Research Focus

Depth of strategic reasoning in games

Andrew M. Colman

School of Psychology, University of Leicester, University Road, Leicester LE1 7RH, UK

Orthodox game theory assumes indefinitely recursive reasoning ('I think that you think that I think...'), but human decision-makers, who are limited by bounded rationality, cannot handle limitless layers of complexity. Recent research corroborates earlier findings that human players tend to operate at only one or two levels of strategic depth.

In conventional decision theory, it is assumed that rational agents invariably choose options that maximize their payoffs relative to their beliefs. For interactive decisions, further assumptions are required about what players expect their co-players to do. Game theory therefore incorporates 'common knowledge' assumptions [1,2], based on players' models of their co-players: every player knows everything about the game, knows that every player is rational, knows that every player knows all this, knows that every player knows that every player knows it, and so on. This implies a form of indefinitely iterated recursive reasoning ('I think that you think that I think...') that was discussed by Keynes [3] in a famous passage likening stock market investment to newspaper competitions to choose the prettiest face from an array, the prize going to competitors whose choice matches the most popular choice: 'It is not a case of choosing those which, to the best of one's judgement, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practise the fourth, fifth and higher degrees.' ([3], p. 156).

Cognitive limitations

In orthodox game theory, players are assumed to pursue recursive reasoning indefinitely and to attribute the same strategic depth to their co-players. But human decision makers have bounded rationality, and working memory is limited in capacity, probably to between four and seven chunks of information [4–6].

Research into cognitive processing of recursively embedded sentences has shown that three or more levels of embedding generate errors of comprehension and recall [7,8]. A typical four-level embedded sentence is:

'The movie (that the script (that the novel (that the producer (whom she thanked) discovered) became) was made into) was applauded by the critics.'

Embedded sentences can be rearranged as right-branching sentences:

'She thanked the producer who discovered the novel that became the script that was made into the movie that was applauded by the critics.'

But in complicated cases they remain difficult to understand and remember, and the same problem arises with strategic reasoning under the common knowledge assumption.

Structure of mental models

Hedden and Zhang recently reported two ingenious experiments designed to investigate recursive reasoning in sequential dyadic games [9]. Players made one-off decisions in 32 games such as that illustrated schematically in Fig. 1. Participants, who were assigned the role of Player I, first predicted Player II's choice at the second decision node (if reached), then indicated their own opening move.

Player I's prediction was used to diagnose the theory-of-mind (TOM) reasoning (zeroth-order or first-order) being attributed to Player II, and by implication whether Player I was using first-order or second-order TOM reasoning, respectively. According to Hedden and Zhang, zeroth-order reasoning leads to myopic choices that maximize the player's payoff with 'no understanding of the desires, beliefs, or thoughts of others' ([9], p. 4). First-order reasoning maximizes against zeroth-order reasoning by the co-player. Second-order reasoning maximizes against first-order reasoning by the co-player. In Fig. 1, for example, a second-order Player I predicts that the assumedly first-order Player II will *Stay* at the second node, expecting a payoff of 3 rather than 2, because of an expectation that Player I will reply to *Switch* with *Switch* (to receive 4 rather than 2); hence, at the first node, Player I chooses *Stay*, expecting 3 rather than 1. A first-order Player I, on the other hand, predicts that the assumedly

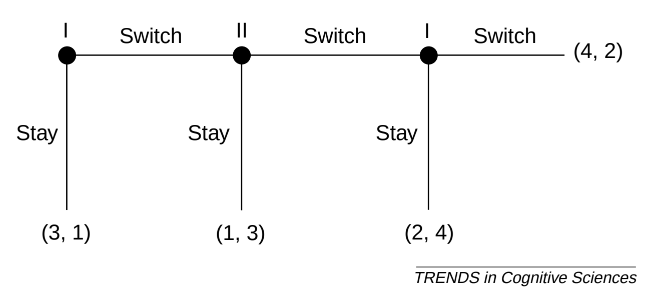

Fig. 1. In this sequential game, Players I and II alternate in choosing either *Stay* or *Switch*, starting at the left. A choice of *Stay* terminates the game immediately, with payoffs to Players I and II shown in parentheses in that order at the bottom, and after successive choices of *Switch–Switch–Switch*, the game terminates with payoffs (4, 2) shown on the right.

Corresponding author: Andrew M. Colman (amc@le.ac.uk).

zeroth-order Player II will ignore Player I's point of view and *Switch*, myopically hoping for 4 rather than 3 after Player I's reply; hence, at the first node, Player I chooses *Switch*, expecting 4 rather than 3.

Hedden and Zhang's results suggested that players generally began with first-order reasoning. When pitted against first-order co-players, some began to use second-order reasoning, but even in the final block of four trials (out of 32), about 30% (in their Experiment 1) or 40% (in Experiment 2) still manifested first-order reasoning. Against co-players who switched from zeroth-order to first-order reasoning at the halfway point, a modest but significant tendency was found, in both experiments, for players to switch from first-order to second-order predicting and choosing, and vice versa.

