Toward a Theoretical Foundation for Multi-Agent Coordinated Decisions

Wynn C. Stirling, Michael A. Goodrich, and Richard L. Frost

Electrical and Computer Engineering Department
459CB Brigham Young University, Provo UT 84602
wynn@ee.byu.edu

Abstract

A system of M agents *coordinates* if the agents possess and use information regarding the existence, decisions, or decision-making strategies of each other. Inter-agent relationships may be expressed through a *coordination function*, which is a joint probability mass (or density) function defined over a $2M$-dimensional *coordination space*. The coordination function expresses the joint relationships between all agents in the system as characterized by agent goals and abductive considerations such as cost, hazard, and resource consumption. The coordination function is used to derive two joint utilities, termed *accuracy* and *rejectability*, and Levi's joint rule of epistemic utility is applied to identify the *satisficing set* of decision vectors that represent jointly rational behavior for the system. By intersecting the satisficing sets obtained under differing model assumptions, a robust set of satisficing decisions may be obtained.

Topic Areas: Cooperation and coordination; Conceptual and theoretical foundations.

Introduction

Coordination occurs with a multi-agent system if any of its members use information concerning the existence, decisions, or perceived decision-making strategies of other agents. To account for the variety of ways in which agents may interact, a formalism for multi-agent coordination would be useful. Such a formalism should characterize each agent's individual attributes, abilities, and goals. It should permit the modeling of inter-agent relationships in either an existing system or in a system to be synthesized. Furthermore, it should provide a means of decision-making for each agent in the system as a function of both the agent's individual agenda and its relationships to the other members of the system. Ideally, such a formalism should admit a convenient mathematical implementation, it should be computationally tractable, and it should lead to intuitive understanding.

Perfect coordination may occur with a multi-agent system only under very special circumstances. First, the system must be *logically closed*; that is, conditions of *hyperrationality* and *indeterminacy* must be satisfied. Hyperrationality means that all agents know <u>all</u> of the logical consequences of their assumed knowledge, and indeterminacy means that all agents know <u>only</u> the logical consequences of their assumed knowledge (Bacharach 1994). Second, the system must be *homogeneously rational*; that is, all agents must act according to the same concept of rationality. These conditions imply that the multi-agent system is perfectly modeled, all agents have access to the same information, and all are fully aware of each other's decision-making dispositions and capabilities. If any of these requirements are violated, a condition of *imperfect coordination* exists. Since most interesting problems involve situations of partial or imprecise information, we concentrate our attention to systems that involve imperfect coordination.

Two approaches currently dominate the study of multi-agent systems with imperfect coordination: (a) formalisms based on Bayesian decision theory, and (b) formalisms based on heuristics. A standard approach to address decision problems with incomplete information is to study Bayesian equilibrium points. This approach requires the specification of a prior distribution over a set of modeling assumptions, and adopting the decision that minimizes the Bayes risk. In particular, minimax theory, wherein each agent adopts the decision that minimizes its maximum risk, is a conservative approach to this problem that can be used to evaluate the robustness of the decision maker with respect to modeling uncertainty. The minimax approach assumes that nature (the agent who chooses the model structures) is noncooperative, and will attempt to thwart the performance of the agents by selecting the structure according to a least-favorable prior distribution.

Heuristic solutions are judged according to the reasonableness, practicality, and defensibility of the procedure used to arrive at the solution. Such solution techniques are especially applicable to cooperative systems, where communication is required but it is important to conserve resources. A well-designed heuristic solution will accommodate partial and imprecise information, and will provide a solution that, though perhaps not optimal, can be substantiated by the available evidence. Doing so without regard for performance, however, may lead to significantly decreased functionality. Of fundamental importance to the study of multi-agent systems, therefore, is the issue of making decisions in the presence of uncertainty that also exhibit acceptable performance.

In this paper we present an additional paradigm for multi-agent decision making. Building on Simon's notion of *satisficing* and Levi's theory of *epistemic utility*, the paradigm is a principle based, mathematically rigorous concept of a "middle ground" between optimal decisions based solely on precise performance criteria, and heuristic decisions based primarily on empirical considerations. We first present a summary of satisficing decision-making, we next illustrate how this theory applies to multi-agent systems, and we then demonstrate the application of this theory to a simple example.

Satisficing Decision-Making

Many researchers have advocated decision-making philosophies distinct from utility maximization, particularly for situations of limited resources and information (Simon 1955; Slote 1989; Cherniak 1986; Scheffler 1994). Perhaps the most well-known instantiation of such concepts is Simon's notion of *satisficing*, which requires the specification of an *aspiration level*, such that once this level is met, the corresponding solution is deemed adequate. The practical problem with this notion is the question of how to define the aspiration level: in the absence of principle-based justification, if one chooses to settle for a given aspiration level, is that not simply a heuristic judgment that the cost of further searching exceeds the benefits of the search? The aspiration level must be specified according to principles and must be made mathematically precise. Otherwise, satisficing is, at its root, a heuristic concept.

