Phase Transition Thresholds when Finite Automata Play Supergames under Finite Costs for Complexity

Omar Muhammad[1] and Amrit Mahendra Joseph[1]

[1]Indian Institute of Science

Abstract

Cooperation in repeated strategic interactions can depend on strategies that remember past behaviour and respond to deviations. When implementing these strategies carries a recurring complexity cost, players may benefit from adopting simpler implementations, potentially undermining cooperation. We study this trade-off in the infinitely repeated Prisoner's Dilemma by modelling strategies as deterministic finite automata. Each player's utility is their long-run average payoff minus a recurring maintenance cost proportional to the number of automaton states. Within this framework, we investigate four critical cost thresholds that characterize the collapse of equilibrium behaviours as maintenance costs increase, from the availability of cooperative outcomes to exclusively mutual defection.

1 Introduction and Motivation

The standard analysis of repeated games often abstracts away the resources required to implement a strategy. Bounded rationality brings these limitations into the analysis: remembering past behaviour and responding to deviations requires memory and computation, whose cost may affect a player's choice of strategy.

We model strategies as Moore machines: deterministic finite automata with an action assigned to each state. Transitions update the state after observing the opponent's action. Charging for each state makes players weigh the benefits of cooperation against the cost of implementing the strategy that sustains it.

In programming languages and formal methods, reactive synthesis studies how to construct a program from a specification of its behaviour over time. The resulting program can be represented by a finite-state machine that repeatedly reads inputs and produces outputs. When the inputs come from another participant with its own objectives, the success of an implementation can depend on that participant's incentives. Rational synthesis addresses this issue by constructing a system together with strategies for the other participants that they have no incentive to abandon (Fisman et al., 2010).

Our setting motivates a resource-aware version of this problem. Suppose the desired specification is that both participants eventually cooperate forever. A pair of finite-state strategies must satisfy this specification, but must also remain an equilibrium when each participant accounts for the cost of its machine. Minimizing implementation size therefore interacts with the incentives that sustain the specified behaviour. Our threshold analysis identifies cost regimes in which such equilibrium implementations can or cannot exist, providing feasibility conditions for this synthesis objective.

We study this trade-off in the infinitely repeated Prisoner's Dilemma through four thresholds: the loss of the full equilibrium payoff set, the loss of eventual mutual cooperation, and the disappearance of non-defection outcomes under Nash and semi-perfect equilibrium.

2 Related Work

Folk theorem results show how future punishment can sustain a wide range of equilibrium payoffs (any feasible and individually rational vector) including cooperation in repeated games when there is no charge for complexity (Friedman, 1971). Aumann (1981) proposed finite automata for modelling strategies, since the infinite memory required for many standard supergame strategies is physically unrealizable and computationally unrealistic. Neyman (1985) showed that restrictions on strategy complexity can support cooperation even in the finitely repeated Prisoner's Dilemma, despite the usual backward-induction argument. This was extended in Megiddo and Wigderson (1986) showing that the sustainability of cooperation persists even when strategies are modelled by size-restricted Turing Machines.

Rubinstein (1986) and Abreu and Rubinstein (1988) study strategy complexity as a cost. Rubinstein (1986) uses lexicographic preferences: players maximize average payoff first and minimize states second. This can destabilize strategies with unused punishment states, which prescribe adverse responses to an opponent's deviation:

a player can obtain the same payoff against the opponent using a smaller machine. However, this does not rule out all Nash equilibria with eventual mutual cooperation. Rubinstein's stronger semi-perfect equilibrium requirement rules out mutual cooperation in the eventual cycle, while permitting mutual defection and certain periodic exploitation cycles. Banks and Sundaram (1990) include transition structure in the complexity measure and obtain stage-game Nash play in every period.