Methodological problems

There are a number of methodological problems with Hedden and Zhang's experiments. The most obvious is the absence of monetary incentives. Players were simply told that 'the goal should be to earn as many points as possible in each game without regard for the number of points earned by the opponent' ([9], p. 12). Without tangible incentives, extraneous arguments in players' utility functions tend to influence their choices. For example, some players try to 'beat the opponent' by maximizing *relative* (own minus co-player's) points, and others strive for *equality* of points [10,11].

Second, Hedden and Zhang's operational definition of zeroth-order TOM reasoning is slightly obscure. In Fig. 1, every Player I who predicted that Player II would *Switch* at the second decision node was classified as attributing zeroth-order reasoning to Player II. But why should a Player II with 'no understanding of the desires, beliefs, or thoughts of others' choose *Switch*? Choosing *Stay* guarantees Player II a payoff of 3, whereas *Switch* yields either 4 or 2, depending on Player I's reply, and a zeroth-order reasoner does not consider Player I's reply. Move selection under zeroth-order reasoning seems imponderable. The operational diagnosis of Player I's first-order reasoning therefore rests on a debatable interpretation of Player II's zeroth-order reasoning. Other researchers have offered a more determinate characterization of zeroth-order reasoning, as we will see.

Third, Hedden and Zhang 'prompted' players before each choice to 'make a prediction of the opponent's choice' ([9], p. 12). Research has shown that players tend to choose with greater strategic sophistication after being prompted to predict co-players' choices [12,13]. Belief prompting is a form of facilitative priming, and it would not be unreasonable to expect shallower strategic reasoning in its absence.

Stahl–Wilson approach

In spite of these methodological problems, Hedden and Zhang's results broadly corroborate those of earlier experiments, not cited in their article. Most importantly, Stahl and Wilson [14] distinguished between '*level-0*' types, who choose randomly with uniform probability choice functions; '*level-1*' types, who believe that their co-players are level-0 types; '*level-2*' types, who believe that their co-players are either level-0 or level-1 types; '*naive Nash*' types, who believe that their co-players will choose game-theoretic equilibrium strategies; and '*worldly*' types, who constitute a residual category.

Stahl and Wilson performed an experiment with large monetary incentives and no belief prompting [14]. They used twelve simultaneous-choice 3 × 3 games, some of which could be solved by iterated elimination of dominated strategies, that is, by taking each player's point of view in turn and eliminating any strategies yielding worse payoffs for that player irrespective of the co-player's choice. Most of the players avoided their own dominated strategies but did not do so iteratively: that is, they did not incorporate avoidance of dominated strategies into their models of their co-players. Ignoring the residual 'worldly' category, mathematical modelling revealed that 28% were level-0 types, 34% level-1 types, 4% level-2 types, and 34% naive Nash types.

Related findings and conclusions

It is worth mentioning briefly some other related findings. Experimental research into Stackelberg reasoning, in which players in simultaneous-choice games reason as if their choices could be anticipated by their co-players, has confirmed that many players easily manage first-order strategic reasoning [15]. Experimental 'beauty contest' games, named after the passage from Keynes quoted near the beginning, have confirmed that most players are limited to first-order or second-order reasoning [16]. This corroborates the findings, mentioned earlier, on cognitive processing of recursively embedded sentences [7,8]. All these disparate findings, together with those of Stahl and Wilson [14] and most recently Hedden and Zhang [9], converge on the conclusion that boundedly rational human agents generally operate at first-order or second-order depth of strategic reasoning. Future research will reveal whether there are any special circumstances in which deeper or shallower levels are commonly used.

References

1 Lewis, D.K. (1969) *Convention: A Philosophical Study*, Harvard University Press

2 Aumann, R.J. (1976) Agreeing to disagree. *Ann. Stat.* 4, 1236–1239

3 Keynes, J.M. (1936) *The General Theory of Employment Interest and Money*, Macmillan

4 Baddeley, A. (1986) *Working Memory*, Oxford University Press

5 Miller, G.A. (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63, 81–97

6 Simon, H.A. (1974) How big is a chunk? *Science* 183, 482–488

7 Miller, G. and Isard, S. (1964) Free recall of self-embedded English sentences. *Inform. Control.* 7, 293–303

8 Christiansen, M.H. and Chater, N. (1999) Toward a connectionist model of recursion in human linguistic performance. *Cogn. Sci.* 23, 157–205

9 Hedden, T. and Zhang, J. (2002) What do you think I think you think? Strategic reasoning in matrix games. *Cognition* 85, 1–36

10 Van Lange, P.A.M. (1999) The pursuit of joint outcomes and equality in outcomes: an integrative model of social value orientation. *J. Pers. Soc. Psychol.* 77, 337–349

