Conditional Games: A Framework for Coordination

Wynn Stirling

***Abstract*—Game theory is often proposed as a framework within which to model coordination. Neoclassical game theory, however, focuses exclusively on individual preferences, whereas coordination requires a concept of group preference as well as individual preferences. Conditional game theory differs from classical theory in two fundamental ways. First, it involves a utility structure that permits agents to define their preferences conditioned on the preferences of other agents, and second, it accommodates a notion of group rationality as well as individual rationality. The resulting framework permits a notion of group preferences to be defined, and leads to the development of a metric to characterize the intrinsic ability of the members of a group to coordinate.**

***Index Terms*—GameTheory, Multiagent Systems, Utility Theory**

I. INTRODUCTION

Coordination is a sophisticated and challenging capability for a multiagent system to achieve. Indeed, the dictionary definition of *coordinate* is instructive: "To place or arrange (things) in proper position relative to each other and to the system of which they form parts; to bring into proper combined order as parts of a whole" [1]. The reason coordination can be challenging is that it requires rational behavior of the individuals (the parts) to be reconciled with rational behavior for the system (the whole). Since game theory provides a mathematical framework within which to synthesize rational behavior among multiple decision makers, it has emerged as an important mathematical tool with which to characterize coordination for multiagent systems [2–8].

In this paper we restrict attention to finite strategic (normal-form) games; that is, single-stage games such that the players simultaneously take only one action each, after which each receives its payoff. For many scenarios, a multi-stage game can be described by a series of single-stage games, particularly when the key issue of concern is that of coordination.

Formally, a finite single-stage strategic game consists of (i) a set of autonomous decision makers, or *players*, denoted $\boldsymbol{\mathcal{X}} = \{X_1, \ldots, X_n\}$, where $n \geq 2$, (ii) a finite action set $\mathcal{A}_i$ for each X_i, and (iii) a utility $u_{X_i}\colon \boldsymbol{\mathcal{A}} \to \mathbb{R}$ for each X_i, $i = 1, \cdots, n$, where $\boldsymbol{\mathcal{A}} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_n$ is the product action space. For any *action profile* $\mathbf{a} = (a_1, \ldots, a_n) \in \boldsymbol{\mathcal{A}}$, the utility u_{X_i} provides a cardinal preference ordering for X_i over $\boldsymbol{\mathcal{A}}$. These utilities are *categorical* in the sense that u_{X_i} unconditionally defines the preference ordering for X_i, regardless of the preferences of other players.

In addition to the categorical structure of the utilities, it is usually assumed that each X_i possesses a logical structure that defines how it should play the game. The most widely used logical structure is the doctrine of *individual rationality*: each X_i should act in a way that maximizes its own utility, regardless of the effect doing so has on others. Under the assumption that each player subscribes to this notion and believes that all others do so as well, they each will solve their corresponding constrained optimization problem, resulting in a Nash equilibrium.

Rational behavior for the individuals, however, does not necessarily correspond to rational behavior for the group. Although individuals are free to substitute the interests of others for their own, doing so simply redefines the game without changing the solution concept, which is still based on individual rationality and therefore does not admit an explicit notion of group preference. As Luce and Raiffa observe, "general game theory seems to be in part a sociological theory which does not include any sociological assumptions . . . it may be too much to ask that any sociology be derived from the single assumption of individual rationality" [9, p. 196]. Shubik points out that, even though a group may arrive at a decision, the choice does not necessarily correspond to a preference for the group: "It may be meaningful, in a given setting, to say that a group 'chooses' or 'decides' something. It is rather less likely to be meaningful to say that the group 'wants' or 'prefers' something" [10, p. 124]. Conventional game theory eschews the notion of group preference, primarily on the grounds that the group is not a "superplayer" that possesses the power to make decisions.

Coordination is not the same as cooperation. While agents may cooperate to their mutual advantage, that does not necessarily mean that a concept of preference for the group as a whole exists. It might be expected that the solution to a so-called cooperative game (a game such that players may enter into binding agreements) would possess some notion of group rationality. A cooperative game permits a subset of players to enter into a coalition such that each receives a payoff that is greater than it would receive if it acted alone. However, each player enters into a coalition solely on the basis of benefit to itself and, even though each may be better off for having joined, no notion of group benefit is relevant.

