

Categorical Decision Making by People, Committees, and Crowds

Lav R. Varshney*, Joong Bum Rhim[†], Kush R. Varshney* and Vivek K Goyal[†]

*IBM Thomas J. Watson Research Center, USA

email: {lrvvarshn,krvarshn}@us.ibm.com

[†]Research Laboratory of Electronics, Massachusetts Institute of Technology, USA

email: {jbrhim,vgoyal}@mit.edu

Abstract—Choosing among alternatives is a basic decision problem faced by people in all aspects of life, whether individually or collectively. Results in cognitive science suggest that people perform approximately Bayes-optimal decision making but that cognitive limitations require the coarse categorization of ensembles of problems rather than the application of optimal decision rules on a problem-by-problem basis. These observations motivate the development of a mathematical theory for Bayesian hypothesis testing with quantized prior information.

This paper reviews recent results in minimum Bayes risk quantizer design and its economic implications. In the context of individual decision making, the theory explains differentials in false alarm and missed detection error rates for majority and minority subpopulations without appealing to a taste for discrimination. In group decision making by majority vote, quantizer design becomes a strategic form game. Nash equilibria are guaranteed to exist but often are not Pareto optimal. The analysis reveals precise senses in which a team of agents performs best when it is diverse and shares common goals. Finally, the implications of the theory for crowdsourcing are discussed.

I. INTRODUCTION

Decision making under risk is one of the most basic actions taken by people, either individually or collectively. Decisions include the choice of whether or not to make a job offer to an applicant [1], [2], whether or not to accept a job offer [3], [4], whether or not a software project is in good health [5], whether or not a rule is being violated [6], and whether or not to enter a contest. Such choice can be cast as mathematically equivalent to detecting a signal in noise. Many economic analyses of choice use the decision rule that maximizes expected utility, the likelihood ratio test (LRT). The key parameter of the LRT is the threshold. When an optimal strategy is employed, it is a function of the prior probabilities of the choices as well as the costs of various outcomes [7].

Note that LRTs are not only optimal detection rules and convenient mathematically, but psychology experiments suggest that human decision makers also employ them [8], [9]. Psychology experiments also suggest that humans are able to use prior probabilities in decision rules when available in natural formats [10].

If several decision problems with differing prior probabilities are to be solved, then the optimal strategy is to also set thresholds differently each time. Examples of such scenarios include foul calls by sports referees and arrests by police officers, as each player or citizen has a different prior probability

of committing a foul or crime. Now consider a decision maker or committee of decision makers that categorizes population members according to prior probability, e.g. a police officer using categories like *law-abiding*, *delinquent*, *criminal*, and *nefarious* in the decision to arrest a citizen instead of using the citizen's precise individual prior probability of committing a crime.

There are two motivations for considering decision makers who categorize according to prior probabilities: limited cognitive ability and limited training, both forms of *bounded rationality*. Using a different threshold for different decisions puts much strain on a human decision maker; when decision makers make decisions on members of a population, information processing constraints lead to categorical and coarse thinking [11]–[13]. If the prior probabilities are learned by the decision maker from a finite number of noisy samples, then categorization helps prevent the statistical phenomenon of overfitting [14], which people try to avoid [15], [16].

In this summary paper, we review a model of decision making under uncertainty that incorporates population categorization (and consequently limited threshold precision), mathematized using quantization theory [17]. Further connections to the literature and details on categorical decision making by individuals are provided in [18], [19] and on categorical decision making by groups/committees are provided in [20], [21]. The discussion on crowdsourcing that appears at the end of this paper does not appear elsewhere.

First we consider a single decision maker, aiming to understand the bias based on race, the so-called racial profiling that has been observed in several decision-making scenarios of the type described, including foul calls by National Basketball Association (NBA) referees, arrests for minor offences by police, searches of stopped vehicles by police, and jury decisions [6], [22]–[24]. Our model provides an explanation that does not invoke invidious motive by the decision maker. To reach this information-based explanation for discrimination, rather than a taste-based, statistical, or implicit one [25]–[27], we require a few ingredients in addition to limited precision in the LRT threshold. These additional ingredients arise from social structure [28] and social cognition [29]. The influence of social structure on categorical decision making yields a characterization of social welfare loss using a new concept called the *price of segregation*.

Next we consider a committee of decision makers that combine local decisions through majority vote [30], a form of distributed detection [31]. Our aim is to determine properties of optimal groups of categorical decision makers. When individual members of groups have different costs and benefits for global actions [32], so-called *preference heterogeneity* [33], both local decision making and local quantizer design become strategic form games [34]. By determining Nash equilibrium strategies, we find that there is deviation from Pareto optimality, and so optimal groups should have preference homogeneity. For the case of identical costs and benefits, the team-theoretic setting [35], [36], we further show that limiting to identical quantizers among agents leads to loss in performance. Diversity is beneficial. Thus decision making teams with diversity of training but aligned objectives perform better than other decision making groups.

Finally, we close with some comments on how coarse thinking may affect crowdsourcing competitions, aiming to explain the empirical fact that competitors typically expend cumulative resources in excess of the cumulative prize purse [37]–[39].

II. INDIVIDUAL DECISION MAKING

In this section, we develop a model of LRT decision making in which the prior probabilities that go into the threshold are optimally quantized. The model is further extended to have separate optimal quantizers for different subpopulations of objects.

A. Bayes Risk and Likelihood Ratio Test Detection Rule

Consider the signal detection scenario in which a decision maker uses a noisy observation Y to determine whether an object (e.g. player in the officiating context) is in state h_0 or state h_1 . State h_0 corresponds to a null hypothesis such as no foul committed, whereas h_1 corresponds to an alternative hypothesis such as foul committed. Noisy observations on whether or not a foul was committed are modeled by the likelihood functions $f_{Y|H}(y|h_0)$ and $f_{Y|H}(y|h_1)$. The object has prior probability p_0 of being in state h_0 and $p_1 = 1 - p_0$ of being in state h_1 , that is $p_0 = \Pr[H = h_0]$ and $p_1 = \Pr[H = h_1]$. In a population of objects, each object may have a different prior probability. That is, different players may have different prior propensities of committing a foul. The population is modeled by a probability density function $f_{P_0}(p_0)$ supported on the unit interval; this is a probability distribution on probabilities.

The detection rule $\hat{h}(y)$ of the decision maker is the following LRT:

$$\frac{f_{Y|H}(y|h_1)}{f_{Y|H}(y|h_0)} \underset{\hat{h}(y)=h_0}{\overset{\hat{h}(y)=h_1}{\gtrless}} \frac{c_{10}a}{c_{01}(1-a)}, \quad (1)$$

where c_{ij} is the non-negative cost of deciding h_j when the true state is h_i ; we assume correct decisions incur no cost. The parameter a weights the decision rule so as to allow the

incorporation of prior beliefs. There are two types of errors, with the following probabilities:

$$\begin{aligned} p_E^I &= \Pr[\hat{h}(Y) = h_1 | H = h_0], \\ p_E^{II} &= \Pr[\hat{h}(Y) = h_0 | H = h_1]. \end{aligned}$$

The Bayes risk, the performance of the decision rule, may be expressed in terms of those error probabilities as:

$$J(p_0, a) = c_{10}p_0p_E^I(a) + c_{01}(1 - p_0)p_E^{II}(a). \quad (2)$$

If the parameter a is set so $a = p_0$, then the decision rule (1) is the Bayes optimal decision rule, i.e., it minimizes (2).