Under a modified evolutionary stability criterion with lexicographic costs, Binmore and Samuelson (1992) require outcomes to maximize the sum of the players' payoffs. Cooper (1996) instead subtracts a linear state cost from time-average payoffs and obtains a folk-theorem-type result: every payoff between mutual defection and mutual cooperation can be approximated by a neutrally stable strategy for sufficiently small positive costs. van Veelen and García (2019) extend this result to discounting and establish a uniform invasion barrier for the one-state defector. These results distinguish the existence of cooperative equilibria from their evolutionary stability. Our focus is on how the existence of Nash and semi-perfect equilibria changes as the per-state cost increases.

3 Model and Equilibrium Concepts

Let $G = \langle S_1, S_2, u_1, u_2 \rangle$ be a two-player normal-form game where S_i and $u_i : S_1 \times S_2 \to \mathbb{R}$ are respectively a finite set of actions and the payoff function for player i. Its supergame repeats G at periods $t = 1, 2, \ldots$, with each player observing the opponent's action after each period. A supergame strategy determines a player's choice of action at period t as a function of the previous $t-1$ outcomes.

We restrict to strategies implemented by Moore machine $M_i = \langle Q_i, q_i^0, \sigma_i, \mu_i \rangle$, where Q_i is a finite state set, $q_i^0 \in Q_i$ is the initial state, $\sigma_i : Q_i \to S_i$ is the output function, and $\mu_i : Q_i \times S_{3-i} \to Q_i$ is the transition function. At period t, player i in state q_i^t plays $s_i^t = \sigma_i(q_i^t)$ and moves to $q_i^{t+1} = \mu_i(q_i^t, s_{3-i}^t)$. Write $|M_i| = |Q_i|$, and let $M_i(q)$ denote the same machine started at q, retaining its full state set. The joint run (q_1^t, q_2^t) is deterministic on the finite product $Q_1 \times Q_2$, hence ultimately periodic. We define player i's supergame payoff $\pi_i(M_1, M_2)$ as their average payoff over the eventual joint cycle, equivalently the long-run average of the generated payoffs.

For a per-state maintenance cost $\lambda \geq 0$, we define utility for a player as $U_i(M_1, M_2) = \pi_i(M_1, M_2) - \lambda |M_i|$.

Definition 1. *A pair (M_1, M_2) is a Nash equilibrium if neither player can increase utility by changing their machine at the start, holding the opponent's machine fixed: $U_1(M_1, M_2) \geq U_1(M_1', M_2)$ and $U_2(M_1, M_2) \geq U_2(M_1, M_2')$ for all finite replacement machines M_1', M_2'.*

Definition 2. *A pair (M_1, M_2) is a semi-perfect equilibrium (SPE) if neither player can increase utility by replacing their machine at any period $t \geq 1$ along the induced play, while the opponent continues from its current state: $U_1(M_1(q_1^t), M_2(q_2^t)) \geq U_1(M_1', M_2(q_2^t))$ and $U_2(M_1(q_1^t), M_2(q_2^t)) \geq U_2(M_1(q_1^t), M_2')$ for all finite replacement machines M_1', M_2'.*

Every SPE is an NE. We specialize to the Prisoner's Dilemma, where each player chooses whether to cooperate (C) or defect (D):

	C	D
C	R, R	S, T
D	T, S	P, P

Each entry lists the row and column players' payoffs. Here R is the reward for mutual cooperation, P the punishment payoff for mutual defection, T the temptation payoff for defecting against a cooperator, and S the sucker's payoff for cooperating against a defector. We assume $T > R > P > S$ and $2R > T + S$: each player individually benefits from defection regardless of the other's action, yet mutual cooperation uniquely maximizes their combined payoff. Let M_C and M_D be the one-state machines that always cooperate and always defect, respectively.

4 Threshold Results

We state the main results with proof sketches. A payoff profile is the pair $v = (\pi_1(M_1, M_2), \pi_2(M_1, M_2))$ of players' game payoffs, before subtracting state costs.