True coordination, however, requires a notion of preference for the group as well as for its members, and is an important consideration when designing a multiagent system that is intended to serve some useful purpose. The challenge with a game-theoretic approach, therefore, is to define utilities for the agents and a solution concept that together lead to simultaneously acceptable global (system level) and local (individual level) performance.

Historically, game-theoretic treatments of coordination, as discussed by [11–17]. are based upon categorical preference orderings. When the payoffs of the players are juxtaposed

Wynn Stirling is with the Faculty of Electrical and Computer Engineering, Brigham Young University, 105 FPH, Provo, Utah, USA, 84602 wynn@ee.byu.edu.

978-1-4244-8738-7/11/$26.00

in a payoff array, behavioral attributes such as exploitation, compromise, altruism, and cooperation become apparent. Once the payoffs are defined, however, the rationale behind their generation is irrelevant to the solution. Unfortunately, it is exactly these behavioral attributes that come into play when seeking a "coordinated" solution. Schelling observes that, "In the pure-coordination game, the player's objective is to make contact with the other player through some imaginative process of introspection, of searching for shared clues" [11, p. 17]. A condition of "tacit coordination" occurs when ... "one is trying to guess what the other will guess ones self to guess the other to guess, and so on ad infinitum" [11, pp. 92-93]. Along these same lines, Bicchieri argues that coordination is unlikely unless the players have significant social knowledge about each other, either known *a priori* or obtained as a result of learning [13]. Coordination, under this view, is the result of social evolution. These concepts of coordination require the players to know things about each other in addition to the information contained in the payoff array.

Unless such information is encoded in the payoffs, however, it is not part of the formal structure of the game. We shall term coordination that derives from the extra-game knowledge that the players have regarding each other *extrinsic coordination*, since it is derived exogenously from considerations that are not part of the formal structure of the game as defined by the payoffs. Thus, the notions of coordination espoused by Schelling and Bicchieri, as well as others, are concepts of extrinsic coordination.

The ability to coordinate extrinsically depends upon the propensities, introspective skills, and concepts of rationality employed by the players as they attempt to construct a notion of group preference. While such behavior may lead to the effective playing of a game for humans, artificial agents will not possess such propensities and skills unless they are explicitly programed into their logic.

In contrast to extrinsic coordination, we may also consider the concept of *intrinsic coordination*, which arises due to the social influences that exist among the members of a group as objectively encoded in the preference relationships, rather than derived from the subjective social attitudes and rationality concepts of the players. If all members possess categorical utilities, then no notion of group preference exists and, consequently, they possess no intrinsic ability to coordinate (although they may cooperate if their payoffs are compatible). However, if the members are able to influence each other's preferences, then a notion of group preference may emerge, and the members may be able to coordinate.

We view coordination as a neutral concept; neither necessarily positive, such as cooperation, nor negative, such as competition (such as would occur with an athletic contest or a military engagement). Instead, we shall comply with the dictionary definition and require coordination to involve some form of explicit sociality among the agents. Our goal is twofold. First, we extend the utility structure to account explicitly for social relationships and, second, we use those utilities to generate simultaneous notions of group and individual preference. We then introduce a metric to define an intrinsic measure of the group's ability to coordinate that is independent of the notions of rationality, propensities, skill, and all other factors that are not encoded into the utility structure. This metric will address the ecological issue of how fit, or suitable, a multiagent system is to function in its environment.

II. INFLUENCE

Neoclassical decision theory is not concerned with the process of defining utilities; it is only concerned with how they are used once defined. Friedman put it this way: "The economist has little to say about the formation of wants; this is the province of the psychologist. The economist's task is to trace the consequences of any given set of wants" [18, p. 13]. Consequently, any social influence involved in the process of defining utilities becomes irrelevant once they are defined.

Our approach, however, is to move further upstream toward the headwaters of how preferences are formed by explicitly accounting for the possibility that the preferences of others may influence an agent's preferences. We say that X_j *influences* X_i if X_i's preferences are affected by X_j's preferences. Without knowledge of X_j's preferences, X_i is in a state of suspense with respect to its own preferences. Essentially, X_j's preferences propagate through the group to affect X_i's preferences, thereby generating a social bond between them. Once such a bond exists, it is possible to define a notion of joint preference for the two agents viewed simultaneously, and it is possible to extract individual preference orderings from this joint preference ordering since, once X_j's preferences are revealed, X_i need no longer remain in suspense. It is thus possible for both group and individual preferences to co-exist. We illustrate this concept with a simple example.

