Convergence of the Iterated Prisoner's Dilemma game
[*]

Martin Dyer[†] Leslie Ann Goldberg[‡] Catherine Greenhill[§]
Gabriel Istrate[¶] Mark Jerrum[‖]

May 17, 2000

Abstract

Co-learning is a model involving agents from a large population, who interact by playing a fixed game and update their behaviour based on previous experience and the outcome of this game. The Highest Cumulative Reward rule is an update rule which ensures the emergence of cooperation in a population of agents without centralized control, for various games and interaction topologies. We analyse the convergence rate of this rule when applied to the Iterated Prisoner's dilemma game, proving that the convergence rate is optimal when the interaction topology is a cycle and exponential when it is a complete graph.

Keywords: Prisoner's dilemma, convergence, co-learning, Highest Cumulative Reward rule

1 Introduction

The setup of the *theory of learning in games* [3] is fairly simple: agents from a large population interact by playing a fixed game and update their behavior based on previous experience and the outcome of this game.

---

[*] This work was supported in part by the EPSRC Research Grant "Sharper Analysis of Randomised Algorithms: a Computational Approach" and by the ESPRIT Projects RAND-APX and ALCOM-FT.

[†] School of Computer Studies, University of Leeds, Leeds LS2 9JT, United Kingdom, email: dyer@scs.leeds.ac.uk

[‡] Department of Computer Science, University of Warwick, Coventry CV4 7AL, United Kingdom, email: leslie@dcs.warwick.ac.uk

[§] Department of Mathematics and Statistics, University of Melbourne, Parkville VIC 3502, Australia, email: csg@ms.unimelb.edu.au Supported by an Australian Research Council Postdoctoral Fellowship

[¶] Center for Nonlinear Science and CIC-3 Division, Los Alamos National Laboratory, Mail Stop B258, Los Alamos, NM 87545, email: istrate@lanl.gov

[‖] School of Computer Science, University of Edinburgh, King's Building, Edinburgh EH9 3JZ, United Kingdom, email: mrj@dcs.ed.ac.uk

Particular models are obtained by specifying the game, the interaction topology and the update mechanism. The main thrust of the theory is to explain the emergence of various types of game-theoretic equilibria as the outcome of an evolutionary process (rather than seeing them as steady-state properties).

Such models have recently been quite popular in Complex Systems Theory [2], and also in distributed Artificial Intelligence, mainly in connection with *multi-agent Reinforcement Learning* (see e.g. [7] for a discussion). A particularly interesting framework, dubbed *co-learning* was introduced by Shoham and Tennenholtz [9]. They were interested in mechanisms that ensure the emergence of cooperation in a population of agents without centralized control and introduced a rule very similar to reinforcement learning, called *highest cumulative reward* (HCR), which guarantees this for a variety of interaction topologies and games. In particular they studied two games, the *convention game* and the *cooperation* game. The latter is our focus of study, and is often called the *Iterated Prisoner's Dilemma.*

One weakness of the co-learning framework, discussed by Shoham and Tennenholtz, is its asymptotic character, that is the lack of specific predictions on the time needed for the emergence of cooperation (henceforth called *absorbtion time*). They wrote [9]:

> It would be natural to expect that subsequent investigations would provide finer and finer lower and upper bounds, increasing our understanding of HCR. Unfortunately this has not been our experience. What we found instead was that rather specific properties of particular games being played flavor the dynamics so strongly that it appears extremely difficult to arrive at the level of a particular update function.

Kittock [4] further undertook a study of this model, in particular providing evidence that for HCR on Iterated Prisoner's Dilemma this time is polynomial on cycles and exponential on complete graphs. His paper, however contains only heuristic and experimental, rather than rigorous results.

In this note we investigate the absorbtion time of the HCR strategy by proving such results for two of the cases studied by Kittock; namely, for cycles and complete graphs.