In (2), the dependence of the Bayes risk and error probabilities on p_0 and a has been explicitly noted. The error probabilities depend on a through $\hat{h}(\cdot)$, given in (1).

B. Quantized Prior

The choice $a = p_0$ in the LRT (1) is the fully rational one. An essential piece of our model, however, is that the decision maker is bounded to be a coarse thinker and must use the same prior belief parameter a for different objects. Thus the decision maker only has access to the object's category label when making decisions about that object. Categorization of objects (in the prior probability space) is modeled as a quantizer for the population distribution $f_{P_0}(p_0)$. A K -point quantizer of $f_{P_0}(p_0)$ partitions the interval $[0, 1]$ into K regions $\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_K$. For each of these quantization regions \mathcal{R}_k , there is a representation point a_k to which elements are mapped. This value a_k may be thought of as the prior probability for a prototype member of the k th category. For regular quantizers, the regions are subintervals $\mathcal{R}_1 = [0, b_1]$, $\mathcal{R}_2 = (b_1, b_2]$, \dots , $\mathcal{R}_K = (b_{K-1}, 1]$ and the representation points a_k are in \mathcal{R}_k . A quantizer is a nonlinear function $q_K(\cdot)$ such that $q_K(p_0) = a_k$ for $p_0 \in \mathcal{R}_k$.

When constrained, the decision maker uses the prior belief parameter $a = q_K(p_0)$ in the likelihood ratio test (1). There are many different possible quantization functions or categorizations of objects; following the notion of costly rationality [40], $q_K(\cdot)$ should be optimal in terms of decision-making performance. We require that $q_K(\cdot)$ (for fixed K) minimize $D = E[J(P_0, q_K(P_0)) - J(P_0, P_0)]$, where the expectation is with respect to $f_{P_0}(p_0)$. As such, the quantization fidelity criterion is the difference between the quantized Bayes risk and the optimal unquantized Bayes risk. This criterion is an expectation of a novel quantization distortion function termed *Bayes risk error*.

Definition 1. Let Bayes risk error $d(p_0, a)$ be the difference between the Bayes risk functions $J(p_0, a)$ and $J(p_0, p_0)$:

$$d(p_0, a) = J(p_0, a) - J(p_0, p_0). \quad (3)$$

C. Optimal Categorization

In general the design of an optimal quantizer does not have a closed-form solution. Nevertheless, nearest neighbor, centroid, and zero probability of boundary conditions can be developed for mean Bayes risk error (MBRE) minimizing quantizer design; these lead to a Lloyd-Max algorithm to

design a locally optimal quantizer for $f_{P_0}(p_0)$ [18]. There are also quantizer design algorithms for discrete-valued $f_{P_0}(p_0)$ based on dynamic programming that find the globally optimal solution [41, Section 4.2.3].

D. Separate Quantizers for Different Subpopulations

Now consider the situation where the decision maker must deal with subpopulations that are distinguished according to a socially observable part of identity like race [42]. For ease of connection to empirical studies, we restrict to two groups and use ‘black’ and ‘white’ to denote them. The rational coarse-thinking decision maker should ignore the dimension of race altogether and simply partition along the p_0 dimension [43, Theorem 8.4.2], but social cognition constraints [44]–[46] prevent the decision maker from doing so.

Automaticity of racial categorization results in two quantizers designed separately for the two populations. The total quota on representation points, K_t , is split into some number of points for whites and some number for blacks, denoted $K_t = K_w + K_b$. The separate quantizers may then be denoted $q_{K_w}(\cdot)$ and $q_{K_b}(\cdot)$.

We can extend the definition of MBRE to two subpopulations as:

$$D^{(2)} = \frac{m_w}{m_w + m_b} E[J(P_0, q_{K_w}(P_0))] + \frac{m_b}{m_w + m_b} E[J(P_0, q_{K_b}(P_0))] - E[J(P_0)], \quad (4)$$

where m_w and m_b are the number of whites and blacks relevant to the decision maker in both social and economic life (which are cognitively intertwined [47]). The goal is to minimize this extended MBRE by finding the optimal quantizers $q_{K_w}(\cdot)$ and $q_{K_b}(\cdot)$ and the optimal allocation of representation points K_w and K_b .

The model we propose assumes that the two populations are identical. Thus $q_{K_w}(\cdot)$ and $q_{K_b}(\cdot)$ should be designed as discussed in Section II-C. The problem reduces to the straightforward minimization of MBRE over all $K_t - 1$ possible allocations of K_w and K_b . If m_w is larger than m_b , it is better to allocate more representation points to whites whereas if m_b is larger than m_w , it is better to allocate more representation points to blacks.

In this section, we have proposed a Bayesian likelihood ratio test that incorporates prior probability quantization to model human decision making. The model has been further extended for decision making on distinct racial populations through separate quantization functions to incorporate social cognition factors. The next section quantitatively and qualitatively shows one implication of this model; further implications are given in [19].

III. PRICE OF SEGREGATION

Due to social segregation and the racial isolation it induces, there is greater intra-population interaction than inter-population interaction. Whites interact more with whites whereas blacks interact more with blacks.

In the mathematical model, one would expect the $m_w/(m_w + m_b)$ of a white decision maker to be greater

than the $m_w/(m_w + m_b)$ of a black decision maker. Due to optimal representation point allocations and the fact that Bayes risk performance does not get worse when the number of quantization levels K is increased [18], a white decision maker would perform worse than a black decision maker when dealing with blacks and a black decision maker would perform worse than a white decision maker when dealing with whites, judging quality by expected Bayes risk. Under appropriate values of decision maker preferences, c_{01} and c_{10} , the decision making model generates the in-race bias or out-of-race bias that has been observed empirically [6], [22]–[24].

Decision makers of different races exhibit different biases because they have different K_w and K_b allocations due to different $m_w/(m_w + m_b)$ ratios. This ratio is not the actual fraction of whites whose actions are assessed by the decision maker, but is determined in part by the decision maker’s segregated social life. If decision makers of all races have a bias that matches the true white fraction, then the phenomenon of racial bias would actually achieve optimal social welfare. Different decision-making biases by different decision makers, however, cannot simultaneously be societally optimal.

Our model fixes limitations of human information processing, automaticity of racial classification, and intertwining of social and economic life. Social segregation causes mismatch between social and economic lives and is therefore the root cause of non-optimal racial bias. In analogy with notions of welfare loss in economic theory, a *price of segregation* is defined here as a way to measure the deleterious effect of segregation.

