for different behavioral models. Formally, let the state of the world be θ which is drawn from a finite set Θ according to a prior distribution $\mu_0 \in \Delta(\Theta)$. Let τ be the information scheme ML chooses. Upon receiving the realization of the signal σ based on the information scheme, human can choose an action a from an action set A. To incorporate human behavioral models, we specify human behavior with two functions, a belief updating function $\omega(\mu_0, \sigma)$, which denotes the posterior distribution induced by the signal σ and prior μ_0, and a decision function $P(a|\omega(\mu_0, \sigma))$, characterized by a distribution of decisions given the posterior. Let $V(a, \theta)$ be ML's utility when the receiver takes action a and the state of the world be θ. In this task, we assume $V(a, \theta)$ is given and known. We will address how to design this objective function in Thrust 3.[2] With these notations, the information design problem can be

[1] We note that our focus is on the design of *information structure*. While the presentations of information, such as layout, color, font size, or language usages, are also important aspects of information design, they are not the focus of this proposed research.

[2] In some domains, such as recommendation system, the utility could be easier specified (e.g., the utility could be 1 when

formulated as follows: $\max_\tau \ \mathbb{E}_{\theta,\sigma\sim\tau}[\sum_a P(a|\omega(\mu_0,\sigma))V(a,\theta)]$.

Following the seminal work by Kamenica and Gentzkow [52], without loss of generality, we can limit the space of information structure to be the distributions of posteriors, i.e., $\tau \in \Delta(\Delta(\Theta))$, with the constraint that the induced posterior need to be *plausible*, i.e., $\tau \in \mathcal{K}$, where $\mathcal{K} = \{\tau \in \Delta(\Delta(\Theta)) : \exists\tau$ such that $\forall\mu \in \mathsf{supp}(\tau), \mu = \omega(\mu_0,\sigma)\}$. Note that, when humans are Bayesian, this constraint reduces to *Bayesian-plausibility*, i.e., $\mathbb{E}_{\mu\sim\tau}[\mu] = \mu_0$. Therefore, the information design problem can also be written as

$$\max_{\tau\in\Delta(\Delta(\Theta))} \ \mathbb{E}_{\theta,\mu\sim\tau}\left[\sum_{a\in\mathcal{A}} P(a|\mu)V(a,\theta)\right] \quad \text{s.t.} \quad \tau \in \mathcal{K}. \tag{2}$$

Research questions. With the optimization formulation in place, in this task, we aim to characterize and explore the information design problem with different human models. First, consider the classical case that humans are Bayesian rational, the decision function $P(a|\omega(\mu_0),\tau))$ is a delta function that puts all the probability mass on the action that maximize the receiver's payoff. When putting this decision function back to the optimization problem, the objective is non-continuous and the optimization is in general NP-hard to solve. On the other hand, when we consider the discrete choice model (let $\mu = \omega(\mu_0,\tau)$), the decision function is in the form of: $P(a|\mu) = \frac{\exp\left(\beta\widehat{u}^R(a|\mu)\right)}{\sum_{a'}\exp\left(\beta\widehat{u}^R(a'|\mu)\right)}$, which is a continuous softmax function. With this human behavioral model, the information design problem can be formulated as a convex optimization problem, and there exist efficient algorithms to find the optimal information design when the space of the information design is small. The above discussion highlights the need to understand how different human behavior models impact the problem of optimal information design. In this task, we will assume the knowledge of human behavioral models and address the corresponding optimization problem. Depending on the human models, there will be two types of optimization problems to be addressed:

- Optimization with non-continuous objective: When the human decision model follow the expected utility theory or prospect theory (and possibly other variants), since human decision making will be in the form of choosing an action that maximizes the (possibly distorted) payoff function, the objective of the optimization problem will be non-continuous, and we cannot directly apply the standard first-order methods to solve the optimization problem. In this type of problem, we plan to utilize the techniques from recent research efforts in algorithmic persuasion [27, 29, 6] to characterize the equilibrium solution and the computational complexity. On a high-level, this line of approach often involves utilizing the duality theory to characterize the properties of the optimal solution. The characterizations can help reduce the search space for optimal solutions and make the optimization more efficient.
- Optimization with continuous objective: When the human decision model follow the discrete choice model or other models that lead to stochastic decision making, the optimization objective can usually be written as a continuous differentiable function. This enables the first-order optimization methods, such as gradient descent, to be applied. In this type of problems, we plan to characterize the computational complexity and convergence to the optimal solution with different human models of belief updating and decision making. My prior work on the study of complexity and convergence rate of secure convex optimization [91] will serve as the technical foundation for addressing this problem.