We present a formal definition of the aspiration level that is not based on heuristic considerations. This formalism leads naturally to a constructive procedure for identifying the satisficing decisions. The key to this development is the philosophical concept of epistemic utility.

Levi's *epistemic utility theory* (Levi 1967; 1980; 1984) asserts that any collection of cognitive propositions must be evaluated according to two independent desiderata: (a) truth value, and (b) informational value of rejection. The invocation of a truth metric is an essential component of any cognitive decision-making scheme, but Levi asserts that considerations of truth only are not adequate: for a proposition to be retained as a serious possibility it must have a high truth value *relative to* the informational value that would accrue to the agent if the proposition were to be rejected.

While truth value is a well-known and appreciated concept, the notion of rejection value may appear to be a bit novel. This concept has its roots in the notion of *abduction*, the philosophical idea of appealing to the best explanation. An inference that appraises a proposition only in terms of the value of the consequence, independent of considerations of truth, is an *abductive* inference. If a proposition were of of great value, then the agent naturally would desire to be true. Although this desire should not translate into an assertion of truth, it rightly would motivate the agent to obtain considerable evidence against the proposition before rejecting it.

Levi's theory implements the abductive concept by ascribing a metric for *informational value of rejection*, in addition to a metric for truth value, to a set of propositions to be investigated. These two metrics will generally be derived from independent principles. Incorporating both metrics into a decision-making procedure grants the agent a means of adjusting its decision to account for importance as well as credibility. To illustrate, suppose one were to receive a coded message that could be decoded, with equal probability of being correct, in two ways: m_1 = "You owe a one-dollar library fine," or m_2 = "You have won a million dollars, tax free." In terms of abductive considerations only, the person would be more prone to reject proposition p_1: *m_1 was sent*, than p_2: *m_2 was sent*, since rejecting p_1 implies more monetary reward than does rejecting p_2. Thus, p_1 has a higher informational value of rejection than does p_2.

Whereas classical decision theory typically invokes a single utility for ranking decisions, Levi's decision methodology employs two utility functions to characterize independently the truth value and rejection value of the propositions being considered. By structuring these utilities as probabilities, the agent apportions a unit of truth value and a unit of rejection value among the propositions, and thereby permits a direct comparison of each agent's truth value with its rejection value.

Let X be an agent, and let $(U, \mathcal{B})$ be a measurable space, where $U = \{u_\lambda, \lambda \in I\}$ is a *decision space*, with I being a countable index set. We restrict our development to the discrete case; the continuous case admits a parallel development, but will be omitted in the interest of brevity. U consists of all propositions that are under consideration by X. $\mathcal{B}$ is a σ-field of events in U; in this development, it is convenient to take $\mathcal{B}$ as as the power set of U. Let $(U, \mathcal{B}, P_A)$ be a probability space, where $P_A : \mathcal{B} \mapsto [0,1]$ is a probability describing the truth value of elements of $\mathcal{B}$. Also, let $(U, \mathcal{B}, P_R)$ be a probability space, where $P_R : \mathcal{B} \mapsto [0,1]$ is a probability describing the rejection value of elements of $\mathcal{B}$.

For any $b \in (0, \infty)$, the *expected epistemic utility* of a set $G \in \mathcal{B}$ is

$$\epsilon(G) = P_A(G) - bP_R(G) = \sum_{u \in G} [P_A(u) - bP_R(u)].$$

The set

$$S_b = \arg\max_{G \in \mathcal{B}} \epsilon(G) = \{u \in U : P_A(u) \geq bP_R(u)\}. \quad (1)$$

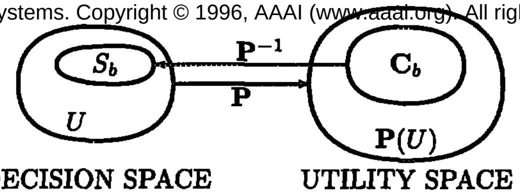
constitutes a decision set according to Levi's rule of epistemic utility (Levi 1980; Stirling & Morrell 1991):

Levi's Rule of Epistemic Utility

Given a set of propositions U, X should fail to reject all and only those propositions in the set S_b. The parameter b is termed the rejectivity *index, and reflects X's propensity for rejection.*

S_b is the *satisficing* set with respect to rejectivity b, and is the largest collection of propositions for which the truth value exceeds the rejection value (scaled by b). Equation (1) is called the *satisficing likelihood ratio test*, or SLRT. Any element of S_b is a *satisficing* decision. Figure 1 illustrates the structure of this decision rule. The two utility functions are mappings from the decision space, U, to a utility space, $\mathbf{P}(U)$, where $\mathbf{P}(u) = [P_A(u), P_R(u)]$. The SLRT generates a comparison set consisting of the region of utility space, $\mathbf{C}_b$, where accuracy dominates rejectability. The satisficing decisions then correspond to the inverse image of $\mathbf{C}_b$ under the utility mapping $\mathbf{P}$.