Let π_{true} be the fraction of whites in the economic decision-making setting. A particular decision maker that leads a segregated life, on the other hand, will have a white ratio $\pi_{\text{seg}} = m_w/(m_w + m_b)$. The MBRE, from the perspective of society, under the true white fraction is

$$D^{(2)}(\pi_{\text{true}}) = \pi_{\text{true}} E[J(P_0, q_{K_w}(\pi_{\text{true}})(P_0))] + (1 - \pi_{\text{true}}) E[J(P_0, q_{K_b}(\pi_{\text{true}})(P_0))] - E[J(P_0, P_0)]$$

whereas the MBRE, from the perspective of society, under the segregated white fraction is

$$D^{(2)}(\pi_{\text{seg}}) = \pi_{\text{true}} E[J(P_0, q_{K_w}(\pi_{\text{seg}})(P_0))] + (1 - \pi_{\text{true}}) E[J(P_0, q_{K_b}(\pi_{\text{seg}})(P_0))] - E[J(P_0, P_0)].$$

The difference between these two is the price of segregation:

$$\Pi = D^{(2)}(\pi_{\text{true}}) - D^{(2)}(\pi_{\text{seg}}). \quad (5)$$

The price of segregation Π depends strongly on the discontinuous, integer-valued $K_w(\cdot)$ function, and is also discontinuous. The price of segregation is a non-decreasing function of the level of segregation mismatch $|\pi_{\text{true}} - \pi_{\text{seg}}|$. An example of the price of segregation for a particular society and several different values of π_{true} is shown in Figure 1. Notice that if the level of mismatch is small, there may be no price of segregation, similar to how entire regions in the probability simplex have the same optimal Huffman code [48].

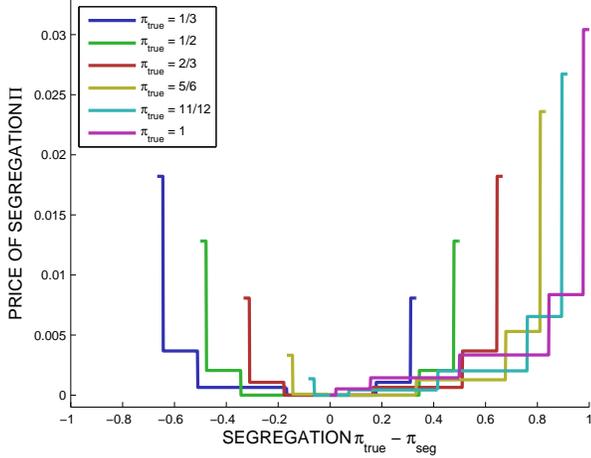


Fig. 1. The price of segregation Π as a function of the level of segregation mismatch $\pi_{\text{true}} - \pi_{\text{seg}}$ for several values of π_{true} . The distribution of the prior probability is beta(5,2), the measurements are observed through additive white Gaussian noise with unit signal to noise ratio, the Bayes costs are $c_{10} = c_{01} = 1$, and $K_t = 7$.

The model predicts that greater homogeneity of social interaction among people would mitigate the price of segregation by driving the π_{seg} for all decision makers closer to π_{true} . This draws a connection to intergroup contact theory [49], [50]. One branch of contact theory suggests that contact reduces prejudice since it allows individuals the chance to see previously unnoticed similarities and counter-stereotypic characteristics and behaviors in one another [50], [51], a conclusion similar to model predictions.

Unexpectedly, social interaction is not linear in the overall ratio of subgroup populations [52], perhaps due to structural reasons [28], so a policy directive to mitigate the price of segregation would be difficult to formulate.

IV. GROUP DECISION MAKING

Section II developed a mathematical model of an individual memory-constrained decision maker. Here we consider a group of such decision makers, which make local decisions and combine them using majority vote. The basic problem formulation is depicted in Figure 2 for the case of three decision making agents (assumed throughout this section).

A. Bayes Risk and Quantized Priors

For each i , Agent i (marked D_i) observes Y_i satisfying likelihood function $f_{Y_i|H}$ and sends a local decision $\hat{h}_i \in \{h_0, h_1\}$ to a fusion center. The observations are assumed to be conditionally independent given H . A fusion center determines $\hat{h} \in \{h_0, h_1\}$ by majority rule.

Agent i has local error probabilities

$$p_{E,i}^I = \Pr(\hat{h}_i(Y_i) = h_1 | H = h_0), \text{ and}$$

$$p_{E,i}^{II} = \Pr(\hat{h}_i(Y_i) = h_0 | H = h_1).$$

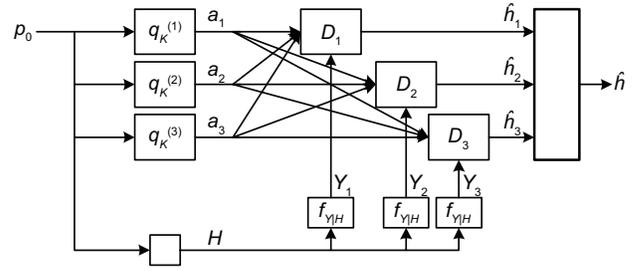


Fig. 2. A schematic diagram depicting the group decision making problem (for three agents). The environment generates a Bernoulli signal H using prior probability p_0 , which is quantized by three separate quantizers; the results a_i are used by local agents D_i . Each agent also has access to H corrupted by noise and all agents' Bayes costs $c_{10}^{(i)}$ and $c_{01}^{(i)}$. The fusion center determines \hat{h} from the local decisions \hat{h}_i .

Since global errors occur exactly when the majority of agents make local errors, global error probabilities can be expressed in terms of the local error probabilities:

$$p_E^I = \Pr(\hat{H} = h_1 | H = h_0) \quad (6)$$

$$= p_{E,1}^I p_{E,2}^I + p_{E,2}^I p_{E,3}^I + p_{E,3}^I p_{E,1}^I - 2p_{E,1}^I p_{E,2}^I p_{E,3}^I,$$

$$p_E^{II} = \Pr(\hat{H} = h_0 | H = h_1) \quad (7)$$

$$= p_{E,1}^{II} p_{E,2}^{II} + p_{E,2}^{II} p_{E,3}^{II} + p_{E,3}^{II} p_{E,1}^{II} - 2p_{E,1}^{II} p_{E,2}^{II} p_{E,3}^{II}.$$

All error probabilities depend on p_0 and on the $\hat{h}_i(\cdot)$, which are all assumed to take the form of threshold tests with real-valued thresholds.

When p_0 is known, the goal of Agent i is to minimize the expected value of the i th Bayes risk

$$\tilde{J}_i = c_{10}^{(i)} p_0 p_E^I + c_{01}^{(i)} (1 - p_0) p_E^{II}, \quad (8)$$

where $c_{10}^{(i)}$ and $c_{01}^{(i)}$ are the Bayes costs for Agent i . Through the definitions of p_E^I and p_E^{II} , it is clear that \tilde{J}_i is the conditional mean of the Bayes cost given $P_0 = p_0$.

As depicted in Figure 2, Agent i quantizes p_0 to $a_i = q_K^{(i)}(p_0)$ due to information-processing limitations. Thus, Agent i makes decisions to minimize i th perceived Bayes risk

$$\bar{J}_i = c_{10}^{(i)} a_i p_E^I + c_{01}^{(i)} (1 - a_i) p_E^{II}. \quad (9)$$

The decision threshold, and consequently $p_{E,i}^I$ and $p_{E,i}^{II}$, of each agent is determined based on $\{\bar{J}_i, i = 1, 2, 3\}$. However, the true Bayes risk is \tilde{J}_i , where p_E^I and p_E^{II} in (8) are affected by the perceived Bayes risks. We define MBR for the quantizer $q_K^{(i)}$ of $f_{P_0}(p_0)$:

$$E[\bar{J}_i] = \int_0^1 [c_{10}^{(i)} p_0 p_E^I(a_i) + c_{01}^{(i)} (1 - p_0) p_E^{II}(a_i)] f_{P_0}(p_0) dp_0. \quad (10)$$

The MBR of an agent differs from that of other agents, but depends on the quantizers of all agents. Thus designing quantizers is a strategic form game. Later we will find Nash equilibrium quantizer designs, but first we discuss a game-theoretic formulation of decision-making itself.