When considering a multiagent system, a key issue is the degree of harmony, or concordance, that exists within the community as its members influence each other in various ways. It may happen that the social bonds are so strong that unanimity will result, that is, that all X_i agree that some $\mathbf{a}^*$ is best for all. At the other extreme, the agents may be so diametrically opposed to each other that the group will be dysfunctional. Although the group's behavior may not be near either of these extremes, some degree of conflict, or discord, will generally exist. Thus, when considering a system whose members must coordinate, a critical issue is the concordance, or the degree of harmony, among its members as a function of their individual preferences. To formalize a concept of concordance, we introduce the following definitions.

Definition 1 Let $\boldsymbol{\mathcal{X}} = \{X_1, \ldots, X_n\}$ be a multiagent system. A *conjecture* for X_i is an action profile, denoted $\mathbf{a}_i$, that is hypothesized as X_i's most preferred outcome.

A *joint conjecture* for $\boldsymbol{\mathcal{X}}$ is an n^2-dimensional vector of action profiles, denoted $(\mathbf{a}_1, \ldots, \mathbf{a}_n) \in \boldsymbol{\mathcal{A}}^n$, where $\mathbf{a}_i$ is a conjecture for X_i, $i = 1, \ldots, n$. □

Definition 2 For a multiagent system $\boldsymbol{\mathcal{X}} = \{X_1, \ldots, X_n\}$, a *concordant utility* $U_{\boldsymbol{\mathcal{X}}}$ is a real-valued function defined over

$\mathcal{A}^n$ that corresponds to the degree of harmony for the system for each joint conjecture $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$. □

The concordant utility is a generalization of individual utility which, rather than providing a preference ordering for a single agent over the space $\mathcal{A}$ of action profiles, provides a concordance ordering for the entire group over the product space $\mathcal{A}^n$ of joint conjectures. When $\mathbf{a}_1 = \cdots = \mathbf{a}_n$, the concordant utility measures the degree of harmony if all members conjecture the same action profile. When the conjectures are different, $U_{\mathcal{X}}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ measures the degree of concordance that exists among the members of the group if each X_i were to view $\mathbf{a}_i$ as its most-preferred outcome. The expression $U_{\mathcal{X}}(\mathbf{a}_1, \ldots, \mathbf{a}_n) > U_{\mathcal{X}}(\mathbf{a}'_1, \ldots, \mathbf{a}'_n)$ means that $(\mathbf{a}'_1, \ldots, \mathbf{a}'_n)$ causes a more severe conflict for the group than does $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$.

The concordant utility provides a measure of the severity of disputes among the agents, and thus, provides a notion of group preference; namely, that some intrinsic notion of consistency and congruity is desirable. Such a notion of group preference may not lend itself to an easily understood operational interpretation. Rather, it is an emergent notion that arises as the members of the group interact, and accounts for the combined effects of social attributes such as propensities toward cooperation, competition, selfishness, altruism, and so forth.

III. CONDITIONAL UTILITIES

We now address the issue of how to construct a concordant utility. Clearly, simply defining such a complex function from some set of first principles or desired behavior would quickly become intractable, due to the complexity of such an undertaking. An alternative and more tractable approach is to synthesize a concordant utility from simpler components. To motivate this approach, we draw an analogy from probability theory, and consider the construction of a multivariate probability mass function. Rather than construct such a function directly, the more usual approach is to construct it from conditional mass functions via the chain rule.

Our approach is to apply a similar procedure to utility theory by introducing the notion of a conditional utility. A natural mathematical mechanism with which to account for social influence is to expand the notion of utility in a way that permits such influence to be explicitly modeled.

To illustrate, suppose that X_1 influences X_2; that is, X_2's preferences are affected by X_1's preferences,and let us consider the following hypothetical proposition: given the antecedent that X_1 conjectures $\mathbf{a}_1$, the consequent is the preference ordering for X_2 under that hypothesis. We may formalize this notion as follows.