4.1 From free states to positive costs

Proposition 1 (No costs for states)**.** *In the repeated Prisoner's Dilemma with $\lambda = 0$, a payoff profile $v = (v_1, v_2)$ is a Nash equilibrium payoff of the machine game if and only if v is a rational convex combination of $(R, R), (S, T), (T, S), (P, P)$ and $v_i \geq P$ for $i \in \{1, 2\}$. The same characterization holds for semi-perfect equilibrium payoffs.*

Infinitesimal state costs can be shown to be equivalent to lexicographic preferences. Under this setup, unused punishment states cannot be retained merely to deter deviations. Rubinstein's semi-perfect equilibria exclude mutual cooperation from the eventual cycle, while allowing mutual defection and individually rational exploitation cycles (Rubinstein, 1986). Nash equilibrium can still support eventual mutual cooperation (Abreu and Rubinstein, 1988).

Let V^* be the set of payoff profiles characterized in Proposition 1. We ask how much state maintenance can cost before particular equilibrium outcomes become impossible. The following suprema give the upper cost limits for preserving the entire set V^*, sustaining mutual cooperation in NE, and sustaining any payoff other than mutual defection in SPE and NE, respectively:

$$\begin{aligned}
\lambda_1 &= \sup\{\lambda \geq 0 : \text{every } v \in V^* \text{ is an NE payoff}\},\\
\lambda_2 &= \sup\{\lambda \geq 0 : \text{an NE yields } (R,R)\},\\
\lambda_3 &= \sup\{\lambda \geq 0 : \text{an SPE yields } v \neq (P,P)\},\\
\lambda_4 &= \sup\{\lambda \geq 0 : \text{an NE yields } v \neq (P,P)\}.
\end{aligned}$$

Here (R,R) means eventual mutual cooperation; a finite introductory phase does not affect the average payoff. The distinction between NE and SPE matters because a machine chosen at the start may cease to be optimal after its introductory states have been used.

4.2 Best responses and existence of the thresholds

Proposition 2 (Upper bound and existence)**.** *For any state-maintenance cost $\lambda > R - P$, the unique Nash equilibrium of the machine game is (M_D, M_D), yielding (P,P). All four thresholds therefore exist, with*

$$0 \leq \lambda_1 \leq \lambda_2 \leq \lambda_4 \leq R - P, \qquad 0 \leq \lambda_3 \leq \lambda_4.$$

Lemma 1 (Abreu and Rubinstein, 1988, Lemma 1)**.** *Against a fixed opponent M_2, a machine with at most $|M_2|$ states attains the maximum game payoff over all finite-machine responses.*

Lemma 2 (Finite-cost and lexicographic equilibria)**.** *For $\lambda > 0$, every NE has $|M_1| = |M_2|$, and each machine maximizes the game payoff against the other. Consequently, every finite-cost NE is a lexicographic NE, and every finite-cost SPE is a lexicographic SPE.*

If one machine were larger, Lemma 1 would give a smaller response with at least the same payoff. Equal sizes then imply payoff optimality, since a higher payoff is attainable without adding states. An equal payoff with fewer states is also a profitable deviation.

Lemma 3 (One-state machines)**.** *Under strictly positive state costs: (i) a pair of one-state machines is an NE if and only if both are M_D; (ii) any NE with payoff (P,P) uses (M_D, M_D); and (iii) in every NE with payoff different from (P,P), both machines have at least two states.*

To prove Proposition 2, note that in a non-defection NE, both machines need at least two states and $\pi_1 + \pi_2 \leq 2R$. Each player can guarantee $P - \lambda$ with M_D, giving $2P - 2\lambda \leq U_1 + U_2 \leq 2R - 4\lambda$, hence $\lambda \leq R - P$. Proposition 1 makes all four defining sets nonempty; the bounds make their suprema finite.