In the sequel, assume that observation noise is additive white Gaussian.

TABLE I
GAME I: DECISION RULES

- The set of players $\mathcal{I} = \{1, 2, 3\}$.
- The set of available strategies $S_i = \mathbb{R}$, which is a set of possible decision thresholds s_i for Agent i .
- The payoff function $u_i = -\bar{J}_i, \forall i \in \mathcal{I}$, which is the negative perceived Bayes risk.

TABLE II
GAME II: QUANTIZER FOR PRIORS

- The set of players $\mathcal{I} = \{1, 2, 3\}$.
- The strategy of Agent $i \in \mathcal{I}$ is quantizer representation points and cell boundaries $s_i = (a_{1,i}, \dots, a_{K,i}, b_{1,i}, \dots, b_{K-1,i})$.
- The payoff function of Agent $i \in \mathcal{I}$ is $v_i = -\mathbb{E}[\bar{J}_i]$, the negative mean Bayes risk.

B. Equilibrium Detection Rules

Agents have heterogeneous preferences, thereby restricting collaboration. However game theory provides useful methods to analyze agents' decision-making strategies under competition [34]. It is straightforward to describe the decision-making problem in strategic form, Game I, with players \mathcal{I} , strategies $(S_i)_{i \in \mathcal{I}}$ for choosing decision thresholds, and payoffs $(u_i)_{i \in \mathcal{I}}$ [34]. Several properties can be proven [21]; proofs are omitted here.

Theorem 1. *Dominant strategies do not exist in Game I.*

In addition to the lack of existence of dominant strategies, the only dominated strategies for any agent are $s_i = \infty$ and $s_i = -\infty$. Therefore, Game I is not solvable by iterative dominance. Each agent's decision rule depends on other agents' rules. Thus, we consider Nash equilibrium strategies.

Theorem 2. *A pure Nash equilibrium always exists in Game I.*

As detailed in [21], a Nash equilibrium decision-making strategy for each agent can be computed directly. Due to conflict, however, a computed Nash equilibrium is not Pareto optimal. As shown in Figure 3, the operating point of the Nash equilibrium is located in the interior of the operating region. The agents can improve their performance by changing their decision rules (true beyond the example). Since agents have different Bayes costs, however, they cannot agree on how to improve decision rules, incurring a *penalty of discord*.

Just as we had assumed the individual decision maker uses LRT decision making, we now assume that the group of agents adopt threshold-based Nash equilibrium decision rules.

C. Equilibrium Categorization

Quantizer design is also a strategic form game, Game II. There are $2K - 1$ degrees of freedom in a strategy when agents use K -point quantizers: K for representation points $\{a_{k,i}\}_{k=1}^K$ and $K - 1$ for cell boundaries $\{b_{k,i}\}_{k=1}^{K-1}$.

Theorem 3. *Game II does not always have a dominant strategy.*

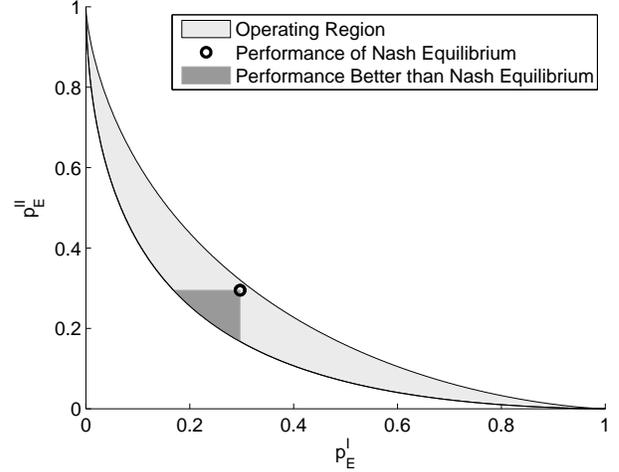


Fig. 3. The penalty of discord for three agents whose Bayes costs are $c_{10}^{(1)} = 1, c_{01}^{(1)} = 4, c_{10}^{(2)} = 4, c_{01}^{(2)} = 1, (c_{10}^{(3)} = 4, \text{ and } c_{01}^{(3)} = 4,$ for $a_i = 0.5$, and Gaussian observation noise with unit signal to noise ratio.

Since there need not exist a dominant strategy, we consider Nash equilibria. Finding equilibrium strategies has high computational complexity for $K \geq 2$ because each variable depends on some others.

Moreover, the dependency varies with quantizer partition structure. By partition structure, we mean how cells in the quantizer of Agent i overlap with the cells in the quantizer of Agent j . To find the best set of quantizers, we need to find the equilibrium set of quantizers for each possible partition structure and choose the best one. The number of partition structures is given by De Bruijn's $s(3, K - 1)$:

$$\frac{(3(K - 1))!}{((K - 1)!)^3} \sim \frac{1}{2} \sqrt{3} \frac{27^{K-1}}{\pi(K - 1)}.$$

Although equilibrium quantizer design is computationally difficult, it can be carried out.

If the quantizers have the same partitions, then each partition is independent of other partitions and the only remaining dependency is among agents. Suppose all quantizers have the same partition, with fixed boundaries

$$0 = b_0 < b_1 < b_2 < \dots < b_{K-1} < b_K = 1.$$

The problem of designing quantizers reduces to the problem of choosing representation points.

Alternatively for common partitions, equilibrium decision rules of agents within each partition can be determined, thereby directly controlling decision making. Interestingly, the Nash equilibrium of the latter is exactly the single-agent centroid condition [21]. That is, each agent locally optimizes quantization after coordinated partition regions are imposed.

In this section, we have described a completely novel quantizer design game. We have also argued that group decision making suffers from a penalty of discord under heterogeneous preferences.

V. HOMOGENOUS PREFERENCES

In the previous section, we had seen that there is inherently a penalty of discord when agents have differing preferences. We had also seen that some level of coordination can strongly simplify equilibrium strategies. In this section, we consider what happens when the group of memory-limited decision makers has homogenous preferences, i.e. they are a team in the sense of Marschak and Radner [35], [36] with identical Bayes costs c_{01} and c_{10} . This setting (with unbounded decision makers) has been studied as distributed detection in engineering [31] and has also been studied in economics and political science [30] to understand juries, committees, and elections.

Although the problem is identical as before, considering the team-theoretic limit of the game-theoretic formulation leads to some further insights.