Definition 3 Let $\{X_{i_1}, \ldots, X_{i_{p_i}}\}$ be a subgroup of $\mathcal{X}$ that influences X_i and let $(\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_{p_i}})$ be a joint conjecture for $\{X_{i_1}, \ldots, X_{i_{p_i}}\}$. A *conditional utility* $u_{X_i|X_{i_1}\cdots X_{i_{p_i}}}(\cdot|\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_{p_i}})$ is a real-valued function defined over $\mathcal{A}$ that specifies the preference ordering for X_i given $(\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_{p_i}})$. That is, $u_{X_i|X_{i_1}\cdots X_{i_{p_i}}}(\mathbf{a}_i|\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_{p_i}}) > u_{X_i|X_{i_1}\cdots X_{i_{p_i}}}(\mathbf{a}'_i|\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_{p_i}})$ means that X_i prefers $\mathbf{a}_i$ to $\mathbf{a}'_i$, given that X_{i_l} conjectures $\mathbf{a}_{i_l}$, $l = 1, \ldots, p_i$. □

Example 1 Suppose X_1 and X_2 are to purchase an automobile. X_1 is to choose the manufacturer, either foreign (F) or domestic (D), and X_2 is to choose the model, either a convertible (C) or a sedan (S). Let $\mathcal{A}_1 = \{F, D\}$ and $\mathcal{A}_2 = \{C, S\}$, and let $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2$ denote the product action space.

Let X_1 define a categorical utility u_{X_1}, and let X_2 define a conditional utility $u_{X_2|X_1}(\cdot|\mathbf{a}_1)$ for each $\mathbf{a}_1 \in \mathcal{A}$. The expression $u_{X_2|X_1}(F, C|F, C) > u_{X_2|X_1}(D, S|F, C)$ then means that X_2 prefers a foreign-made convertible to a domestic-made sedan, given the hypothesis that X_1 most prefers a foreign-made convertible. The important feature of this model structure is that it allows X_2 to define its preferences given a hypothesis concerning X_1's preferences without actually knowing X_1's true preferences. □

IV. AGGREGATION

Consider a multiagent system $\mathcal{X} = \{X_1, \ldots, X_n\}$, and suppose each X_i is influenced by a subset $\{X_{i_1}, \ldots, X_{i_{p_i}}\}$ of $\mathcal{X}$, and there fore possesses a conditional utility of the form $u_{X_i|X_{i_1}\cdots X_{i_{p_i}}}$. If $p_i = 0$, then X_i possesses a categorical utility u_{X_i}. Our goal is to develop a forumla to combine these utilities to form the concordant utility. To proceed, we first introduce the notion of subjugation and offer the following principle to guide our search for an acceptable aggregation formula.

Coherence. *The interests of no individual should be categorically subjugated to the interests of the group in all situations.*

The coherence principle means that no agent will be categorically disenfranchised, meaning that it is placed in a situation such that no choice that is good for the group will be good for it.

Definition 4 A multiagent system is *socially coherent* if no agent can be subjugated; that is, if $\mathbf{a}_i$ is preferred to $\mathbf{a}'_i$ by X_i, then there must exist a joint conjecture $(\mathbf{a}^*_1, \ldots, \mathbf{a}^*_{i-1}, \mathbf{a}^*_{i+1}, \ldots, \mathbf{a}^*_n)$ for the subgroup $\mathcal{X} \setminus \{X_i\}$ such that $(\mathbf{a}^*_1, \ldots, \mathbf{a}^*_{i-1}, \mathbf{a}_i, \mathbf{a}^*_{i+1}, \ldots, \mathbf{a}^*_n)$ is preferred by the group to $(\mathbf{a}^*_1, \ldots, \mathbf{a}^*_{i-1}, \mathbf{a}'_i, \mathbf{a}^*_{i+1}, \ldots, \mathbf{a}^*_n)$. □

Avoiding individual subjugation ensures that the preferences of an individual are not so irreconcilably contrary to the preferences of the system that, no matter what the system prefers, the preferences of the affected individual can never be accommodated. Although individual subjugation is not always avoided in societies, avoiding subjugation is an important feature of a system that is designed to cooperate.