For $b = 1$, equal weight is placed on truth value and rejection value, and Levi's rule corresponds to the decision to fail to reject all propositions for which the truth value is at least as great as the benefit that would accrue should they be rejected (put in the vernacular, do not reject a proposition if the "good" outweighs the "bad"). The larger b is, the more propositions will be rejected. Setting $b = 0$ causes none of the propositions to be rejected. Letting $b \to \infty$ causes all of the propositions with positive rejection value to be rejected. Rejectivity is a design parameter in much the same way



$$\mathbf{P} = [P_A, P_R]$$
$$\mathbf{C}_b = \text{Comparison set} = \{[x, y] \in \mathbf{P}(U) : x \geq by\}$$
$$S_b = \mathbf{P}^{-1}(\mathbf{C}_b)$$

Figure 1: Illustration of the SLRT.

weighting factors are employed in conventional optimal decision theory to "tune" the performance of the procedure.

In its original cognitive context, Levi's approach to decision-making addresses the question of "what to believe." For practical decision-making, however, the context shifts to the question of "how to act." Consequently, it is necessary to interpret the notions of truth value and rejection value in a practical (that is, action, rather than belief) context. The following operational definitions characterize these notions:

Accuracy Truth value is replaced by *accuracy*, meaning conformity to a standard. Agents designed for practical purposes are usually *teleological*, or goal-oriented. The degree to which a possible action succeeds in achieving the goal is a measure of its accuracy. To illustrate, suppose a mobile robot has a goal of reaching a desired position, and attempts to reach this goal by implementing appropriate steering and throttle commands. Those commands that decrease the distance between it and the goal are more accurate than other commands.

Rejectability Actions may also be evaluated strictly in terms of their economic consequences. Resource consumption, elapsed time, exposure to hazard, and other such costs are consequences of taking action, regardless of the teleological consequences. The amount of benefit (such as fuel savings for a mobile robot) that accrues to the agent as a result of not taking a given action is the *rejectability* of the action.

The accuracy and rejectability properties of each element of U are encoded into the accuracy and rejectability utilities, P_A and P_R, respectively.

One of the most obvious differences between epistemic utility theory and conventional utility maximization-based decision theory is the introduction of two utilities, rather than one. A possible objection to this approach is that it over parameterizes the problem and introduces arbitrariness into the framework. This is a valid concern—parsimony is an essential feature of any well-designed decision rule, and if the goal of the theory were to maximize utility, the objection may have merit. But the goal of epistemic utility theory is not to compare propositions collectively, rank them according to an absolute standard, and accept only the best one. Rather, its goal is to *compare the attributes* of each proposition separately, perform a binary ranking for each, and fail to reject all that meet or exceed a threshold. Two utilities are essential for this purpose.

Although epistemic utility theory uses probability theory, it differs fundamentally from statistical theory in the way probability is interpreted. The discipline of statistics uses probability theory to model randomness in the behavior of data. By contrast, epistemic utility theory uses probability theory to generate the accuracy and rejectability utilities that embody the attributes, abilities, and goals of the system of agents. In other words, probability theory is employed as a *synthesis* tool to generate desired behavior, rather than an *analysis* tool to characterize observed behavior. Applying probability theory to synthesis rather than analysis represents a type of duality that can be exploited by reinterpreting for synthesis the various analysis techniques that have been developed for data modeling, such as Bayes theory and Markov theory.

A distinctive feature of the satisficing paradigm is that it results in a set of solutions, rather than a singleton, or point, solution. If action is to occur, however, a point-valued decision must be implemented. Therefore, S_b must be reduced to a single decision according to some tie-breaking mechanism. Simply choosing one of the members of the strongly satisficing set at random would be an acceptable tie-breaker. This approach may be unnecessarily arbitrary, however, since auxiliary information could be available that is not relevant to the definition of the aspiration level but may be useful for tie-breaking. For example, the agent may be of a conservative disposition, and choose, from all of the elements in S_b, the one with the lowest rejectability: $u_R = \arg\min_{u\in S_b}\{P_R(u)\}$. Alternatively, it may feel aggressive, and choose the one with the highest accuracy: $u_A = \arg\max_{u\in S_b}\{P_A(u)\}$. A compromise between these two extremes would be to chose the *maximally discriminating* decision rule: $u_D = \arg\max_{u\in S_b}\{P_A(u) - bP_R(u)\}$. All of these choices are satisficing, but none is intrinsically better than the others. These particular tie-breakers extend the methodology beyond the basic concepts of epistemic utility theory. Though, technically, they are constrained maxima, adopting such a tie-breaker does not constitute a reversion to the optimality paradigm. The choice of tie-breaker does not directly influence the solution methodology, and it may be changed without affecting the structure of the satisficing set. This feature allows the agent to delay the application of the tie-breaker until it is necessary to take action.