A. Detection Rules

For notational simplicity, consider the decision rule for Agent 1 in a setting with three agents. The optimal decision for Agent 1 minimizes Bayes risk (9) with $i = 1$. Suppose that Agents 2 and 3 have some fixed decision rules. By rewriting (6) and (7), we have

$$\begin{aligned} p_E^I &= (p_{E,2}^I + p_{E,3}^I - 2p_{E,2}^I p_{E,3}^I) p_{E,1}^I + p_{E,2}^I p_{E,3}^I \\ &\triangleq A_{11} p_{E,1}^I + A_{12}, \\ p_E^{II} &= (p_{E,2}^{II} + p_{E,3}^{II} - 2p_{E,2}^{II} p_{E,3}^{II}) p_{E,1}^{II} + p_{E,2}^{II} p_{E,3}^{II} \\ &\triangleq B_{11} p_{E,1}^{II} + B_{12}, \end{aligned}$$

where A_{11} , A_{12} , B_{11} , and B_{12} are nonnegative quantities that do not depend on the decision rule of Agent 1. The optimal decision rule $\hat{h}_1(y_1)$ for Agent 1 can be expressed as an LRT

$$\frac{f_{Y_1|H}(y_1|h_1)}{f_{Y_1|H}(y_1|h_0)} \underset{\hat{h}_1(y_1)=h_0}{\overset{\hat{h}_1(y_1)=h_1}{>}} \frac{c_{10}p_0 A_{11}}{c_{01}(1-p_0)B_{11}} \triangleq \eta_1,$$

which is similar to the optimal decision rule for a single agent.

Under the Gaussian observation model, Agent 1 has optimal decision rule

$$y_1 \underset{\hat{h}_1(y_1)=h_0}{\overset{\hat{h}_1(y_1)=h_1}{>}} \frac{h_1 - h_0}{2} + \frac{\sigma^2}{h_1 - h_0} \ln \eta_1 \triangleq \lambda_1. \quad (11)$$

The optimal decision rules of Agents 2 and 3 are analogous, with thresholds λ_2 and λ_3 . Note that $(p_{E,1}^I, p_{E,1}^{II})$, $(p_{E,2}^I, p_{E,2}^{II})$, and $(p_{E,3}^I, p_{E,3}^{II})$, depend directly on λ_1 , λ_2 , and λ_3 , respectively. The optimal value of λ_1 in (11) is a function of λ_2 and λ_3 through η_1 (hence A_{11} and B_{11}), and similarly for λ_2 and λ_3 . Thus, the λ_i cannot be optimized independently.

B. Bayes Risk and Quantized Priors

As depicted in Figure 2, the decision rule of Agent i is based on a quantized version of the prior probability. Thus, Agent i makes decisions to minimize perceived Bayes risk \bar{J}_i , (9) with common Bayes costs c_{01} and c_{10} .

The agent's decision rule, and consequently $p_{E,i}^I$ and $p_{E,i}^{II}$, are determined based on \bar{J}_i . The (true) MBR with decision

rules impacted by quantization of the prior probability is $E[\bar{J}]$, (10) with common Bayes costs c_{01} and c_{10} , which is the appropriate fidelity criterion for quantizer design.

C. Optimal Categorization for Identical Agents

Suppose that all agents use the same quantizer. This causes all perceived Bayes risks to be equal to

$$\bar{J} = c_{10} a p_E^I + c_{01} (1-a) p_E^{II}.$$

We further assume that all decision rules are identical, which has long been conjectured to be optimal in the distributed detection literature [53]. In the Gaussian observation model, identical decision rules correspond to all λ_i thresholds taking one common value λ . Identical decision rules imply that all local probabilities of error are equal.

Now consider any odd number of agents using the decision rule

$$y_i \underset{\hat{h}_i(y_i)=h_0}{\overset{\hat{h}_i(y_i)=h_1}{>}} \lambda \quad (12)$$

for some λ . With this rule, $\hat{h}_i(y_i) = h_1$ implies $\hat{h}_j(y_j) = h_1$ for any j such that $y_j > y_i$. Likewise, $\hat{h}_i(y_i) = h_0$ implies $\hat{h}_j(y_j) = h_0$ for any j such that $y_j < y_i$. Thus, comparing the median y_i to λ determines the majority-rule decision. Moreover, the performance of the collaborating group is the same as that of a single agent with observation likelihood function $f_{Y|H}$ determined by the median of the group's observations.

Theorem 4. *Suppose $2n + 1$ agents make decisions using the rule (12) for some λ , where the observation of Agent i is $Y_i = H + W_i$. The collaborative performance by majority rule is equal to the performance of a single agent using rule (12) with observation $Y = H + V$, where $V = \text{median}(\{W_i\}_{i=1}^{2n+1})$.*

The distribution of V can be derived from the distribution of $\{W_i\}$ using the theory of order statistics [54].

Fact 1. *Suppose $\{W_i\}_{i=1}^{2n+1}$ are i.i.d. $\mathcal{N}(0, 1)$ and $V = \text{median}(\{W_i\})$. Then*

$$\sqrt{n}V \longrightarrow \mathcal{N}(0, \frac{\pi}{2}) \text{ as } n \rightarrow \infty.$$

On the other hand, letting M be the mean of the $\{W_i\}$,

$$\sqrt{n}M \longrightarrow \mathcal{N}(0, 1) \text{ as } n \rightarrow \infty.$$

The ratio between the two asymptotic variances, $2/\pi$, quantifies the loss in using majority voting rather than full measurement fusion for large committees; the loss is less for smaller committees [20].

The equivalence theorem, Theorem 4, implies that optimization of the quantizer in the multi-agent model can be converted to the optimization problem for a single agent discussed in Section II-C.

D. Optimal Categorization for Diverse Agents

Now remove the restriction that agents use the same quantizer. Differently quantized prior probabilities make the agents' perceived Bayes risks differ even though the agents have the same Bayes costs. In this section, we limit attention to three agents and an equal number of levels K for each agent's quantizer.

Since we assume that the agents do not have access to the true prior p_0 , the decision rule is optimized based on the quantized prior probabilities a_i . One collaborative way is to minimize the perceived Bayes risk averaged over the three agents, which we call the *perceived common risk*:

$$\begin{aligned} \bar{J}_C &= \frac{1}{3}(\bar{J}_1 + \bar{J}_2 + \bar{J}_3) \\ &= \frac{1}{3}c_{10}(a_1 + a_2 + a_3)p_E^I + \frac{1}{3}c_{01}(3 - a_1 - a_2 - a_3)p_E^{II}. \end{aligned}$$

As it turns out, it is possible to show a performance equivalence between three agents using quantizers $(q_K^{(1)}, q_K^{(2)}, q_K^{(3)})$ and three agents using a shared quantizer q_S . The equivalence enables the optimization of $(q_K^{(1)}, q_K^{(2)}, q_K^{(3)})$ and shows that when these quantizers are different, the performance achieved is commensurate with having finer quantizers [20].

Theorem 5. Consider three quantizers $(q_K^{(1)}, q_K^{(2)}, q_K^{(3)})$ and another quantizer $q_{3K-2}^{(S)}$. Agents that use $(q_K^{(1)}, q_K^{(2)}, q_K^{(3)})$ and those that use $(q_{3K-2}^{(S)}, q_{3K-2}^{(S)}, q_{3K-2}^{(S)})$ achieve the same performance for any p_0 if $\bigcup_{i=1}^3 B_i = B_S$ and $q_{3K-2}^{(S)}(p_0) = \frac{1}{3} \sum_{i=1}^3 q_K^{(i)}(p_0)$, where B_1, B_2, B_3 , and B_S are the sets of cell boundaries of $q_K^{(1)}, q_K^{(2)}, q_K^{(3)}$, and $q_{3K-2}^{(S)}$, respectively.

Recall that the optimization of $q^{(S)}$ to be shared by three identical agents can be done as in Section V-C.

