# Finite Automata to Represent Bounded Rationality

Ryan M. Golbeck
Department of Computer Science
University of British Columbia
Vancouver, BC V6T 1Z4, Canada

December 9, 2005

**Abstract**

Game theory makes the assumption that each agent in a game is rational: the agents are expected utility maximizers. This paper explores the use of finite automata to model agents and the restriction of the allowable automata as a way of bounding the rationality of the agents. We analyze several major results for their effectiveness at explaining the experimental results that are inconsistent with theoretical predictions.

## 1 Introduction

Game theory makes the assumption that each agent in a game is rational: the agents are expected utility maximizers. This assumption leads to rationality constructs such as "backward induction," where a rational agent will compute his/her strategy in an extended form game by starting at the leaves of the game tree and working upward. However, backward induction often predicts results which are entirely different than results obtained empirically. This divergence is most often noted in the "Finitely Repeated Prisoners' Dilemma," where theoretical results predict that each agent will defect in every round and experimental results show extended periods of cooperation[1, 2].

Bounded rationality has emerged as a major paradigm for explaining why theoretical and experimental results do not always line up in game theory. We still make the assumption that each agent is an expected utility maximizer, but we attempt to restrict the agents' computational ability so that they may not be able to compute the "most rational" outcome.

This paper looks at modeling agents' actions using finite automata. We then evaluate four ways of restricting the allowable automata as a means of placing bounds on the agents' rationality: imposing a maintenance cost, considering an alternate solution concept, directly limiting the complexity of the automata, and limiting the time agents are allowed to compute their automata.

| | c | d |
|---|---|---|
| C | 2,2 | 0,3 |
| D | 3,0 | 1,1 |

Figure 1: Standard Prisoners' Dilemma Payoff

# 2 Bounded Rationality with Finite Automata

The use of finite automata (FA) has emerged as the most studied model for representing bounded rationality in agents. In this model, each agent's pure strategy space is the space of modified FA called Moore machines. A Moore machine is a FA that labels each state with an output. This output corresponds to the action that that agent plays at his/her next choice node. The transition function maps the current state together with the current information set to the next state in the machine. For iterated stage games, such as iterated prisoners' dilemma, we can simplify the transition function to be a map from the current state and the opposing player's action to the next state.

More formally, in the iterated game $G = (N, A, u)$, agent $i$'s strategy is:

- A set of states, $Q_i$.
- An initial state, $q_i^0 \in Q_i$.
- A action labeling function, $L_i : Q_i \to A_i$, mapping each state to a pure action.
- And a transition function, $\delta_i : Q_i \times A \to Q_i$, mapping a state and an action profile to the next state.

So, an agent's allowable pure strategies are exactly the automata that the agent is permitted to use. Further, each agent's mixed strategy space is a straight-forward extension of mixed strategies in a normal form game: the set of distributions over the set of all Moore machines available to player $i$.

## 2.1 Examples

Consider the standard prisoners' dilemma game with the payoff matrix shown in figure 1.

In the iterated prisoners' dilemma, many well known strategies are easy to represent using FA:

1. The tit-for-tat strategy, in which the agent's first action is to cooperate,

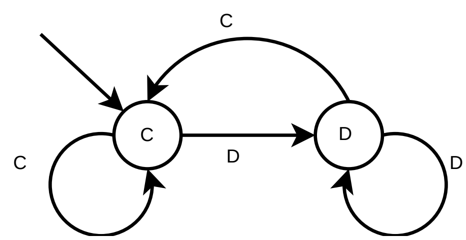


Figure 2: Tit-for-tat Moore Machine for Iterated Prisoners' Dilemma

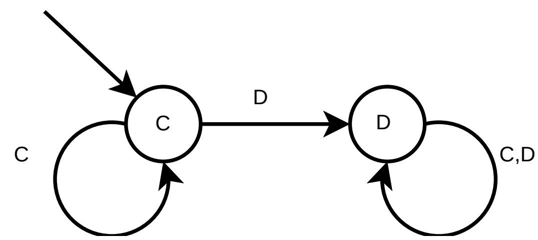


Figure 3: Trigger Strategy Moore Machine for Iterated Prisoners' Dilemma

and every round the agent plays his/her opponent's last move. We get:

$$
\begin{aligned}
&Q_i = \{q_i^c, q_i^d\} \\
&q_i^0 = q_i^c \\
&L(q_i^a) = a \\
&\delta(q_i^a, a') = q_i^{a'}
\end{aligned}
$$

See figure 2 for a graphic representation for this strategy.

2. For the trigger strategy, where the agent always cooperates until the opposing agent defects, we get:

$$
\begin{aligned}
&Q_i = \{q_i^c, q_i^d\} \\
&q_i^0 = q_i^c \\
&L(q_i^a) = a \\
&\delta(q_i^a, d) = q_i^d \\
&\delta(q_i^a, c) = a
\end{aligned}
$$

See figure 3 for a graphical representation of this strategy.

