Games and Economic Behavior 87 (2014) 204–223

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Learning with bounded memory in games ☆

Daniel Monte

Sao Paulo School of Economics - FGV, Brazil

ARTICLE INFO

Article history:
Received 18 June 2010
Available online 27 March 2014

JEL classification:
C72
D82
D83

Keywords:
Bounded memory
Incomplete information games
Repeated games

ABSTRACT

We study learning with bounded memory in zero-sum repeated games with one-sided incomplete information. The uninformed player has only a fixed number of memory states available. His strategy is to choose a transition rule from state to state, and an action rule, which is a map from each memory state to the set of actions. We show that the equilibrium transition rule involves randomization only in the intermediate memory states. Such randomization, or less frequent updating, is interpreted as a way of testing the opponent, which generates inertia in the player's behavior and is the main short-run bias in information processing exhibited by the bounded memory player.

1. Introduction

In this paper we assume that people categorize the world in a coarse way. This is consistent, for example, with the fact that consumer reports often come in the form of finite ratings. It is also consistent with the view that it may not be possible to distinguish beliefs about two agents or products that differ only by a negligible fraction (in a particular dimension). In fact this is also how some authors in psychology model working memory.[1] To illustrate, assume that an agent thinks of his opponent as being someone "trustworthy", "not trustworthy" or "unclear", instead of having a precise posterior distribution about the opponent's true underlying type.

In our model, bounded memory captures this coarse categorization. Apart from bounded memory, the agent is rational, and updating from one category to another is part of her strategy. A bounded memory player has only a fixed number of memory states available. All she knows about the history of the game is her current memory state. The player is aware of her memory constraints, and her strategy is to choose a transition rule from state to state and an action rule, which is a map from each memory state to the set of actions. We show how this agent updates her beliefs and the implications of this chosen updating rule in a strategic setting. In particular, we show that the equilibrium updating rule in a reputation game involves randomization at the intermediate memory states (as opposed to what has been found in decision problems), and we argue that this inertia is a short-run bias in information processing in the behavior of the bounded memory player.

Given that the player is only constrained in her memory, we view memory as a "conscious" process, i.e. the player is subject to "incentive compatibility constraints" (sequential rationality constraints extended to games with bounded memory).

☆ This paper is a substantially revised version of my Ph.D. dissertation at Yale University. I am very thankful to Ben Polak, Dirk Bergemann, and Stephen Morris for their constant guidance. Many people have helped me throughout the process and I am very thankful to them. I am particularly thankful for conversations with Attila Ambrus, Faruk Gul, Pei-Yu Lo, George Mailath, Tymofiy Mylovanov, and Arthur Robson.

E-mail address: daniel.monte@fgv.br.

[1] See for example Miller (1956), and Cowan (2001).

http://dx.doi.org/10.1016/j.geb.2014.03.005
0899-8256/

The action that the player chooses at each memory state and the transition from state to state must be optimal given her beliefs at that state, and taking as given her *own* strategy – both action and transition rules – at all other states. In contrast to sequential rationality in models with perfect recall, incentive compatibility constraints here do not imply an optimal continuation strategy. The reason a bounded-memory player must take her own strategy at all periods (including future periods) as given when deciding on an action or on which state to move is that if she deviates today, she will not remember it tomorrow.

Conscious memory distinguishes our model from other models of bounded rationality in the literature, in particular, from the standard finite automata.[2] Like ours, a standard automaton has a fixed set of states, a transition rule, and an action rule. However, a standard automaton is assumed to fully (and credibly) commit to a strategy at an ex-ante stage, and, hence, it does not face incentive compatibility constraints.[3] The idea of incentive compatibility constraints, or sequential rationality in bounded memory, was introduced by Piccione and Rubinstein (1997) and Wilson (2003), but these authors studied single-person decision problems for which commitment ex-ante has been shown to be ineffective. Here, we study games, in which the inability to commit matters.