4.3 The breakdown of the Folk Theorem: λ_1

Lemma 4 (State-cost bounds)**.** *For a profile with payoff v_i and $|M_i| > 1$, a necessary condition for NE is $\lambda \leq (v_i - P)/(|M_i| - 1)$. In particular, given $\lambda > 0$, every machine in an NE satisfies*

$$|M_i| \leq 1 + \left\lfloor \frac{T - P}{\lambda} \right\rfloor.$$

Proposition 3 (Immediate loss of the full payoff set)**.** *For any finite state cost $\lambda > 0$, some payoff in V^* ceases to be a Nash equilibrium payoff. In particular, $\lambda_1 = 0$.*

Indeed, $\pi_i - \lambda|M_i| \geq P - \lambda$ and $\pi_i \leq T$ give Lemma 4. The resulting finite bound on machine size leaves only finitely many equilibrium payoffs for any fixed $\lambda > 0$, whereas V^* is infinite.

4.4 Mutual cooperation in Nash equilibrium: λ_2

Grim trigger (cooperates until the opponent defects, then defects forever) fails at positive costs because its unused punishment state can be discarded. A punishment phase used at the beginning of play can avoid this deviation.

Proposition 4 (Threshold for mutual cooperation)**.** *The maximum cost threshold for sustaining the mutual cooperation payoff* (R, R) *in a Nash equilibrium is*

$$\lambda_2 = \max_{N \in \mathbb{Z}_{>0}} g(N), \qquad g(N) = \frac{N(R-P) - (T-R)}{N(N+1)}.$$

For sufficiency, consider an $(N+1)$-state password machine that plays N defections and then cooperates, advancing only when the opponent matches its action and resetting otherwise. Self-play gives $(D, D)^N (C, C)^\infty$. Reaching mutual cooperation requires a deviator to count N consecutive inputs D before outputting C, requiring at least $N+1$ states. Smaller machines earn at most $(NP+T)/(N+1)$, attained by M_D; larger machines cannot improve on R when $g(N) \geq 0$. The binding inequality is

$$R - \lambda(N+1) \geq \frac{NP+T}{N+1} - \lambda \quad \Longleftrightarrow \quad \lambda \leq g(N).$$

For necessity, state minimization and Abreu and Rubinstein (1988, Theorem 1*) imply that any equilibrium yielding (R, R) has a common machine size $n \geq 2$, one cooperative cycle state, and matching actions in its introductory phase. Under constant input D, the defection states cannot form a cycle, so M_D earns at least $P + (T-P)/n$. The equilibrium inequality then gives $\lambda \leq g(n-1)$.

4.5 Semi-perfect equilibrium and alternating exploitation

Proposition 5 (Mutual cooperation is not an SPE)**.** *For every* $\lambda > 0$*, no semi-perfect equilibrium yields the payoff* (R, R)*.*

After the (C, C) cycle begins, replacing a multi-state machine by M_C preserves payoff R and saves states; two one-state cooperators are not an NE.

Proposition 6 (Upper bound on λ_3)**.** *For* $\lambda > 0$ *with* $\lambda > (T+S)/2 - P$*, the unique semi-perfect equilibrium is* (M_D, M_D)*. Hence*

$$\lambda_3 \leq \max\{0, (T+S)/2 - P\}.$$

By Lemma 2 and Rubinstein's characterization, a non-defection SPE cycle has total payoff $T+S$. Comparing two machines of at least two states with the guaranteed utilities $P - \lambda$ gives $2P - 2\lambda \leq T + S - 4\lambda$.

To obtain lower bounds, consider the two-state alternator M^a, with $\sigma(c) = C$, $\sigma(d) = D$, and

$$\mu(c, \cdot) = d, \qquad \mu(d, C) = c, \qquad \mu(d, D) = d.$$

Player 1 starts at c and Player 2 at d. The pair $(M^a(c), M^a(d))$ alternates between (C, D) and (D, C), earning each player $(T+S)/2$ and visiting both states every cycle.