Multi-Agent Epistemic Utility

For the single-agent case, the decision space consists of *unipartite decisions* that pertain to only one agent. For the multi-agent case, the decision space consists of *multipartite decisions*, consisting of as many components as there are agents. Since epistemic utility theory is expressed in the language of probability, it is equipped to deal equally well with both single- and multi-agent decision problems. Single-agent decision problems require univariate accuracy and rejectability probabilities, and multi-agent decision problems require multivariate probabilities. The basic theory is unaffected by the dimensionality of the decision space. In fact, the rich methodology of multivariate probability and statistical analysis can be brought to bear on the coordination problem by suitably reinterpreting such fundamental notions as independence, correlation, conditioning, etc.

To set up the multi-agent coordinated decision problem in terms of epistemic utility theory for a system $\{X_1, \ldots, X_M\}$, we define the *multipartite decision space*

$$\mathbf{U}_{1\cdots M} = U_1 \times \cdots \times U_M,$$

where $U_i = \{u_{i\lambda_i}, \lambda_i \in I_i\}$ is the (unipartite) decision space for X_i. Elements of $\mathbf{U}_{1\cdots M}$ are denoted $\mathbf{u}_{1\cdots M} = (u_{1\lambda_1}, \ldots, u_{M\lambda_M})$, with element $u_{i\lambda_i}$ corresponding to the decision X_i adopts. The *joint accuracy* utility is a multivariate probability

$$P_{A_1\cdots A_M} : \mathbf{U}_{1\ldots M} \mapsto [0, 1],$$

and the *joint rejectability* utility is a multivariate probability

$$P_{R_1\cdots R_M} : \mathbf{U}_{1\ldots M} \mapsto [0, 1].$$

Accuracy and rejectability utilities, however, are not the fundamental entities of multi-agent epistemic utility theory. Although in the single-agent case it is usually assumed that accuracy and rejectability are modeled independently, in the multi-agent case there may well be dependencies between one agent's accuracy and another agent's rejectability. To account for this possibility, we consider a more fundamental quantity, which we term the *coordination* function.

To define the coordination function, we first introduce the *coordination space.* For an M-agent system, the coordination space is a $2M$-dimensional space $\mathcal{C}_M = (U_1 \times U_1) \times \cdots \times (U_M \times U_M)$. Elements of this space are *coordination vectors* of the form

$$\mathbf{c}_M = \{u_1^a, u_1^r; \ldots; u_M^a, u_M^r\}$$

where u_i^a and u_i^r represent the accuracy and rejectability variables[1], respectively, for agent X_i. The coordination function, denoted $C : \mathcal{C}_M \mapsto [0, \infty)$, is a probability and therefore admits the structure

$$C(u_1^a, u_1^r, \ldots, u_M^a, u_M^r) = P_{A_1 R_1 \cdots A_M R_M}(u_1^a, u_1^r, \ldots, u_M^a, u_M^r).$$

The coordination function characterizes both the accuracy and rejectability features. Hence, the joint accuracy and rejectability probabilities can be obtained as

$$P_{A_1 \cdots A_M}(u_1^a, \ldots, u_M^a) = \sum_{u_1^r \in U_1} \cdots \sum_{u_M^r \in U_M} C(u_1^a, u_1^r, \ldots u_M^a, u_M^r) \quad (2)$$

$$P_{R_1 \cdots R_M}(u_1^r, \ldots, u_M^r) = \sum_{u_1^a \in U_1} \cdots \sum_{u_M^a \in U_M} C(u_1^a, u_1^r, \ldots u_M^a, u_M^r), \quad (3)$$

For a system of M agents, the coordination function can be very complex, and it may be difficult to specify its structure. Since the function is a probability, however, we can apply Bayesian theory and express the function as the product of conditional probabilities. For example, suppose $M = 2$. The coordination function may be factored into products of conditional probabilities. For example, we may write

$$\begin{aligned} C(u_1^a, u_1^r, u_2^a, u_2^r) &= \\ & P_{A_1 R_1 | A_2 R_2}(u_1^a, u_1^r | u_2^a, u_2^r) P_{A_2 R_2}(u_2^a, u_2^r) \\ =\ & P_{A_1 | R_1 A_2 R_2}(u_1^a | u_1^r, u_2^a, u_2^r) P_{R_1 | A_2 R_2}(u_1^r | u_2^a, u_2^r) \\ & P_{A_2 | R_2}(u_2^a | u_2^r) P_{R_2}(u_2^r), \end{aligned}$$

where $P_{\cdot|\cdot}$ is a conditional probability. The advantage of factoring the coordination function into conditional components is that behavior conditioned on various situations is often easier to specify than unconditional joint behavior. As noted by Pearl (Pearl 1988), conditional probabilities permit local, or specific responses to be characterized; they possess modularity features similar to logical production rules.