# 3 Results and Analysis

## 3.1 Bounding Rationality by Utility

A straight forward way to limit the complexity of the allowable automata, in infinitely repeated games, is to incorporate its complexity into the utility function of the agent. Rubinstein[5] incorporates the complexity by imposing a cost on each agent for each state of the automata. These payments are infinitesimal compared to the payoffs of the game, so an agent will always choose a larger machine if that machine will result in an increase in utility.

Incorporating this cost to explain the bounded rationality of agents can be justified in a number of ways. Since a state machine can be used to remember previous games, this cost can be viewed as the cost of maintaining the necessary memory to remember the history, or it could also represent the cost of learning the strategy.

Although, intuitively, imposing additional costs for the complexity of the automata seems to be a valid way to constrain an agent's rationality, doing so eliminates strategy profiles that were previously Nash equilibria. The trigger strategy profile for the iterated prisoners' dilemma is no longer a Nash equilibrium. Since each agent has "punishment" states in his/her FA that are never used, the agent will be better off by removing these states.

The trigger strategy explains some experimental results that show agents cooperating in iterated games. And, unfortunately, by excluding this equilibrium, we are losing a valid explanation for cooperation.

## 3.2 Semi-Perfect Equilibrium

Additionally, with infinitely repeated games, Rubinstein presents a natural extension to the Nash equilibrium solution concept called *semi-perfect-equilibrium* (SPE). This equilibrium is analogous to subgame-perfect equilibrium in that the strategy profile $A = (M_1, M_2, \ldots, M_n)$ is in semi-perfect-equilibrium if for every round $t$ of the iterated game, $A$ is also a Nash equilibrium of the game if it were to start at round $t$.

Following this solution concept, if a strategy profile $A$ is in equilibrium, then every state in each $M_i$ is used infinitely often. If it was not, and the state $q' \in M_i$ was not used after round $t$, then there exists an $M_i'$ that omits state $q'$ so that agent $i$ would be better off with $M_i'$ starting at round $t + 1$.

Forcing every state to be used infinitely often is more restrictive than adding costs for complexity, which requires only that a state be used at least once.

Under SPE, the space of solutions for iterated prisoners' dilemma shrinks to include only the payoff profiles that are of the form

$$(u_1, u_2) = \alpha(3, 0) + (1 - \alpha)(0, 3) > (1, 1)$$

for rational $\alpha$, or where $(u_1, u_2) = (0, 0)$.

So, similarly to imposing a cost for complexity, the trigger strategy is not an equilibrium strategy under SPE. Since punishment states must be used infinitely

| | L | R |
|---|---|---|
| U | 2,3 | 3,5 |
| M | 5,4 | 2,1 |
| D | 1,0 | 3,1 |

Figure 4: Payoff Matrix for Midterm Game

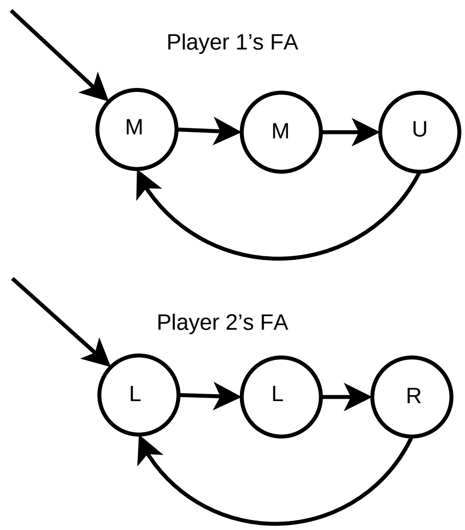


Figure 5: Equilibrium Strategies for Midterm Game

often, they must be incorporated into the normal play of the machine. This restriction, however, does bear some resemblance to real life situations: many institutions, and even human facilities, degrade and are discarded if unused for long periods of time.

Although, it seems difficult to build strategies incorporating punishment phases into the normal course of play, sometimes following a Nash equilibrium without any deviation is enough to punish the other players for deviating.

For example, consider the game with payoffs from figure 4. It can be shown that the strategies given by the pair of machines in figure 5 are in semi-perfect equilibrium.

Unfortunately, since SPE also eliminates the trigger strategy as an equilibrium, we lose an explanation for rational cooperation in iterated stage games.

## 3.3 Bounding Rationality By Complexity

Now, consider bounding the rationality of the agents by explicitly restricting the complexity of their FA. This restriction represents a direct bound on the computation and memory limits of each agent.

In finite repeated prisoners' dilemma played $N$ times, Neyman claims in [3] that for automata with size bounded by $N-1$, there are equilibrium strategy profiles where each agent always cooperates. Further, he makes the claim (which is later proved by Papadimitriou and Yannakakis in [4]) that for any $k$, there exists an $N$ such that if the size of each of the automata are in the range $(N^{1/k}, N^k)$, there is an equilibrium where the payoffs to each player are within $\frac{1}{k}$ of the cooperation payoff. Neymen's result is also improved on in [4] to show that in the iterated prisoners' dilemma, for any $\epsilon > 0$, if at least *one* of the players' automata is bounded by $2^{c_\epsilon N}$, for some constant $c_\epsilon$, then there is an equilibrium with average payoff within $\epsilon$ of the cooperative payoff to each player.