Given Theorem 5, agents with diverse K -level quantizers $(q^{(1)}, q^{(2)}, q^{(3)})$ cannot perform better than the identical agents using $q^{(S)}$. A mapping described in [20] from the optimized $q^{(S)}$ to $(q^{(1)}, q^{(2)}, q^{(3)})$, leads to the design of optimal diverse quantizers.

The fact that diverse quantizers with K levels perform just as well as identical quantizers with $3K - 2$ levels demonstrates that there is a significant *value of diversity*. An example is shown in Figure 4.

In summary, decision making by committees of memory-constrained agents achieves the best performance when the penalty of discord is avoided, but the value of diversity is earned.

VI. CROWDSOURCING CONTESTS

There has been growing interest in harnessing the wisdom of crowds, beyond simple information aggregation through voting, as described in the previous sections. Indeed crowdsourcing contests with prizes are an emerging business trend, whether to design pieces of software code, t-shirts, or car components [55]. There has also been a resurgence in philanthropic/governmental prizes for meeting technological goals, such as the deployment of suborbital spacecraft [38].

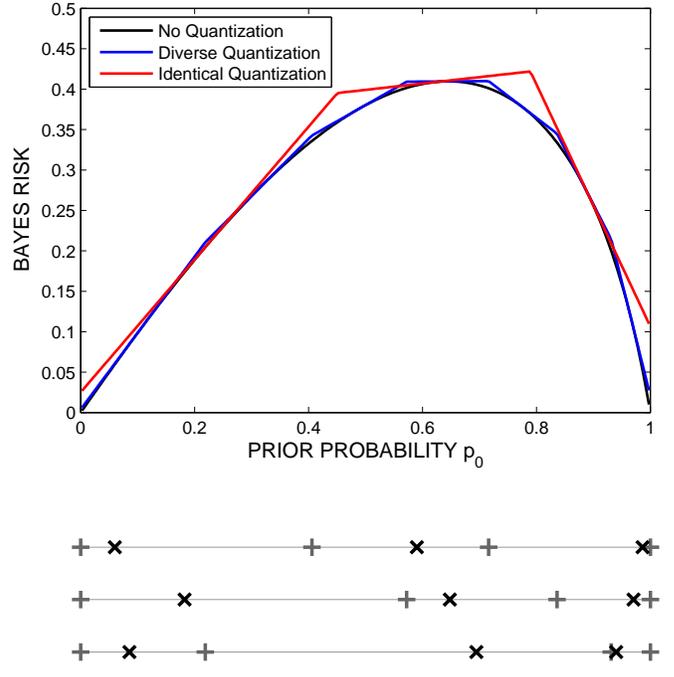


Fig. 4. The value of diversity is the difference between mismatched Bayes risk for optimal performance under identical K -level quantizers and mismatched Bayes risk for optimal performance under diverse K -level quantizers, whose reproduction and boundary points are shown below the plot. The setting here is Bayes costs $c_{01} = 4$, $c_{10} = 1$, P_0 distributed uniformly on $[0, 1]$, $K = 3$, and Gaussian observation noise with unit signal to noise ratio.

These winner-take-all competitions have the property that many participants exert effort or expend resources to win the prize, but the losers' efforts or expenditures go uncompensated. A simple and natural way to model such competitions is as auctions, where all players pay the amount of their bids, but only the participant that bids the highest wins the prize [56]. These are called *all-pay auctions* [57].

A. Bidding Rules

Suppose there are n bidders D_i in an all-pay auction, $i = 1, 2, \dots, n$, each submitting a sealed bid η_i for a prize valued by bidder i at θ_i . All players forfeit their bids, but the highest bidder wins the prize.

For each player a bid strategy involves choosing a non-negative bid value η_i . Choosing a bid value $\eta_i = 0$ means the player has decided not to compete in the auction.

When there is complete information, the payoff to bidder i is

$$J_i(\eta_1, \dots, \eta_n) = \begin{cases} -\eta_i, & \text{if there is } j \text{ such that } \eta_j > \eta_i, \\ \frac{\theta_i}{m} - \eta_i, & \text{if } i \text{ in } m - 1\text{-way tie for win,} \\ \theta_i - \eta_i, & \text{if } \eta_i > \eta_j \text{ for all } j \neq i. \end{cases} \quad (13)$$

One can interpret the θ_i as though each competitor has some prior ability to convert an entry into a prize. Players with higher θ can be thought of as stronger players. To see this, suppose the utility to player i of winning a prize of Z by putting forth effort η_i is $u_i^* = U_i(Z) - \alpha_i \eta_i$, where α_i is the

TABLE III
GAME III: BIDDING RULES

-
- The set of players $\mathcal{I} = \{1, 2, \dots, n\}$.
 - The strategy of Competitor $i \in \mathcal{I}$ is a bid $\eta_i \in \mathbb{R}^+$.
 - The payoff function of Competitor $i \in \mathcal{I}$ is J_i , (13).
-

marginal cost of effort to player i . Since behavior is invariant to affine transformations, the utility function can be rewritten as $u_i = u_i^*/\alpha_i = \theta_i - \eta_i$, where $\theta_i = U_i(Z)/\alpha_i$. So differences in the θ_i may be due to differences in the abilities of players.

The all-pay auction game is given as Game III. Nash equilibrium strategies for the all-pay auction with complete information depend strongly on the strict ordering of player prior strengths θ [57]. Without going into full detail, let us briefly describe the Nash equilibria (assuming $n > 2$):

- C1 With identical strengths ($\theta_1 = \theta_2 = \dots = \theta_n$), there exists a unique symmetric equilibrium and a continuum of asymmetric equilibria. In any equilibrium, the expected sum of bids is θ_1 , and the expected payoff to each player is zero. The same result carries over to the case when there are several players that are tied for strongest, and other weaker players; the weaker players do not actively bid and the strongest players behave as if the weaker players do not exist.
- C2 When there is a strongest player, a second strongest player, and other players, i.e. $\theta_1 > \theta_2 > \theta_3 \geq \dots \geq \theta_n$, then there is a unique equilibrium, where only the two strongest players actively bid. The strongest player earns payoff $J_1 = \theta_1 - \theta_2$, whereas other players earn payoffs of zero.
- C3 With strengths $\theta_1 > \theta_2 = \dots = \theta_m \geq \theta_{m+1} \dots \geq \theta_n$, there exists a continuum of Nash equilibria. In any equilibrium, players $m+1$ through n do not actively bid. In any equilibrium, the strongest player earns expected payoff $J_1 = \theta_1 - \theta_2$ and all other players earn payoff zero. Different equilibria yield different expected sums of bids, all less than $2\theta_2$.

Note in particular that in the all-pay auction with identical strengths, the equilibrium strategies result in the sum of the expected bids equalling the value of the prize, so-called full dissipation of the rent.¹ Indeed, for any all-pay auction with complete information in equilibrium, there should be no overdissipation.

Just as we had assumed the individual decision maker uses LRT decision making and group of agents adopt threshold-based Nash equilibrium decision rules, we now assume that competitors in all-pay auctions use Nash equilibrium strategies from the all-pay auction with complete information [57].

¹People are said to be seeking *rents* when trying to obtain benefits for themselves through the political arena. Lobbying for privileges is costly and these expenditures, therefore, dissipate some of the gains to the beneficiaries and cause inefficiency.