[1]The quantities u_a^i and u_r^i may themselves be vectors, but we will refrain from encumbering the notation to account explicitly for that possibility.

Although traditional interpretations of conditional probability do not apply to these conditional expressions, they do admit very useful interpretations in the light of epistemic utility theory. The conditional probability $P_{A_1|R_1 A_2 R_2}(u_1^a | u_1^r, u_2^a, u_2^r)$ represents X_1's conditional accuracy given that X_2 places it's entire unit of accuracy mass on u_2^a and all of its rejectability mass on u_2^r, and X_1 places all of its rejectability mass on u_1^r. Also, $P_{R_1|A_2 R_2}(u_1^r | u_2^a, u_2^r)$ represents X_s's rejectability given that X_2 places its accuracy mass on u_2^a and its rejectability mass on u_2^r.

The conditional probability $P_{A_2|R_2}(u_2^a | u_2^r)$ represents X_2's accuracy given that it places its entire unit of rejectability mass on u_2^r. This function, as well as the unconditional accuracy probability P_{A_2}, represent what X_2's rejectability and accuracy postures would be in an environment devoid of other agents. In a single-agent environment, however, there should be no dependency between an agent's rejectability and its accuracy, so we may set $P_{A_2|R_2}(u_2^a | u_2^r) = P_{A_2}(u_2^a)$. The utilities P_{A_2} and P_{R_2} may be interpreted as *a priori*, or *pre-coordination*, probabilities.

In an ideal case, information and motivation would be present for a system $\{X_1, \ldots, X_M\}$ to function in a state of perfect coordination. Perfect coordination may be obtained in two ways: (a) an omniscient entity may orchestrate the entire multi-agent scenario, compute a joint decision for all agents, and communicate this information to each individual; or (b) each agent may possess a complete understanding of the goals, attributes, and capabilities of all other agents in the system, have access to all relevant information, and duplicate the decision-making process for itself.

Consider a multipartite decision space with joint accuracy and rejectability utilities given by (3) and (4), respectively. Let the joint rejectivity index be denoted $b_{1 \cdots M}$ and define Levi's joint rule of epistemic utility:

Levi's Joint Rule of Epistemic Utility

A system $\{X_1, \ldots X_M\}$, should fail to reject all and only those vector decisions in the set

$$\mathbf{S}_b = \{\mathbf{u}_{1 \cdots M} \in \mathbf{U}_{1 \cdots M} : P_{A_1 \cdots A_M}(\mathbf{u}_{1 \cdots M}) \geq b_{1 \cdots M} P_{R_1 \cdots R_M}(\mathbf{u}_{1 \cdots M})\}. \quad (4)$$

$\mathbf{S}_b$ is the *jointly satisficing* set, and elements of $\mathbf{S}_b$ are *jointly satisficing* decisions. Equation 5 is the *joint satisficing likelihood ratio test* (JSLRT).

A state of imperfect coordination exists in the presence of model uncertainty. In such circumstances, it is imperative that the decision-maker be *robust*, or hyposensitive to modeling errors. The set-valued structure of satisficing decision theory provides a natural mechanism for generating robust decisions: Consider a system $\{X_1, \ldots, X_M\}$ of agents. Let Θ denote a set modeling assumptions to be considered, and compute the accuracy and rejectability utilities for each $\theta \in \Theta$. Next, apply Levi's joint rule of epistemic utility to obtain the satisficing sets $\mathbf{S}_b^i(\theta)$, $i = 1, \ldots, M$. Using these sets, compute the *robust satisficing sets* for each agent as the intersection of the joint satisficing sets: $\mathbf{S}_b^i = \bigcap_\Theta \mathbf{S}_b^i(\theta)$. If $\mathbf{S}_b^i = \emptyset$, X_i is not a robust decision-maker, an important fact to know. If $\mathbf{S}_b^i \neq \emptyset$, it represents the set of all joint decision vectors that are satisficing. Let $\mathbf{u}_D^i$ denote the tie-breaking element of $\mathbf{S}_b^i$, and let u_D^i denote the i-th element of $\mathbf{u}_D^i$. Form the vector $\mathbf{u}^* = \{u_D^1, \ldots, u_D^M\}$. If $\mathbf{u}^* \in \bigcap_{i=1}^M \mathbf{S}_i$, then $\mathbf{u}^*$ is a *strongly robust* decision. If $\mathbf{u}^* \notin \bigcap_{i=1}^M \mathbf{S}_i$, then $\mathbf{u}^*$ is a *weakly robust* decision. The overall robustness of the system may be assessed by computing the set $\mathbf{S} = \bigcap_{i=1}^M \mathbf{S}_b^i$. If this set is not empty, then it is possible for the system to be a fully functional group of agents. If $\mathbf{S} = \emptyset$, the system as a whole is not capable of robust coordination. It may be true, however, that various sub-collections of agents have overlapping robust satisficing sets, in which case the corresponding sub-systems are capable of robust coordination within themselves.