This result is extended to general games, albeit in a weaker form, by Papadimitriou and Yannakakis[4] in the following theorem:

**Theorem 1** (Papadimitriou–Yannakakis)**.** *Let $G$ be an arbitrary game and $p = (p_1, p_2)$ an individually rational Pareto optimal point. For every $\epsilon > 0$ there exists $c_\epsilon > 0$ such that, in the $N$-round repeated game $G$ played by automata with sizes bounded by $2^{c_\epsilon N}$, there is a mixed equilibrium with average payoff within $\epsilon$ of $p_i$ for each player $i$.*

These results show significant promise in modeling the bounded rationality of the agents. Note that if the complexity of the agents' automata are sub-exponential, but still greater than $N$, there are no equilibria for which the agents cooperate in every stage of the game. However, there are equilibria that come arbitrarily close to the cooperation payoff if the game is repeated enough times. This outcome corresponds closely to experimental results where agents achieve long periods of cooperation[1].

## 3.4 Complexity of Computing Best Response

Using FA to model agents' rationality also allows us the possibility of restricting the time that the agents are allowed to use to compute their automata. Bounding the agents in this way leads to a number of questions about the time complexity of computing strategies: what is the complexity of computing the best response to a given strategy?; does a game have a pure equilibrium in FA?; and what is the value of a mixed-equilibrium in a zero-sum game?

We can define a game function to be $g$, a polynomially computable function such that $g(z, x, y) = (a, b)$, where $z$ is an encoded version of the game, $x$ and $y$ are encoded strategies, and $a$ and $b$ are the resultant payoffs. Then, we can use the following theorem [4]

**Theorem 2.** *(Classification Theorem) [Papadimitriou-Yannakakis]*

1. *The class of all languages of the form $\{z; x; b$ : there is a strategy $y$ for player 2 with payoff at least $b\ \}$ is precisely $NP$.*

2. *The class of all languages of the form $\{z$ : there is a pure equilibrium in game $z\}$ is precisely $\sum_2^p$.*

3. *The class of all languages of the form $\{z; b$ : zero-sum game $z$ has an equilibrium with payoff to player 1 at least $b\}$ is precisely $EXP$.*

4. *The class of all languages of the form $\{z; b$ : game $z$ has an equilibrium with payoff at least $b$ to both player 1 and 2 $\}$ is precisely $NEXP$.*

From 1) we have that computing the best response to any given strategy is $NP$-complete, and determining whether there is a pure equilibrium is $\sum_2^p$.

So, it is sufficiently hard to compute best response and equilibrium strategies that we have a plausible explanation that real agents are unable to compute their optimal strategies. Using sub-optimal strategies could produce the results seen in experimentation that do not achieve the "most rational" behaviour.

# 4 Conclusion

Bounded rationality has become an important concept for explaining why game theoretic results do not correspond to experimental data. Finite automata have become a dominate model for expressing the rationality of agents, and we have overview ed four different methods for bounding the agents' rationality by putting restrictions on the FA they are allowed to use.

If we restrict the agents by requiring that they pay a cost per state of the machines they are using, which seems intuitive, we eliminate the tit-for-tat and trigger strategies from the set of Nash equilibria in repeated games. Unfortunately, both of these equilibria result in cooperation among agents; this cooperation is exactly what we are trying to explain with bounded rationality.

Rubinstein attempts to bound the agents by considering a new solution concept: semi-perfect-equilibrium. However, equilibrium strategies under SPE are even more restricted then they are by a cost for state maintenance. SPE results in limiting the solution space payoffs to rational combination of the defect payoffs.

Fortunately, more promising results come from bounding the complexity of the actual machines and the time an agent is allowed to use to compute their strategies. By bounding a FA to be a size sub-exponential in the number of the iterations of the game, we can explain extended periods of cooperation. And, we have seen that computing best response strategies in automata, or even determining if there are pure equilibria, is sufficiently hard for agents that it is plausible agents would arrive at non-rational strategies. Both of these results explain why empirical results show agents playing with long periods of cooperation.

# References

[1] James Andreoni and John H. Miller. Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence. *The Economic Journal*, (103):570–585, 1993.

[2] R. Axelrod. *The Evolution of Cooperation*. Basic Books, New York, 1984.

[3] Neyman. Bounded complexity justifies cooperation in the finitely repeated prisoners' dilemma. *Economics letters*, 19(3):227, 1985.

[4] Christos H. Papadimitriou and Mihalis Yannakakis. On complexity as bounded rationality (extended abstract). In *STOC '94: Proceedings of the twenty-sixth annual ACM symposium on Theory of computing*, pages 726–733, New York, NY, USA, 1994. ACM Press.

[5] Ariel Rubinstein. Finite Automata Play the Repeated Prisores Dilemma. *J. Econ. Theory*, (39):176–188, 1986.