B. Quantized Competitive Knowledge

We relax the assumption of complete information in all-pay auctions as follows. Suppose players are drawn from a population of players, with strength distribution $f_{\Theta}(\theta)$. Now due to information processing limitations, player i only has access to a quantized version of his/her own strength $q_K^{(i)}(\theta_i)$ and quantized versions of the other players' strengths $q_K^{(i)}(\theta_j), j \neq i$. Not knowing oneself or one's competitors completely has been noted in a study of X PRIZE competitors: "many teams have imagined unrealistic chances of their winning while their projects are not competitive enough" [58, p. 32].

Each player makes decisions on whether to actively bid and how much to bid through the Nash equilibrium strategy described above, yielding the random variable realization $H_i = \eta_i$. The utility earned is $J_i(q_K^{(i)}(\theta_1), \dots, q_K^{(i)}(\theta_n))$.

It remains to optimize quantizer design to optimize player categorization. It is clear, however, that categorization of players will condense a possible total ordering of competitor strength into a perceived partial order of strengths. Perhaps this would shift the bidding behavior of a player from a Nash equilibria strategy C2 or C3 into strategy C1, where everyone competes aggressively.

C. Overbidding

Many experimental studies of all-pay auctions have found that participants tend to bid more aggressively than in Nash equilibrium [37]. This also holds in much larger real-world competitions. For example in the Orteig Prize won by Charles Lindbergh, competitors spent a cumulative \$400,000 for a prize of \$25,000; in the Ansari X PRIZE, competitors spent a cumulative \$100 million for a prize of \$10 million [38]. Moreover, the cost of rent seeking in India has been estimated on the order of 30–45 percent of GNP [39]. This qualitatively opposes the adverse selection effect due to information asymmetry that eliminates markets for lemons [59].

One might wonder whether there is an informational explanation for why there is so much overspending. Can bounded rationality and in particular, categorical thinking by competitors explain this phenomenon? It remains to show categorical decision-making can do so, but shifting bidding strategies due to changes in perceived competition is a possible mechanism to do so.

If the putative mechanism stands, it raises the question of how overspending can be quantitatively cast as a *cost of competitive ignorance*. Note that although political economists see overspending as social welfare loss [39], philanthropic prize designers see it as a primary benefit of competitions, terming it leverage [38].

VII. CONCLUSION

Categorical thinking is a new form of bounded rationality that is based on an understanding of human cognition. In this paper, we have summarized recent work on categorical decision making by individuals and by groups, explaining

some puzzling empirical phenomena and characterizing optimal decision-making teams [18]–[21]. We have also suggested a further economic setting—crowdsourcing contests—where categorical thinking may explain empirically observed phenomenon.

As part of the development, we have introduced several quantitative measures of social welfare loss/gain:

- price of segregation,
- penalty of discord,
- value of diversity, and
- cost of competitive ignorance.

Our application of quantization theory to modeling human behavior and in conjunction with game theory is completely novel. Given the rise of behavioral economics, perhaps the time is ripe for further information-theoretic models and analyses to inform an understanding of markets, firms, and civic life.

ACKNOWLEDGMENT

Helpful conversations with P. K. Varshney are appreciated.

