(Cost-of-) Information Design

Yaron Azrieli and Shuo Xu*

April 30, 2019

Abstract

We introduce the cost of information design problem where a decision maker acquires information subject to a cost selected by a designer. We show that when restricted to the family of posterior-separable cost functions, the designer achieves the same level of utility as in the case where she herself chooses the information for the decision maker. The designer fails to achieve the first best if the family of cost functions are further restricted to be invariant to the labeling of the states. We show in an example that the designer induces full information at zero cost when the cost functions are multiples of average reduction of Shannon entropy. We also introduce competition to the cost of information design problem where two designers simultaneously select the cost of information and the decision maker only acquires information from one of the designers. We show in an example that the designers exhibit Bertrand-like behaviors and the unique equilibrium is for the designers to induce full information at zero cost.

Keywords: Information Design, Costly Information Acquisition

JEL Classification: D82, D83

*Dept. of Economics, The Ohio State University, 1945 North High Street, Columbus, Ohio 43210, U.S.A.
1 Extended Abstract

When a decision maker faces a choice problem under uncertainty, he may acquire information regarding the realized state from an information designer. Traditional information transmission models assume that the decision maker passively receives information selected by the designer. The designer observes a signal of the realized state and sends a message to the decision maker. How much information the decision maker receives depends on how informative the message is. In many real-world examples, however, the decision maker decides how much information he would like to acquire. The designer influences the decision maker’s informational choice by designing how difficult it is for the decision maker to acquire information. Can the designer achieve the same level of utility as in the case where she herself chooses the information for the decision maker? How costly the designer should make it be for the decision maker to acquire information?

Consider the example of a financial advisor trying to sell a stock to an investor by choosing how costly it is for the investor to acquire information about the quality of the stock. The quality is either high or low. The investor prefers to buy the stock only when it is of high quality. The advisor gets proceeds if the investor buy the stock. If the advisor were to choose information for the investor, the optimal information should induce the investor to have only two possible beliefs. When the investor receives a low message, he is convinced the stock is of low quality. When the investor receives a high message, he is indifferent between buying and not buying. If the advisor can choose whatever cost function she wants, she can simply make it zero costly for the investor to acquire the optimal information and infinite costly to acquire any other information. The investor evaluates the difference between the value of information and the cost of information. His total utility is positive only when he acquires the optimal information. Such cost function is too extreme. Instead, we show that the advisor achieves the same level of utility if the cost functions satisfy natural properties.

We assume that for any two information structures that induce the same distribution over posteriors, the costs of those two information structures are the same. We say a cost function \( C \) is Posterior-Separable (PS) if there is a convex and continuous function \( c : \Delta(\Omega) \to \mathbb{R} \) with \( c(\mu_0) = 0 \) such that \( C(\tau) = \int_{\Delta(\Omega)} c(\mu) d\tau(\mu) \) for every Bayes plausible distribution over posteriors \( \tau \). We denote the family of all PS cost functions as \( C^{PS} \). The cost functions in this family satisfy the following properties. The cost of a distribution over posteriors equals the weighted average of the cost of reaching the beliefs within the support of the distribution. The cost of null information is zero. Any cost function in this family is monotone with respect to the Blackwell informativeness order and is linear in the distribution over posteriors.

We say a family of cost functions implements the first best, if for any choice problem under uncertainty and any utility of the designer, the least upper bound of the guaranteed expected utility of the designer when she chooses a cost function from the family is the same as her expected utility when she herself chooses the information for the decision maker. Notice that when the decision maker is indifferent between several distributions over posteriors, we do not break ties in favor of the designer. Instead, we take a conservative approach and evaluate her payoff at the worst possible case. The first result of the paper is the following:
**Proposition 1.** The family $C^{PS}$ implements the first best.

The idea is to make it zero costly for the investor to acquire the optimal information of the Bayesian Persuasion setting, and prohibitively costly for the investor to acquire further information. If the investor’s value of the optimal information of the Bayesian Persuasion setting is the same as his value of no information, then the worst possible case for the advisor is to get no sales for sure. To increase the probability of getting a sale, the advisor must make it costless for the investor to acquire information that is almost the same but more informative than the optimal information in the Bayesian Persuasion setting, while also making it prohibitively costly for the investor to acquire further information. We show by construction that the advisor’s utility is arbitrarily close to the utility as in the traditional Bayesian Persuasion setting. For the rest choice problems under uncertainty, we show that the advisor achieves the same level of utility.