To be able to isolate the effects of bounded memory from complexity considerations, we focus on zero-sum games. An infinitely repeated zero-sum game with complete information requires minimal complexity. We look at the incomplete information case in which the probability of a commitment type is sufficiently low, in the spirit of Kreps et al. (1982).

The precise setting of this study is an infinitely repeated two player game with incomplete information. One player is uncertain about his opponent's true type, which we assume to be either a commitment type that is restricted to only one action or a normal type that has opposite preferences to those of the uninformed player. Our main propositions apply to a general class of zero-sum games, but memory restrictions will only bind in games in which learning the true type of the opponent is profitable for the uninformed player. Thus, we focus on games in which the equilibrium payoff of the uninformed player in the complete information game against the commitment type is strictly higher than her equilibrium payoff in the incomplete information game (we want to rule out uninteresting cases, such as when all payoffs are the same). The matching pennies is a canonical example for our stage game. We assume that only the uninformed player has finite memory, whereas the informed player is an unbounded player.

Every period the two agents play a simultaneous finite game. The payoffs are realized and the actions are observed with certainty. Then, the uninformed player updates her beliefs (according to her chosen memory rule) on the opponent's true type.

Equilibrium will always exist and, typically, not be unique. In Propositions 1 and 2, we show conditions that must be satisfied in any equilibrium of the game. In particular, we show that the updating rule must be weakly increasing after observing actions that are consistent with the action a commitment type would play. This implies that the uninformed player's belief on the commitment type is stochastically higher after observing the commitment type's action. In addition, because a different action outside this set leaves no uncertainty in the mind of the uninformed player, she moves to her "lowest" memory state after such an action. We show that, in this "lowest" state, her belief on the informed player being the commitment type is close to zero, and, in her "highest" state, the belief is close to one. This result holds even for the minimal case of only two memory states, or one-bit memory.

In the cases in which the prior belief on the commitment type is sufficiently small, the equilibrium transition rule will require randomization in any equilibrium in which memory states are used in an "efficient" way. Informally, we interpret this result as indicating that when the uninformed player does not have enough memory to keep track of all the actions played by the opponent, she will use randomization to overcome the memory problem and to "test" the opponent before updating.

The role of random transition rules in an optimal finite memory has been studied in single person decision problems. Hellman and Cover (1970) studied the two-hypothesis testing problem with a finite automaton (with ex-ante commitment to the strategy). A decision maker has to make a decision after a very long sequence of signals. However, the decision maker cannot recall all the sequence and, instead, has to choose the best way to store information given his finite set of memory states. A key result of that paper is that, for the discrete signal case, the transition rule is random in the extreme states. The authors concluded that, perhaps counter intuitively, the decision maker uses randomization as a memory-saving device.[4] Kalai and Solan (2003) also study the role and value of simplicity and randomization in a model of dynamic decision making.

Wilson (2003) studied a problem similar to Hellman and Cover (1970). In her model the decision maker was subject to sequential rationality constraints. The optimal memory rule obtained was similar to Hellman and Cover's rule and it included randomization in the extreme states. She showed that this randomization at the extreme states can explain several biases in information processing, such as confirmatory bias and overconfidence/underconfidence bias.

In contrast to the results of Hellman and Cover (1970) and Wilson (2003), we show that in our setting the randomization occurs at the intermediate states and not at the extreme states. Thus, the main behavioral bias exhibited by the bounded memory player in this class of games is infrequent updating in the short-run, which is a result of randomization in the transition rule of memory states associated with intermediate beliefs. We interpret this infrequent updating as excessive inertia and occasional overreaction (when updating does occur).

[2] See, for example, Neyman (1985), Rubinstein (1986), Abreu and Rubinstein (1988) and Kalai and Stanford (1988).

[3] Rubinstein (1986) and Kalai and Neme (1992) also studied automata models with a perfection requirement. The solution concept used in this paper is substantially different, since it requires consistent beliefs, as will be discussed later.

[4] "It is somewhat surprising that randomization is needed at all, since randomization usually decreases information." (Hellman and Cover, 1970, p. 781).