Proposition 7 (Survival threshold of the alternator)**.** *For* $\lambda > 0$*, the pair* $(M^a(c), M^a(d))$ *is a Nash equilibrium if and only if*

$$\lambda \leq \lambda_{\text{alt}} := \min\left\{\frac{T-R}{2}, \frac{T+S}{2} - P\right\}.$$

Whenever it is an NE, it is also an SPE. In the regular regime $2P \leq R+S$*, this threshold is* $(T-R)/2$*.*

The one-state deviations earn P and $(R+S)/2$. Comparing these, less one state cost, with $(T+S)/2 - 2\lambda$ gives the two bounds; when they hold, larger machines cannot improve the game payoff. The same comparisons apply at either point of the cycle.

Proposition 8 (Lower bounds)**.** *If* $T + S > 2P$*, then* $\lambda_3 \geq \lambda_{\text{alt}}$ *and* $\lambda_4 \geq \max\{\lambda_2, \lambda_{\text{alt}}\}$*.*

Thus the proven bounds in the regular regime are

$$\lambda_1 = 0, \quad \lambda_{\text{alt}} \leq \lambda_3 \leq (T+S)/2 - P, \quad \max\{\lambda_2, \lambda_{\text{alt}}\} \leq \lambda_4 \leq R - P.$$

The two constructions motivate the conjectures below; their optimality for λ_3 and λ_4 remains open.

5 Computational Evidence and Future Directions

Preliminary computational testing, exhaustive over machine pairs with at most two states and supplemented by seeded searches up to three states, corroborates the conjectures $\lambda_3 = \lambda_{\text{alt}}$ and $\lambda_4 = \max\{\lambda_2, \lambda_{\text{alt}}\}$ across five tested payoff parametrizations satisfying $2P \leq R+S$. The tested cases include both $\lambda_2 < \lambda_{\text{alt}}$ and $\lambda_2 > \lambda_{\text{alt}}$, suggesting that eventual mutual cooperation and alternating exploitation need not disappear in a fixed order. Establishing these conjectures for machines of arbitrary size remains open. A further direction is to use these thresholds in resource-aware synthesis: constructing finite-state strategies that satisfy an eventual-cooperation specification while remaining stable against cost-sensitive replacement.

References

Dilip Abreu and Ariel Rubinstein. The structure of nash equilibrium in repeated games with finite automata. *Econometrica: Journal of the Econometric Society*, pages 1259–1281, 1988.

Robert J Aumann. Survey of repeated games. *Essays in game theory and mathematical economics in honor of Oskar Morgenstern*, 1981.

Jeffrey S Banks and Rangarajan K Sundaram. Repeated games, finite automata, and complexity. *Games and Economic Behavior*, 2(2):97–117, 1990.

Kenneth G Binmore and Larry Samuelson. Evolutionary stability in repeated games played by finite automata. *Journal of economic theory*, 57(2):278–305, 1992.

David J Cooper. Supergames played by finite automata with finite costs of complexity in an evolutionary setting. *Journal of Economic Theory*, 68(1):266–275, 1996.

Dana Fisman, Orna Kupferman, and Yoad Lustig. Rational synthesis. In *International Conference on Tools and Algorithms for the Construction and Analysis of Systems*, pages 190–204. Springer, 2010.

James W Friedman. A non-cooperative equilibrium for supergames. *The Review of Economic Studies*, 38(1): 1–12, 1971.

Nimrod Megiddo and Avi Wigderson. On play by means of computing machines: preliminary version. In *Theoretical aspects of reasoning about knowledge*, pages 259–274. Elsevier, 1986.

Abraham Neyman. Bounded complexity justifies cooperation in the finitely repeated prisoners' dilemma. *Economics letters*, 19(3):227–229, 1985.

Ariel Rubinstein. Finite automata play the repeated prisoner's dilemma. *Journal of economic theory*, 39(1): 83–96, 1986.

Matthijs van Veelen and Julián García. In and out of equilibrium ii: Evolution in repeated games with discounting and complexity costs. *Games and Economic Behavior*, 115:113–130, 2019.