REFERENCES

- [1] M. Spence, "Job market signaling," *Quart. J. Econ.*, vol. 87, no. 3, pp. 355–374, Aug. 1973.
- [2] S. Schwab, "Is statistical discrimination efficient?" *Am. Econ. Rev.*, vol. 76, no. 1, pp. 228–234, Mar. 1986.
- [3] P. Soelberg, "Unprogrammed decision making," Alfred P. Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, Tech. Rep. 239-67, 1967.
- [4] D. S. Chapman, K. L. Uggerslev, S. A. Carroll, K. A. Piasentin, and D. A. Jones, "Applicant attraction to organizations and job choice: A meta-analytic review of the correlates of recruiting outcomes," *J. Appl. Psychol.*, vol. 90, no. 5, pp. 928–944, Sep. 2005.
- [5] K. Ratakonda, R. Williams, J. Bisceglia, R. W. Taylor, and J. Graham, "Identifying trouble patterns in complex IT services engagements," *IBM J. Res. Develop.*, vol. 54, no. 2, p. 5, Mar.-Apr. 2010.
- [6] J. Price and J. Wolfers, "Racial discrimination among NBA referees," NBER, Working Paper 13206, Jun. 2007.
- [7] A. S. Willisky, G. W. Wornell, and J. H. Shapiro, *Stochastic Processes, Detection and Estimation 6.432 Course Notes*. Cambridge, MA: Dept. Elect. Eng. Comput. Sci., Mass. Inst. Tech., Fall 2003.
- [8] J. A. Swets, W. P. Tanner, Jr., and T. G. Birdsall, "Decision processes in perception," *Psychol. Rev.*, vol. 68, no. 5, pp. 301–340, Sep. 1961.
- [9] M. Glanzer, A. Hilford, and L. T. Maloney, "Likelihood ratio decisions in memory: Three implied regularities," *Psychon. Bull. Rev.*, vol. 16, no. 3, pp. 431–455, Jun. 2009.
- [10] G. L. Brase, L. Cosmides, and J. Tooby, "Individuation, counting, and statistical inference: The role of frequency and whole-object representations in judgment under uncertainty," *J. Exp. Psychol. Gen.*, vol. 127, no. 1, pp. 3–21, Mar. 1998.
- [11] G. A. Miller, "Human memory and the storage of information," *IRE Trans. Inf. Theory*, vol. IT-2, no. 3, pp. 129–137, Sep. 1956.
- [12] R. Fryer and M. O. Jackson, "A categorical model of cognition and biased decision-making," *B. E. J. Theor. Econ.*, vol. 8, no. 1, Jan. 2008.
- [13] S. Mullainathan, J. Schwartzstein, and A. Shleifer, "Coarse thinking and persuasion," *Quart. J. Econ.*, vol. 123, no. 2, pp. 577–619, May 2008.
- [14] V. N. Vapnik, *The Nature of Statistical Learning*. New York: Springer-Verlag, 2000.
- [15] X. Zhu, T. T. Rogers, and B. Gibson, "Human Rademacher complexity," in *Advances in Neural Information Processing Systems 22*, Y. Bengio, D. Schuurmans, J. Lafferty, C. Williams, and A. Culotta, Eds. Cambridge, MA: MIT Press, 2009, pp. 2322–2330.
- [16] R. Hertwig and P. M. Todd, "More is not always better: The benefits of cognitive limits," in *Thinking: Psychological Perspectives on Reasoning, Judgment and Decision Making*, D. Hardman and L. Macchi, Eds. Chichester, UK: John Wiley & Sons, 2005, pp. 213–231.
- [17] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston: Kluwer Academic Publishers, 1992.
- [18] K. R. Varshney and L. R. Varshney, "Quantization of prior probabilities for hypothesis testing," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4553–4562, Oct. 2008.
- [19] L. R. Varshney and K. R. Varshney, "Decision making with quantized priors leads to discrimination," 2010, working paper.
- [20] J. B. Rhim, L. R. Varshney, and V. K. Goyal, "Collaboration in distributed hypothesis testing with quantized prior probabilities," in *Proc. IEEE Data Compression Conf. (DCC 2011)*, Mar. 2011, to appear.
- [21] —, "Conflict in distributed hypothesis testing with quantized prior probabilities," in *Proc. IEEE Data Compression Conf. (DCC 2011)*, Mar. 2011, to appear.
- [22] J. J. Donohue, III and S. D. Levitt, "The impact of race on policing and arrests," *J. Law Econ.*, vol. 44, no. 2, pp. 367–394, Oct. 2001.
- [23] K. L. Antonovics and B. G. Knight, "A new look at racial profiling: Evidence from the Boston Police Department," *Rev. Econ. Stat.*, vol. 91, no. 1, pp. 163–177, Feb. 2009.
- [24] R. M. Bagby, J. D. Parker, N. A. Rector, and V. Kalembo, "Racial prejudice in the Canadian legal system: Juror decisions in a simulated rape trial," *L. Hum. Behav.*, vol. 18, no. 3, pp. 339–350, Jun. 1994.
- [25] G. S. Becker, *The Economics of Discrimination*. Chicago: University of Chicago Press, 1957.
- [26] K. Arrow, "The theory of discrimination," in *Discrimination in Labor Markets*, O. Ashenfelter and A. Rees, Eds. Princeton, NJ: Princeton University Press, 1973, pp. 3–33.
- [27] M. Bertrand, D. Chugh, and S. Mullainathan, "Implicit discrimination," *Am. Econ. Rev.*, vol. 95, no. 2, pp. 94–98, May 2005.
- [28] P. Blau, *Inequality and Heterogeneity: A Primitive Theory of Social Structure*. New York: Free Press, 1977.
- [29] M. Augoustinos, I. Walker, and N. Donaghue, *Social Cognition: An Integrated Introduction*, 2nd ed. London: SAGE Publications, 2006.
- [30] D. Austen-Smith and J. S. Banks, "Information aggregation, rationality, and the Condorcet jury theorem," *Am. Polit. Sci. Rev.*, vol. 90, no. 1, pp. 34–45, Mar. 1996.
- [31] P. K. Varshney, *Distributed Detection and Data Fusion*. New York: Springer-Verlag, 1997.
- [32] G. O. Wiredu, "A framework for the analysis of coordination in global software development," in *Proc. 2006 Int. Workshop Global Softw. Dev. Practitioner*, May 2006, pp. 38–44.
- [33] P. J. Coughlan, "In defense of unanimous jury verdicts: Mistrials, communication, and strategic voting," *Am. Polit. Sci. Rev.*, vol. 94, no. 2, pp. 375–393, Jun. 2000.
- [34] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA: MIT Press, 1991.
- [35] J. Marschak, "Elements for a theory of teams," *Manage. Sci.*, vol. 1, no. 2, pp. 127–137, Jan. 1955.
- [36] R. Radner, "Team decision problems," *Ann. Math. Stat.*, vol. 33, no. 3, pp. 857–881, Sep. 1962.
- [37] C. Noussair and J. Silver, "Behavior in all-pay auctions with incomplete information," *Games Econ. Behav.*, vol. 55, no. 1, pp. 189–206, Apr. 2006.
- [38] K. J. Knight, P. M. Cunio, J. D. Kwan, B. Bhushan, and E. B. Wagner, "Prizes for energy innovation: Incentives for today's challenges," Massachusetts Institute of Technology Industrial Performance Center, Tech. Rep. 10-001, Mar. 2010.
- [39] S. P. Anderson, J. K. Goeree, and C. A. Holt, "Rent seeking with bounded rationality: An analysis of the all-pay auction," *J. Polit. Econ.*, vol. 106, no. 4, pp. 828–853, Aug. 1998.
- [40] R. Radner, "Costly and bounded rationality in individual and team decision-making," in *Understanding Industrial and Corporate Change*, G. Dosi, D. J. Teece, and J. Chytrý, Eds. Oxford: Oxford University Press, 2005, pp. 3–35.
- [41] L. R. Varshney, "Unreliable and resource-constrained decoding," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, Jun. 2010.
- [42] G. A. Akerlof and R. E. Kranton, "Economics and identity," *Quart. J. Econ.*, vol. 115, no. 3, pp. 715–753, Aug. 2000.
- [43] R. G. Gallager, *Principles of Digital Communication*. Cambridge: Cambridge University Press, 2008.
- [44] C. N. Macrae and G. V. Bodenhausen, "Social cognition: Thinking categorically about others," *Annu. Rev. Psychol.*, vol. 51, pp. 93–120, Feb. 2000.

- [45] L. Cosmides, J. Tooby, and R. Kurzban, "Perceptions of race," *Trends Cogn. Sci.*, vol. 7, no. 4, pp. 173–179, Apr. 2003.
- [46] D. Chugh, "Societal and managerial implications of implicit social cognition: Why milliseconds matter," *Social Justice Research*, vol. 17, no. 2, pp. 203–222, Jun. 2004.
- [47] J. N. Wood, "Social cognition and the prefrontal cortex," *Behav. Cogn. Neurosci. Rev.*, vol. 2, no. 2, pp. 97–114, Jun. 2003.
- [48] G. Longo and G. Galasso, "An application of informational divergence to Huffman codes," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 1, pp. 36–43, Jan. 1982.
- [49] G. W. Allport, *The Nature of Prejudice*. Reading, MA: Addison-Wesley, 1954.
- [50] C. Van Laar, S. Levin, S. Sinclair, and J. Sidanius, "The effect of university roommate contact on ethnic attitudes and behavior," *J. Exp. Soc. Psychol.*, vol. 41, no. 4, pp. 329–345, Jul. 2005.
- [51] M. Rothbart and O. P. John, "Social categorization and behavioral episodes: A cognitive analysis of the effects of intergroup contact," *J. Soc. Issues*, vol. 41, no. 3, pp. 81–104, Fall 1985.
- [52] F. Echenique and R. G. Fryer, Jr., "A measure of segregation based on social interactions," *Quart. J. Econ.*, vol. 122, no. 2, pp. 441–485, May 2007.
- [53] J. N. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processing*, H. V. Poor and J. B. Thomas, Eds. Greenwich, CT: JAI Press, 1993, vol. 2, pp. 297–344.
- [54] H. A. David and H. N. Nagaraja, *Order Statistics*, 3rd ed. Hoboken, NJ: Wiley-Interscience, 2003.
- [55] D. Tapscott and A. D. Williams, *Macrowikinomics: Rebooting Business and the World*. New York: Portfolio Penguin, 2010.
- [56] D. DiPalantino and M. Vojnović, "Crowdsourcing and all-pay auctions," in *Proc. 10th ACM Conf. Electron. Commer. (EC'09)*, Jul. 2009, pp. 119–128.
- [57] M. R. Baye, D. Kovenock, and C. G. de Vries, "The all-pay auction with complete information," *Econ. Theor.*, vol. 8, no. 2, pp. 291–305, Jun. 1996.
- [58] B. Bhushan, "The incentives and disincentives of innovation prizes: A survey of the dropout teams from Progressive Insurance Automotive X PRIZE," S.M. thesis, Massachusetts Institute of Technology, Jun. 2010.
- [59] G. A. Akerlof, "The market for "lemons": Quality uncertainty and the market mechanism," *Quart. J. Econ.*, vol. 84, no. 3, pp. 488–500, Aug. 1970.