The question then becomes: what conditions are necessary to impose on the utilities to ensure that individual subjugation does not occur? To address this question, let us turn to an analogous issue. A Dutch book is a gambling situation such that, no matter what the outcome, the gambler will be worse off for having taken the gamble – a situation of sure loss (ones reward is always less than ones stake). A set of preferences is said to be coherent if it is not possible to construct a Dutch book. The Dutch Book Theorem [19, 20] and its converse [21] establish that a belief system is coherent if and only if it complies with a probability measure that describes the degrees of belief regarding the propositions under consideration.

We may now form an exact analogy between the notion coherence in the sense defined by the Dutch Book Theorem and the notion of social coherence introduced above. We thus see that, in order to assure social coherence, the utility structure must conform the structure of probability mass functions. Accordingly, we shall assume that all utilities are mass functions — they are non-negative and normalized to sum to unity; that is, for every X_i and every p_i-element subset $\{X_{i_1}, \ldots, X_{i_{p_i}}\} \subset \boldsymbol{\mathcal{X}} \setminus \{X_i\}$,

$$u_{X_i|X_{i_1}\cdots X_{i_{p_i}}}(\mathbf{a}_i|\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_{p_i}}) \geq 0 \;\forall\; \mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_{p_i}},$$

$$\sum_{\mathbf{a}_i} u_{X_i|X_{i_1}\cdots X_{i_{p_i}}}(\mathbf{a}_i|\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_{p_i}}) = 1 \;\forall\; \mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_{p_i}}.$$

To comply with the probabilistic structure, we shall also assume that the influence relationships among the members of the group are acyclical, that is, the group can be represented by a directed acyclic graph (DAG). Accordingly, each member of the group is influenced by, and only by, its parents. We may then express the concordant utility via the chain rule, yielding

$$U_{\boldsymbol{\mathcal{X}}}(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \prod_{i=1}^{n} u_{X_i|\,\mathrm{pa}\,(X_i)}[\mathbf{a}_i|\,\mathrm{pa}\,(\mathbf{a}_i)], \qquad (1)$$

where $\mathrm{pa}\,(\mathbf{a}_i) = \{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_{p_i}}\}$ is the joint conjecture corresponding to $\mathrm{pa}\,(X_i) = \{X_{i_1}, \ldots, X_{i_{p_i}}\}$, the p_i parents of X_i. If $p_i = 0$, then $u_{X_i|\,\mathrm{pa}\,(X_i)} = u_{X_i}$, a categorical utility.

If the utilities of all agents are categorical, then a concordant utility provides no more information about the system than do the individual utilities. This result is analogous to the probabilistic result that, if events A, B, and C are mutually independent and P is a probability measure, then $P(A \cap B \cap C)$ provides no more information than do the probabilities of the individual events. Appropriating terminology from the probability context, we say that the agents X_i, $i = 1, \ldots, n$ are *mutually independent* if all utilities are categorical.

Definition 5 A *conditional game* is a triple $\{\boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{A}}, U_{\boldsymbol{\mathcal{X}}}\}$ where $\boldsymbol{\mathcal{X}} = \{X_1, \ldots, X_n\}$ is a group of n agents with product action space $\boldsymbol{\mathcal{A}} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_n$ and $U_{\boldsymbol{\mathcal{X}}}$ is a concordant utility. Equivalently, by application of (1), a conditional game can be defined in terms of the conditional utilities $u_{X_i|\,\mathrm{pa}\,(X_i)}$, $i = 1, \ldots, n$. □

If all utilities are categorical, a conditional game becomes a conventional game. The concordant utility encapsulates all of the social relationships that exist in the group but, since it is a function of multiple action profiles and only one action profile can actually be implemented, we must extract information from it that will define rational behavior for the group and for the individuals. To proceed, we observe that, since each agent can control only its own actions, what is of interest is the utility for the group if all agents *make conjectures over, and only over, their own action spaces*.

Definition 6 Consider the concordant utility $U_{\boldsymbol{\mathcal{X}}}$ and let a_{ij} denote the j-th element of $\mathbf{a}_i$; that is, $\mathbf{a}_i = (a_{i1}, \ldots, a_{in})$ is X_i's conjecture. Next, form the action profile $(a_{11}, , a_{nn})$ by taking the i-th element of each X_i's conjecture, $i = 1, \ldots, , n$. Now let us sum the concordant utility over all elements of each $\mathbf{a}_i$ except the ii-th elements to form the *group welfare function* for $\{X_1, \ldots, X_n\}$, yielding

$$w_{\boldsymbol{\mathcal{X}}}(a_{11}, \ldots, a_{nn}) = \sum_{\neg a_{11}} \cdots \sum_{\neg a_{nn}} U_{\boldsymbol{\mathcal{X}}}(\mathbf{a}_1, \ldots, \mathbf{a}_n), \qquad (2)$$

where the notation $\sum_{\neg a_{ii}}$ means that the sum is taken over all a_{ij} except for $j = i$.