Three-Agent Example

We illustrate this theory by applying it to a system consisting of three contestants (Shubik 1982, Page 24):

> SHOOTING MATCH. Contestants X_1, X_2, and X_3 fire at each other's balloon with pistols, from fixed positions. At the beginning, and after each shot, the players with unbroken balloons decide by lot who is to shoot next. The surviving balloon determines the winner.

The decision space for this problem is $\mathbf{U}_{123} = U_1 \times U_2 \times U_3$, where $U_i = \{u_{i1}, u_{i2}\}$, $i = 1, 2, 3$, with

$$\text{Agent } X_1: \begin{array}{lcl} u_{11} & = & \text{Shoot at } X_2\text{'s balloon} \\ u_{12} & = & \text{Shoot at } X_3\text{'s balloon} \end{array},$$

$$\text{Agent } X_2: \begin{array}{lcl} u_{21} & = & \text{Shoot at } X_1\text{'s balloon} \\ u_{22} & = & \text{Shoot at } X_3\text{'s balloon} \end{array},$$

$$\text{Agent } X_3: \begin{array}{lcl} u_{31} & = & \text{Shoot at } X_1\text{'s balloon} \\ u_{32} & = & \text{Shoot at } X_2\text{'s balloon} \end{array}.$$

Let p_1, p_2, and p_3 denote the probability of a hit for the players, and assume that $p_1 > p_2 > p_3$; that is, X_1 is the most skilled shooter and X_3 is the least skilled. An apparently rational approach to this problem is to follow a "domination" policy familiar from two-agent games, resulting in each player adopting the rule: "Shoot first at the balloon of your stronger opponent." As soon as one player is eliminated, the surviving players become strategic dummies and the game simplifies dramatically.

Shubik's critique of this solution reveals, however, that the domination policy leads to the disconcerting results, since it calls for the two weaker players to form a temporary coalition against the strongest player. For example, with $p_1 = 0.8$, $p_2 = 0.6$, and $p_3 = 0.4$, the probabilities of winning are 0.296, 0.333, and 0.370, for X_1, X_2, and X_3, respectively. Thus, the "rational" decision results with the poorest shot having the best chance of winning, and the best shot having the worst chance of winning!

For problems of this type, however, the social interaction between the agents should not be ignored. Consider question of how the most skilled player may assert its superiority. Suppose, as suggested in (Shubik 1982), that X_1 were to threaten X_3 by openly committing to shoot at X_3's balloon if ever X_3 shoots (and misses) his balloon, otherwise he will shoot at X_2's balloon. X_2 gains no advantage from this information, since it's best strategy is still to shoot at X_1's balloon.

If X_3 ignores the threat, the results are slightly more disastrous for X_1, since the probabilities of winning are 0.282, 0.397, and 0.321, respectively, with X_2 clearly having the best chance to win and X_1 being the clear loser. If, however, X_3 succumbs to the threat, the resulting probabilities of winning are 0.444, 0.200, and 0.356, respectively. Thus, X_1 gains the best chance of winning, and X_2 has the worst chance of winning under this scenario.

Classical game theory may be made to accommodate social interaction through the introduction of suitable parameters, such as the preemption levels described in (Brams & Kilgour 1987). A general approach to problems such as this in a game-theoretic framework, however, is not universally recognized. Such issues may be introduced naturally, however, in the framework of epistemic utility, through the specification of the coordination function.

To form the coordination function for this problem, we must first provide clear definitions for accuracy and rejectability. In accordance with our operational definition, we take the standard for accuracy to be the domination principle. Thus, the decision to shoot at the stronger opponent's balloon is considered more accurate than the decision to shoot at the weaker opponent's balloon. (Accuracy, in this context, is not equivalent to the shooting skill of the agent; rather, it deals with the other agents' skill in shooting at it.) To define rejectability, we must identify aspects of the decision that are undesirable--for this problem, the aspect is retaliation. Thus, a decision that leads to retaliation will have higher rejectability than a decision that does otherwise.

