Toward a cognitive game theory

Ivaylo Vlaev[a] and Nick Chater[b]

[a]Department of Experimental Psychology, University of Oxford, Oxford, OX1 3UD, United Kingdom; [b]Institute for Applied Cognitive Science, Department of Psychology, University of Warwick, Coventry, CV4 7AL, United Kingdom.
ivaylo.vlaev@psy.ox.ac.uk   nick.chater@warwick.ac.uk

Abstract: We argue that solving the heterogeneous problems arising from the standard game theory requires looking both at reasoning heuristics, as in Colman's analysis, and at how people represent games and the quantities that define them.

Colman's elegant and persuasive article describes psychological game theory by introducing formal principles of reasoning, and focuses on several nonstandard reasoning processes (team reasoning, Stackelberg reasoning, and epistemic and nonmonotonic reasoning). The goal is to explain psychological phenomena in game-playing that orthodox game theory, and its conventional extensions, cannot explain. We argue that, in addition, a model is needed of how the economic agent perceives and mentally represents the game initially, before any (strategic) reasoning begins. For instance, the perceived utility of various outcomes might change depending on the previous games seen.

As an illustration of such a possible model, here we offer some initial results from a research program that aims to ground ac-

counts of rationality in general, and decision theory in particular, on the underlying cognitive mechanisms that produce the seemingly paradoxical behavior. Colman's sections 2 and 4 discuss the basic underlying assumptions of expected utility theory and game theory. Existing models of rational choice and interactive game-theoretic decision making typically assume that only the attributes of the game need be considered when reaching a decision; that is, these theories assume that the utility of a risky prospect or strategy is determined by the utility of the outcomes of the game, and transforms the probabilities of each outcome. Decisions are assumed to be based on these utilities.

Our results demonstrate, however, that the attributes of the previously seen prospects and games influence the decisions in the current prospect and game, which suggests that prospects and games are not considered independently of the previously played ones (Stewart et al., in press; Vlaev & Chater, submitted). In particular, Stewart et al. (in press) have revealed the phenomena of "prospect relativity": that the perceived value of a risky prospect (e.g., "*p* chance of *x*") is relative to other prospects with which it is presented. This is counter to utility theory, according to which the perceived value of each prospect should be dependent only on its own attributes. Stewart et al. suggest that this phenomenon arises in the representation of the magnitudes that define the prospects, and suggest that the phenomenon has a common origin with related effects in the perception of sensory magnitudes (Garner 1954; Laming 1997; Lockhead 1995).

We have found similar effects, providing a new type of anomaly for orthodox game theory. People play repeated one-shot Prisoner's Dilemma games (Vlaev & Chater, submitted). The degree to which people cooperate in these games is well-predicted by a function of the pay-offs in the game, the cooperation index as proposed in Rapoport & Chammah (1965). Participants were asked on each round of the game to predict the likelihood that their coplayer will cooperate, and then to make a decision as to whether to cooperate or defect. The results demonstrated that the average cooperation rate and the mean predicted cooperation of the coplayer in each game strongly depend on the cooperativeness of the preceding games, and specifically on how far the current game was from the end-points of the range of values of the cooperation index in each session. Thus, in games with identical cooperation indices, people cooperated more and expected more cooperation in the game with higher rank position relative to the other cooperation index values in the sequence. These findings present another challenge to the standard rational choice theory and game theory, as descriptive theories of decision-making under uncertainty, and also to other theories where games are independently considered.

Our proposed account for these results, and also for other problems related to the independence assumption, is that people have poor notions of absolute cooperativeness, risk, and utility, and instead make their judgments and decisions in relative terms, as is described in some existing psychophysical and cognitive theories of perception and judgment of information about magnitudes (intensities of stimulus attributes). Thus, this account departs fundamentally from previous work in this field, by modeling the highly flexible and contextually variable way in which people represent magnitudes, like sums of money, probabilities, time intervals, cooperativeness, and so forth, rather than by assuming that these can be represented on absolute internal psychological scales (i.e., even if these scales exist, they stretch or contract depending on the other stimuli in the environment). We conjecture that the results from the two studies presented here suggest that people use context as a sole determinant of the utility of a strategy, which is a form of a more ecologically adaptive rationality, and therefore any descriptive account of game-theoretic behavior, especially in sequential social dilemmas, should incorporate a model of agents' lower-level cognitive perceptual processes.

This discussion does not answer the paradoxes posed in the target article, but here we would like to make the stronger claim that there are many more phenomena that the standard approach cannot explain (and there will be more to be discovered), and that in order to develop a decent account of human decision behavior in games, a much more radical approach is needed. Our results imply that Colman's version of psychological game theory, as based only on nonstandard forms of reasoning, needs to be supplemented by a more general "cognitive game theory," which grounds decision-making in the underlying cognitive mechanisms that produce the decision behavior. Such a cognitive approach could also include collective rationality criteria (which, as Colman states, are lacking in the standard decision theory), because, for example, categorization of the coplayer as being very similar to me could strongly affect my common belief in each other's rationality, or at least in the similarity of the reasoning processes that we would employ. Also, the perception of the players and the game as being similar to a previous interactive situation, in which the coplayers acted in a certain way (e.g., chose a certain focal point), would enforce the belief that, in the current situation, the coplayers would act in a similar way.