The *individual welfare function* for X_i is the i-th marginal of $w_{\boldsymbol{\mathcal{X}}_n}$, that is,

$$w_{X_i}(a_{ii}) = \sum_{\neg a_{ii}} w_{\boldsymbol{\mathcal{X}}}(a_{11}, \ldots, a_{nn}). \qquad (3)$$

□

The group and individual welfare functions provide a complete *ex post* description of the relationships between the members of a multiagent system as characterized by their *ex ante* conditional utilities, and thus may be used to form decisions that are rational for the group as a whole as well as for the individuals that compose it.

V. COORDINATION

As discussed at the outset, one of the important properties of a cooperative multiagent system is the necessity of its members to coordinate their choices. If we are dealing with categorical preferences, extrinsic coordination can occur if agents happen to have highly compatible preferences. Intrinsic coordination, however, arises because of the social interaction of the members of the group. If social relationships exist, the corresponding conditional utilities will engender an intrinsic aptitude, or capacity, of the group to function as an interrelated whole. Although this endogenous coordination capacity may give rise to harmonious behavior as perceived by an outside observer (or even by the agents themselves), the presence of such behavior itself is not a measure of an intrinsic ability or capacity to coordinate. To measure the overall coordination capacity of a group, we must restrict our attention to an analysis of the utilities that define the system. Furthermore, this coordination capacity must be defined independently of the rationality concepts possessed by the agents.

Fortunately, since the social welfare function is a multivariate mass function and the individual welfare functions are marginals, we may proceed by once again appropriating a concept from probability theory, namely, Shannon's information theory. Our approach is to compute the mutual information associated with a multiagent system and use that quantity to define a notion of coordinatability.

Definition 7 The *entropy*[1] of a multiagent system $\boldsymbol{\mathcal{X}} = \{X_1, \ldots, X_n\}$ with group welfare function $w_{\boldsymbol{\mathcal{X}}_n}$ is

$$H(X_1, \ldots, X_n) = -\sum_{\mathbf{a}} w_{\boldsymbol{\mathcal{X}}}(\mathbf{a}) \log_2 w_{\boldsymbol{\mathcal{X}}}(\mathbf{a}). \tag{4}$$

The entropy of an individual agent X_i is

$$H(X_i) = -\sum_{a_i} w_{X_i}(a_i) \log_2 w_{X_i}(a_i). \tag{5}$$

□

Entropy is an essential concept in the development of Shannon's information theory, and is a measure of the degree of randomness associated with a phenomenon [23]. Essentially, entropy is a measure of the intrinsic potential or capacity to make a wrong guess. If most of the probability mass is concentrated on a single outcome, then the capacity to guess wrong is small (low entropy), but if the probability mass is more or less evenly distributed among the outcomes, then the capacity to make a wrong guess is large (high entropy).

Justification of the use of entropy in a multiagent systems theory setting requires some further explanation. Analogous to the interpretation of entropy as a measure of uncertainty in decision making, it can be interpreted as a measure of the degree of difficulty in making choices. One way to evaluate the difficulty of choosing an action is to consider the opportunity cost, which is traditionally defined as the utility of the next-best action. If most of the utility mass is concentrated on a single action, then the opportunity cost is small (low entropy). On the other hand, if the utility mass is evenly distributed among the actions, then the opportunity cost is large (high entropy).

Definition 8 The *mutual information* associated with a multiagent system is

$$I(X_1; \ldots; X_n) = \sum_{i=1}^{n} H(X_i) - H(X_1, \ldots, X_n) \tag{6}$$

□

It has been shown by [24, 25] that entropy and mutual information used together provide, in the case $n = 2$, a true metric.