We first consider behavior in the absence of coordination. In this situation, each agent makes a unilateral decision independently of the behavior of other agents, so we may construct three separate single-agent decision problems. We will use the hit probabilities, p_i, $i = 1, 2, 3$, to generate the accuracy utilities. For example, X_1 will generate its pre-coordination accuracy function as

$$P_{A_1}(u_1^a) = \begin{cases} \frac{p_2}{p_2+p_3} & u_1^a = u_{11} \\ \frac{p_3}{p_2+p_3} & u_1^a = u_{12} \end{cases},$$

with similar expressions for X_2 and X_3. In the absence of coordination, revenge is moot, a circumstance reflected in assigning the uniform distribution to the rejectability function. Table 1 illustrates the three single-agent accuracy and rejectability functions for $p_1 = 0.8$, $p_2 = 0.6$, $p_3 = 0.4$. Application of Levi's single-agent

X_1			X_2			X_3		
U_1	P_{A_1}	P_{R_1}	U_2	P_{A_2}	P_{R_2}	U_3	P_{A_3}	P_{R_3}
u_{11}	0.6	0.5	u_{21}	0.67	0.5	u_{31}	0.57	0.5
u_{12}	0.4	0.5	u_{22}	0.33	0.5	u_{32}	0.43	0.5

Table 1: Pre-coordination accuracy and rejectability.

rule of epistemic utility with unity rejectivity ($b = 1$) reveals behavior consistent with the domination argument: X_1 shoots at X_2, X_2 shoots at X_1, and X_3 shoots at X_1.

Now let us postulate some structure for the coordination function. The social climate is such that X_1 announces a threat to X_3. X_3 has the option of either ignoring the threat or succumbing to it. X_2 is not influenced by the threat; it would prefer to eliminate X_1 and take its chances with X_3. Thus, X_2's coordinated accuracy and rejectability functions are not influenced by either of the other agents.

An obvious way X_1 could carry out its threat is to change its rejectability should X_3 ever shoot at X_1's balloon. Suppose the game starts at time t_0. Let $P_{R_1}(\cdot; t_0)$ denote X_1's initial rejectability, and suppose X_1 is prepared to switch it's rejectability to a new value if X_3 fires at X_1's balloon. We may represent this situation by permitting the rejectability function to be time-dependent. As long as X_3 succumbs to the threat, the coordination function remains constant, and we express this condition as

$$P_{R_1}(u_1^r; t+1) = P_{R_1}(u_1^r; t), t = t_0, t_0 + 1, \ldots.$$

Suppose, at time $t_1 \geq t_0$, X_3 were to shoot at X_1's balloon (and miss). If X_1 were to carry out its threat, it could do so by modifying its rejectability utility to become

$$P_{R_1}(u_1^r; t) = \begin{cases} 1 & u_1^r = u_{11} \\ 0 & u_1^r = u_{12} \end{cases}, t = t_1 + 1, t_1 + 2, \ldots.$$

X_1's attempt to benefit from its superior skill leaves it vulnerable to X_3's option to either ignore the threat or to succumb. Let $\gamma \in [0, 1]$ denote X_3's propensity to succumb. It accordingly defines the conditional accuracy utility $P_{A_3|R_1}$, as

$$P_{A_3|R_1}(u_3^a|u_{11}) = \begin{cases} 1-\gamma & u_3^a = u_{31} \\ \gamma & u_3^a = u_{32} \end{cases},$$

and keeps $P_{A_3|R_1}(u_3^a|u_{12})$ unchanged. We assume that X_3's rejectability function is independent of X_1, and that other than carrying out its threat if need be, X_1's accuracy and rejectability functions are independent of X_3 and X_2. Consequently, the coordination function factors as

$$\begin{aligned} &C(u_1^a, u_1^r, u_2^a, u_2^r, u_3^a, u_3^r; t) = \\ &\quad P_{A_1}(u_1^a) P_{A_2}(u_2^a) P_{A_3|R_1}(u_3^a|u_1^r) P_{R_1}(u_1^r; t) \\ &\quad \times P_{R_2}(u_2^r) P_{R_3}(u_3^r) \end{aligned}$$

Applying Levi's joint rule of epistemic utility, with $b_{123} = 1$, yields the set of jointly satisficing decision vectors $\mathbf{S}_b$. When $\gamma < 0.5$, the maximally discriminating element of $\mathbf{S}_b$ is $\{u_{11}, u_{21}, u_{31}\}$, and for $\gamma > 0.5$, the maximally discriminating element is $\{u_{11}, u_{21}, u_{32}\}$. Thus, if X_3 is more prone to succumb than to ignore X_1's threat, then the threat is effective and X_1 benefits. Otherwise, X_1 fares poorly as a result of making the threat. Of course, X_1 could also resort to trickery, in that it makes the threat but never intends to actually carry it out. This situation is easily modeled by simply not switching the value of P_{R_1} in the event X_3 shoots at X_1's balloon.