11 Samuelson, C.D. and Allison, S.T. (1994) Cognitive factors affecting the use of social decision heuristics in resource-sharing tasks. *Organ. Behav. Hum. Dec.* 58, 1–27

12 Croson, R.T.A. (2000) Thinking like a game theorist: factors affecting the frequency of equilibrium play. *J. Econ. Behav. Organ.* 41, 299–314

13 Schotter, A. Weigelt, K. and Wilson, C. (1994) A laboratory

investigation of multiperson rationality and presentation effects. *Game Econ. Beh.* 6, 445–468

14 Stahl, D.O. and Wilson, P.W. (1995) On players' models of other players: theory and experimental evidence. *Game Econ. Behav.* 10, 218–254

15 Colman, A.M. and Stirk, J.A. (1998) Stackelberg reasoning in mixed-motive games: an experimental investigation. *J. Econ. Psychol.* 19, 279–293

16 Nagel, R. (1995) Unraveling in guessing games: an experimental study. *Am. Econ. Rev.* 85, 1313–1326

PII: S1364-6613(02)00006-2

Research Focus Response

Two paradigms for depth of strategic reasoning in games

Response to Colman

Jun Zhang and Trey Hedden

Department of Psychology, University of Michigan, Ann Arbor, MI 48109, USA

We contrast two approaches to probing the depth of theory-of-mind (ToM) reasoning by adults in matrix game settings: our own and that of Stahl and Wilson.

As Colman points out [1], although recursive reasoning in games is an essential element of normative game theory, empirical investigations of the sophistication of human players demonstrate that relatively shallow orders of reasoning are more usually used. Strategic reasoning in matrix games in adults using theory-of-mind (ToM) has been probed by Stahl and Wilson [2] and more recently by us [3]. Important differences in these two approaches hold implications for a descriptive theory of bounded rationality.

The Stahl–Wilson paradigm

Stahl and Wilson [2] asked participants to play a series of games against a group of opponents. No feedback was given until all games were completed, but players were allowed to reconsider choices made for previous games. In some cases, inferences could be made based on an analysis of dominance relationships between strategies. A player's depth of reasoning was determined by the extent to which dominated strategies were iteratively eliminated through perspective-taking. Research probing ToM reasoning through dominance relationships of payoffs in games was previously conducted by Perner [4].

This approach has the advantages that: (1) because no belief prompting was given, the depth of reasoning achieved by participants was spontaneous; and (2) all orders of reasoning were operationally defined, although an individual player's strategy was determinable only through statistical modeling.

However, it is not clear that the Stahl–Wilson paradigm probes the order of reasoning players are fully capable of engaging in, as opposed to beliefs about how their opponents will behave. The strategy used by a player will depend upon an estimate of the distribution of strategies used by all other players. As a consequence, some second-order players might revert to lower-order reasoning because of uncertainty about the sophistication of their opponents. It remains to be seen whether a manipulation of background information about opponents' level of intelligence, as performed in Hedden and Zhang [3], would change the distribution of players of various orders.

Although reconsideration of choices masked any learning or sequential effects, it is debatable whether players would exhibit a uniform strategy to Nash-solvable and to dominance-solvable games, because their strategic analyses differ. This is further complicated by the fact that these two kinds of games may well relate to depth of ToM in different and multiple ways. For example, *defect–defect* is the solution to the standard Prisoner's Dilemma game, in the sense of dominance (in addition to the sense of Nash) that requires only zeroth- and first-order consideration, but mutual cooperation becomes a Nash equilibrium under second-level meta-game analysis involving second- and third-order ToM reasoning between the two players [5].

The Hedden–Zhang paradigm

In contrast to Stahl and Wilson, participants in Hedden and Zhang [3] were pitted against a single opponent (a confederate programmed to use either a zeroth- or first-order strategy, or to switch strategies between two consecutive blocks). Hence, players were modeling a particular individual as opposed to estimating the distribution of strategic sophistication in a population. Further, players received feedback on each trial, allowing them to determine the opponent's strategy and to update their ToM model on-line. The availability of reaction-time measures provided corroborative evidence about participants' engagement in shallow and deep reasoning in individual games. This paradigm therefore investigates dynamic changes in the depth of reasoning resulting from progressive interaction, as well as providing an estimate of participants' initial default strategy. For most players, this default was based on first-order reasoning.

Colman [1] notes that drawbacks of this paradigm include that it: (1) relies on a particular operational definition of a zeroth-order opponent's behavior; and (2) involves belief prompting. Both are legitimate concerns.[1] However, (1) would underestimate the probability of

---

Corresponding author: Jun Zhang (junz@umich.edu).

[1] Colman's concern about the motivation of participants when no monetary incentive is involved falls more in the line of standard practices of economists and psychologists. Points earned by opponents were not tallied and displayed in our experiments [3], and care was taken to eliminate participants based on their performance on catch trials incorporated into the experimental design.