Theorem 1: The function

$$d(X_1, X_2) = H(X_1, X_2) - I(X_1; X_2)$$

possesses the following properties:

$d(X_1, X_2) \geq 0$ and $d(X_1, X_2) = 0$ if and only if $X_1 = X_2$

$d(X_1, X_2) = d(X_2, X_1)$ (symmetry)

$d(X_1, X_2) \leq d(X_1, X_3) + d(X_3, X_2)$ (triangle inequality)

where we say that $X_1 = X_2$ if there is a one-to-one function mapping X_1 to X_2; that is, X_1 takes action $a_1 \in \mathcal{A}_1$ if and only if X_2 takes action $a_2 \in \mathcal{A}_2$.

[1]For a detailed discussion of this concept, see [22].

This distance measure is maximum when X_1 and X_2 are independent, in which case $d(X_1, X_2) = H(X_1) + H(X_2)$. As pointed out in [24, 25], a more useful metric is obtained by dividing $d(X_1, X_2)$ by the joint entropy, yielding

$$\mathcal{D}(X_1, X_2) = \frac{d(X_1, X_2)}{H(X_1, X_2)}. \tag{7}$$

This function measures relative distance between the group welfare function and the product of the individual welfare functions. It is clearly symmetric and non-negative by inspection. The proof that $\mathcal{D}$ satisfies the triangle inequality is also provided in [24, 25].

When $n \geq 3$, we may expand the definition of $d(X_1, \ldots, X_n)$ to become

$$d(X_1, \ldots, X_n) = nH(X_1, ..., X_n) - \sum_{i=1}^{n} H(X_i). \tag{8}$$

This function is symmetric, non-negative, and is zero if and only if $X_1 = \cdots = X_n$. We interpret $d(X_1, \ldots, X_n)$ as a *dispersion measure*; that is, a measure of how much the utilities of the members of the group are in conflict. The smaller the dispersion measure, the more tightly the preferences of the agents are clustered. $d(X_1, \ldots, X_n)$ achieves its maximum when the preferences all of the X_i's are mutually independent of each other, in which case $d(X_1, \ldots, X_n) = (n-1)\sum_{i=1}^{n} H(X_i)$.

Definition 9 The *relative dispersion measure* of a multiagent system $\{X_1, \ldots, X_n\}$ is

$$\mathcal{D}(X_1, \ldots, X_n) = \frac{1}{n-1} \frac{d((X_1, \ldots, X_n)}{H(X_1, \ldots, X_n)}. \tag{9}$$

□

Although not a metric when $n \geq 3$, $\mathcal{D}(X_1, \ldots, X_n)$ is symmetric, non-negative, and $D(X, \ldots, X) = 0$, and assumes a maximum value of unity when the agents are mutually independent.

Definition 10 The *coordinatability index* of a group $\{X_1, \ldots, X_n\}$ is the difference between unity and the relative dispersion measure.

$$C(X_1, \ldots, X_n) = 1 - \mathcal{D}(X_1, \ldots, X_n). \tag{10}$$

□

The coordinatability index measures the degree to which the preferences of $X_1, \ldots, X_n$ coincide. If $X_1 = \cdots = X_n$, then $C(X_1, \ldots, X_n) = 1$, the ability to coordination is maximized, and the optimal behavior is for all agents to help themselves by helping others. At the other extreme, if the preferences of the agents are do not influence each other, then $C(X_1, \ldots, X_n) = 0$, the ability to coordinate is minimized, and the optimal behavior is for each agent to do what is best for itself, regardless of the effect on others, which of course might lead to the breakdown of group action. This latter case is the operative condition for classical game theory, where all utilities are categorical. We hasten to add that a lack of an intrinsic capacity to coordinate does not mean that the agents cannot achieve a harmonious solution (recall that the ability to coordinate is not based on a rationality concept). Rather, it means, if they do behave harmoniously, it is by coincidence, based on the juxtaposition of their utilities, and not because of an explicit social relationship between them.

VI. CONCLUSION

Conditional game theory provides a framework within which to model social relationships explicitly via conditional utilities, thus permitting a notion of group, as well as individual, benefit to be defined. Using this structure, it is possible to define a metric to compute the intrinsic ability of a group to coordinate. Coordinatability, however, is a structural attribute of the group, rather than a performance attribute, and provides a quantifiable measure of the intrinsic difficulty in making joint decisions, as well as a measure of the intrinsic ability for the players to coordinate their choices. Consequently, coordinatability may be viewed as meta knowledge that will help the decision makers more completely to understand just how fit, or adapted, they are to address the decision problem they are facing.