3.1.2 Task 1.2: Design information with uncertain human behavior: A bandit approach

In the previous task, we start our investigation by assuming full knowledge of human behavior. While this assumption could be approximately satisfied when we have access to an abundant of human behavior data, it is generally a strong assumption that might not hold in practice. In this task, we consider the scenario

users follow the recommendation and 0 otherwise). However, in domains with more complex objectives, such as homelessness prevention, finding the right objective is a challenging tasks, and we propose to explore this issue by including humans in the loop in Thrust 3.

in which ML can repeated interact with humans over time. This scenario provides the opportunity for ML to infer humans' behavior models by utilizing the interaction of previous rounds (e.g., by observing human decisions on the past presented-information) and then update the information disclosure strategy in the future rounds. More formally, we can formulate this as a multi-armed bandit problem [55, 3, 11], with each possible information disclosure strategy as an arm. Formally, at each time $t = 1, ..., T$, ML chooses an information scheme τ_t, human with unknown but consistent behavior reacts to the information, and ML obtains a payoff $f(\tau_t)$. The goal of ML is to adaptively update the information structure to maximize the total payoff. The performance of bandit is often measured in terms of *regret*, defined as $\mathbb{E}[\mathsf{Reg}(T)] = Tf(\tau^*) - \sum_{t=1}^{T} f(\tau_t)$, where τ^* is the optimal solution of the problem (2) assuming the human behavior model is known.

While the bandit formulation provides a nice foundation for our problem, the main challenge of our setting is that there are infinitely many information policies (infinitely many arms), and standard bandit algorithms would not lead to sublinear regret, i.e., it won't converge to the optimal policy. We plan to address this challenge using the technique in my prior work on adaptive contract design [45], in which we aim to find the optimal contract to crowd workers with unknown cost/effort levels by adaptively updating the contracts and observe their performance. The key intuition is that, when we select a contract, the response we obtained from humans not only provide us information about the posted contract but also the similar contracts (i.e., worker performance should be similar with similar payments). Therefore, we can propagate this information to nearby arms and achieve near-optimal learning. We plan to apply similar idea in information design through mapping the information disclosure policy to the contract. More formally, this feasibility learning problem hinges on the condition that posting similar information scheme leads to similar payoff. Let $f(\tau)$ be the objective as defined in optimization (2), let $D(\tau, \tau')$ denote the *distance* between two information schemes. If we can find a Lipschitz constant L such that $|f(\tau) - f(\tau')| \leq L \cdot D(\tau, \tau')$ for all (τ, τ'), we can adapt the techniques of my prior work to reach efficient learning. In this task, we aim to characterize the conditions for the Lipschitz condition to hold, derive the corresponding Lipschitz constants, and develop the bandit algorithms for different human behavioral models.

3.1.3 Task 1.3: Robust information design

We plan to also study settings in which we cannot learn from past interactions, and the goal is to design an information policy that is *robust* to a range of candidate human models. We plan to address this problem by borrowing ideas from robust contract design [13, 19, 68], in which robustness is defined as the worst-case optimal mechanisms, considering all the possible (unknown) actions players can take. We will start by building a connection between a *contract* in contract design and a *information strategy* in our setting. More specifically, since the main challenge here is due to this non-quantifiable uncertainty about humans, we can similarly define robustness based on the worst-case guarantee of human decisions (induced by unknown human behavior models), over all possible decisions humans might take within a set of human models. Our goal is to design *robust optimal* information strategy: the strategy is *robust optimal* if the worst case performance is (weakly) better than that of all other possible strategies. Consider the scenario where the human modeling comes from a function class $\mathcal{H}$, sender is concerned about the robustness of his information policy, he may evaluate her expected payoff from choosing the policy π as:

$$\max_{\tau \in \Delta(\Delta(\Theta))} \min_{(\omega(\cdot,\cdot), P(\cdot)) \in \mathcal{H}} \mathbb{E}_{\theta, \mu \sim \tau} \left[\sum_{a \in \mathcal{A}} P(a|\mu)) V(a, \theta) \right] \quad \text{s.t.} \quad \tau \in \mathcal{K} \tag{3}$$

My prior work [92] on robust learning, which also utilizes the techniques from robust contract design to design robust decision rules for strategic users, will serve as the technical foundation for this task.