We use the concavification approach proposed by Aumann and Maschler (1995) and Kamenica and Gentzkow (2011). For any information structure, we focus on its induced distribution over posteriors. For any belief within the convex hull of the support of the distribution, the advisor makes it costless for the investor to reach that belief. For any belief outside the convex hull, the advisor can find an affine function such that it is prohibitively costly for the investor to reach that belief, even if the investor’s expected utility reaches its upper bound given that belief. Such affine functions exist because of the separating hyperplane theorem, and we construct a convex cost function by taking the point-wise maximum of those affine functions. Our construction guarantees the induced distribution over posteriors is unique. Such construction also applies when the action set and the state space have higher finite dimensions.

What will happen to the advisor’s utility if we further restrict the family of cost functions? If the restricted family fails to implement the first best, how costly the advisor should make it be for the investor to acquire information? Within the family of posterior separable cost functions $C^{PS}$, we focus on the family of cost functions that are invariant to the labeling of the states. For example, if there are two states of the world, then any cost function from the family would make it equally costly for the investor to reach the two corner beliefs. The property of symmetry restricts the construction of the optimal cost function. If the investor’s relative expected utility of actions is also invariant to the labeling of the states, then the advisor is forced to induce distributions over posteriors where the posterior beliefs within the support are symmetric. For example, if there are two states of the world $\{L, H\}$, and there are two posteriors beliefs $\mu_L, \mu_H$ within the support of the optimal distribution over posteriors, then for any state $\omega \in \{L, H\}$, the sum of the posterior beliefs of the state equals to 1: $\mu_L(\omega) + \mu_H(\omega) = 1$. If the optimal information in the Bayesian Persuasion setting does not have the above property of symmetry, then the advisor fails to achieve her first-best utility.

For concreteness, we focus on the family of cost functions that are multiples of average reduction of Shannon entropy. A cost function $C$ is a Shannon cost function with multiplicative factor $\kappa$, if $\kappa \geq 0$, and $C(\tau) = \kappa \int_{\Delta(\Omega)} [H(\mu_0) - H(\mu)] d\tau(\mu)$, where $H(\mu) = -\sum_{\omega \in \Omega} \mu(\omega) \ln \mu(\omega)$ is the Shannon entropy of a belief $\mu$. We denote this family by $C^{S}_\kappa$.

The literature on rational inattention pioneered by Sims (2003) and popularized by Caplin and Dean
(2013) and Matějka and McKay (2015) heavily use the family of Shannon entropy cost functions and its generalizations. In the rational inattention literature, the cost of acquiring information is exogenous and arises from the decision maker’s difficulty of processing information. In our study, however, the cost of acquiring information is endogenous and arises from the interaction between the designer and the decision maker. Because there is a conflict of interests between the two players, the designer uses the cost function to align the decision maker’s objective.

Given any cost function from the family $C^S_\kappa$, we can derive the optimal distribution over posteriors chosen by the decision maker, for any choice problem under uncertainty. In this paper, we study a class of choice problems under uncertainty, where there are two states of the world, two actions, and the decision maker tries to match the action with the state. The designer strictly prefers the high action, while under the prior belief, the decision maker chooses the low action. The second result of the paper is the following:

**Proposition 2.** If the belief that the decision maker is indifferent between the two actions is less than or equal to $1/2$, then the designer fails to achieve the first-best utility. The optimal cost function is $C^* = 0$ and the decision maker acquires full information.

We also introduce competition between the designers. Can the decision maker benefit by getting access to more designers? What do the cost functions and the induced distribution over posteriors look like in equilibrium? The designers simultaneously select a function from a common family of cost functions, then the decision maker observes the cost functions, and acquires information only from one of the designers. Notice that this assumption is different from what is typically assumed in the competitive persuasion literature. For example, in Gentzkow and Kamenica (2017), the decision maker receives information from multiple information designers. In our study, if the decision maker is different between several distributions over posteriors, we again take the conservative approach and evaluate the designer’s payoff at the worst possible case. If there is a conflict of interests between the designers, then each designer tries to increase her own payoff by inducing a higher payoff to the decision maker than the other designer. The designers exhibit Bertrand-like behaviors.

We solve an example where there are two states of the world and three actions. Both designers dislike the decision maker’s action given the prior belief, but prefer different actions when the decision maker acquires information. Each designer receives utility of 1 if the decision maker chooses her preferred action and utility 0 otherwise. The designers choose a function from the family of posterior separable cost functions $C^{PS}$. The third result of the paper is the following:

**Proposition 3.** The game has a unique equilibrium $(C_1^*, C_2^*)$ and $C_1^* = C_2^* = 0$.

In this example, every designer can guarantee herself a payoff of at least $1/2$. Since there is a conflict of interests between the designers, if one designer receives utility of strictly greater than $1/2$, then the other designer necessarily receives utility of strictly less than $1/2$. By convexity, if the decision maker’s payoff under some cost function is maximal, then the cost function must be zero. Since the designers’ utilities are symmetric, after the Bertrand-like undercutting, the resulting unique equilibrium is for the designers to induce full information at zero cost.
References