This simple example illustrates how rich and complex multi-agent decision-making can become, and it demonstrates the facility of epistemic utility for synthesizing social structure. For another example of using epistemic utility theory as an alternative to classical game-theoretic formulations, see (Moon, Frost, & Stirling 1996) for a discussion of the well-known Prisoner's Dilemma game.

Discussion

Epistemic utility theory provides a decision mechanism possessing two key features that make it attractive for multi-agent decision-making. First, it is a satisficing, rather than an optimizing paradigm. Although optimization is a compelling concept for single-agent scenarios where complete information is available, it loses much of its force in multi-agent contexts, particularly in situations of partial information. Satisficing decisions, on the other hand, are, by construction, compatible with the amount of information available. They are inherently comparative, rather than superlative in nature, and identify the set of good, but not necessarily best, decisions.

A second key feature of epistemic utility in the multi-agent context is that it permits modeling of the social interactions between the agents by means of the coordination function, which may be factored into conditional accuracy and rejectability functions via Bayes theory. This feature provides a means of constructing mathematical expressions for non-optimal and idiosyncratic behavior.

As developed thus far, epistemic utility-based coordination theory is embryonic. Some of the key questions that remain to be addressed include the following issues.

1. It must deal with imperfect coordination, where each agent has its own perception of other agents' attributes, abilities, and goals that is subject to error. This problem is at the crux of any viable theory of coordination. Epistemic utility theory, with its foundations couched in probability, provides a medium to characterize the types of perceptual errors that may occur within a system, and to analyze the resulting performance of individual agents and of the system as a whole.

2. Each agent should be capable of learning from its own and other agents' performances within the system. Learning may be accomplished in the epistemic utility context by using this information to modify the coordination function. A natural way to do this is to use learning automata theory to adapt the conditional and marginal probabilities that comprise the coordination function. The learning of these structures corresponds to the learning of principles, or *cognitive learning*, rather than *response learning*, which does not rely on knowledge of the internal workings of an agent.

3. It should be capable of iterative decision-making, or deliberation. In this process, an agent makes a decision and then calculates its effect on other agents, predicts their responses, and uses this information to re-evaluate its own decision. By repeating this process, the agent may converge to an equilibrium decision, and thereby improve its performance. The successful application of deliberation to the Prisoner's Dilemma problem (viewed as a two-agent coordinated decision problem) is discussed in (Moon, Frost, & Stirling 1996).

Historically, much of the development of probability theory has been devoted to the *analysis of observed behavior* for estimation and prediction of performance. Epistemic utility-based coordination theory, however, employs probability theory primarily to express principles for the *synthesis of desired behavior.* There appears to be a type of duality between these two concepts, much in the spirit of the duality that exists between observation and control in classical system theory. Exploration of this duality may lead to the development of a viable theory for the synthesis of multi-agent societies.

References

Bacharach, M. 1994. The Epistemic Structure of a Theory of a Game. *Theory and Decision* 37:7–48.

Brams, S. J., and Kilgour, D. M. 1987. Optimal Threats. *Operations Research* 35(4):524–536.

Cherniak, C. 1986. *Minimal Reality.* MIT Press.

Levi, I. 1967. *Gambling with Truth.* Cambridge, Massachusetts: MIT Press.

Levi, I. 1980. *The Enterprise of Knowledge.* Cambridge, Massachusetts: MIT Press.

Levi, I. 1984. *Decisions and Revisions.* London: Cambridge University Press.

Moon, T. K.; Frost, R. L.; and Stirling, W. C. 1996. An Epistemic Utility Approach to Coordination in the Prisoner's Dilemma. *BioSystems* 37(1-2):167–176.

Pearl, J. 1988. *Probalistic Reasoning in Intelligent Systems.* San Mateo, Ca: Morgan Kaufmann.

Scheffler, S. 1994. *The Rejection of Consequentialism (Revised Edition).* Oxford University Press.

Shubik, M. 1982. *Game Theory in the Social Sciences.* Cambridge, Massachusetts: MIT Press.

Simon, H. A. 1955. A Behavioral Model of Rational Choice. *Quart. J. Economics* 59:99–118.

Slote, M. 1989. *Beyond Optimizing.* Harvard University Press.

Stirling, W. C., and Morrell, D. R. 1991. Convex Bayes Decision Theory. *IEEE Transactions on Systems, Man, and Cybernetics* 21(1):173–183.