REFERENCES

[1] J. A. H. Murray, H. Bradley, W. A. Craigie, and C. T. Onions, Eds., *The Compact Oxford English Dictionary*. Oxford: The Oxford Univ. Press, 1991.

[2] J. S. Rosenschein, "Rational interaction: Cooperation among intelligent agents," Ph.D. dissertation, Stanford University, 1985.

[3] M. R. Genesereth, M. L. Ginsberg, and J. S. Rosenschein, "Cooperation without communication," in *Proceedings of the Fifth National Conference on Artificial Intelligence (AAAI-86)*, 1986, pp. 51–57.

[4] G. Weiss, Ed., *Multiagent Systems*. Cambridge, MA: MIT Press, 1999.

[5] S. Parsons and M. Wooldridge, "Game theory and decision theory in multi-agent systems," *Autonomous Agents and Multi-Agent Systems*, vol. 5, pp. 243–254, 2002.

[6] S. J. Russell and P. Norvig, *Artificial intelligence: A Modern Approach*, 2nd ed. Upper Saddle River, N.J.: Prentice-Hall, 2003.

[7] Y. Shoham and K. Leyton-Brown, *Multiagent Systems*. Cambridge, UK: Cambridge University Press, 2009.

[8] J. R. Marden, G. Arslan, and J. S. Shamma, "Cooperative control and potential games," *IEEE Trans. Systems, Man, Cybernet.*, vol. 39, no. 6, pp. 1393 – 1407, December 2009.

[9] R. D. Luce and H. Raiffa, *Games and Decisions*. New York: John Wiley, 1957.

[10] M. Shubik, *Game Theory in the Social Sciences*. Cambridge, MA: MIT Press, 1982.

[11] T. C. Schelling, *The Strategy of Conflict*. Oxford: Oxford University Press, 1960.

[12] D. K. Lewis, *Convention*. Cambridge, MA: Harvard University Press, 1969.

[13] C. Bicchieri, *Rationality and Coordination*. Cambridge: Cambridge University Press, 1993.

[14] T. W. Malone and K. G. Crowston, "Toward an interdisciplinary theory of coordination," 1991, technical Report No. 120. Center for Coordination Science, SS WP#3294-91-MSA, MIT.

[15] ——, "The interdisciplinary study of coordination," in *Organizing Business Knowledge*, T. W. Malone, K. G. Crowston, and G. A. Herman, Eds. Cambridge, MA: MIT Press, 2003.

[16] R. W. Cooper, *Coordination Games*. Cambridge, UK: Cambridge University Press, 1999.

[17] S. Goyal, *Connections*. Princeton, NJ: Princeton University Press, 2007.

[18] M. Friedman, *Price theory*. Chicago, Il: Aldine Press, 1961.

[19] F. P. Ramsey, "Truth and probability," in *The Foundations of Mathematics and Other Logical Essays*, R. B. Braithwaite, Ed. New York, NY: The Humanities Press, 1950.

[20] B. de Finetti, "La prévision: ses lois logiques, ses sources subjectives," *Annales de l'Institut Henri Poincaré*, vol. 7, pp. 1–68, 1937, translated as 'Forsight. Its Logical Laws, Its Subjective Sources', in *Studies in Subjective Probability*, H. E. Kyburg Jr. and H. E. Smokler (eds.), Wiley, New York, NY, 1964, pages 93–158.

[21] J. Kemeny, "Fair bets and inductive probabilities," *Journal of Symbolic Logic*, vol. 20, no. 1, pp. 263–273, 1955.

[22] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley, 1991.

[23] C. Shannon and W. Weaver, *The Mathematical Study of Communication*. Urbana, Il: University of Illinois Press, 1949.

[24] M. Li, J. H. Badger, X. Chen, S. Kwong, P. Kearney, and H. Zang, "An information-based sequence distance and its application to whole mitochondrial genome phylogeny," *Bioinformatics*, vol. 17, no. 2, pp. 149–154, 2001.

[25] A. Kraskov, H. Stöbauer, R. G. Andrzejak, and P. Grassberger, "Hierarchical clustering based on mutual information," 2003, *ArXiv q-bio/0311039 (http://arxiv.org/abs/q-bio/0311039)*.