2 Preliminaries

We first discuss the particular dynamics we are concerned with, the HCR update rule with 1-step memory from [4], without any other reference to the motivation and specific details of the framework from [9, 4]. We are given a population of $n \geq 4$ agents situated at the vertices of a connected graph $G = (V, E)$. Each agent has a state $X(i) \in \{-1, 1\}$, *cooperate*, encoded as 1 or *defect*, encoded as -1. At each stage we choose a pair $\{i, j\} \in E$ uniformly at random and replace $X(i)$ and $X(j)$ by $X(i)X(j)$. The process has 1-step memory because each agent i remembers its current state $X(i)$ but does not remember any previously-held states. From now on we call this process the *HCR process*, which should be considered as an abbreviation for "repeated application of the HCR update rule with 1-step memory for the Iterated Prisoner's Dilemma". Note that, from each player's point of view, the update rule is the so-called Pavlov update rule for Iterated Prisoner's Dilemma [1, 6]). We refer to the 1 values as *pluses* and the -1 values as *minuses*. The state X^* with $X^*(i) = 1$ for all $v \in V$ is an absorbing state of this process. If G contains no isolated vertices then X^* is the unique absorbing state and there exists a sequence of moves which can transform X to X^*, for every $X \in \{1, -1\}^V$.

We are interested in the *absorbtion time*; that is, the time required for the HCR process to reach the stable state. Shoham and Tennenholtz [8] proved that the absorbtion time is at least $\Omega(n \log n)$. In this paper, we investigate two special classes of graphs, namely cycles and complete graphs. We prove the following theorems, showing that the HCR process has optimal mixing time when G is a cycle, and exponential mixing time when G is a complete graph.

Theorem 1 *Let G be a cycle on n vertices and let $\varepsilon > 0$ be given. The probability that the HCR process has not reached the stable state in*

$$\frac{49}{2} n \log\left(\frac{7n}{4\varepsilon}\right)$$

steps is at most $14\varepsilon/47$.

Theorem 2 *Let T_n denote the absorbtion time of HCR process on the graph K_n starting from a configuration X_0 with at most $.61n$ nodes labelled 1. Then there exists $c > 1$ such that with probability $1 - o(1)$ we have $T_n \geq c^n$.*

3 Optimal absorbtion on cycles

Let G be a cycle on the the vertex set $[n] = \{0, \ldots, n-1\}$. That is, G has n edges $\{i, i+1\}$ for $0 \leq i < n$. Here, and throughout the paper, addition and subtraction on vertices is performed modulo n.

We define a potential function $\psi : \{1, -1\}^V \to \mathbb{R}$ to measure the distance of a given state X from the absorbing state X^*. First, we must introduce some terminology. Let $X \in \{1, -1\}^V$ be given. A *run* in X is an interval $[i, j]$ where $0 \leq i,\ j < n$, such that $X(\ell) = -1$ for $\ell = i, i+1, \ldots, j-1, j$ and $X(i-1) = 1$, $X(j+1) = 1$. (It is possible to have $j < i$, since we are working modulo n.) Clearly all runs are disjoint. We can define the set $\mathcal{R}(X)$ of all runs in X. By convention, the all-minuses configuration is *not* considered a run, since it has no bordering pluses.

Suppose that $r = [i, j]$. The *length* of the run r, denoted by $\ell(r)$, equals the number of minuses in the run. We will refer to a run of length ℓ as an ℓ-run. A 1-run will also be called a *singleton* and a 2-run will also be called a *pair*. Then the potential function ψ is given by

$$\begin{aligned} \psi(X) \;=\; & |\{i : X(i) = -1\}| + \beta \cdot |\mathcal{R}(X)| + \gamma \cdot |\{r \in \mathcal{R}(X) : r \text{ is a singleton}\}| \\ & + \delta \cdot |\{r \in \mathcal{R}(X) : r \text{ is a pair}\}|. \end{aligned}$$

The parameters β, γ and δ will be set below. Note that a singleton is a barrier to absorbtion since a singleton minus cannot be changed to a plus in one step. The singleton must first become part of a longer run. So we set $\gamma > 0$ to penalise singletons. On the other hand, pairs give the opportunity for two minuses to be changed at one step. Thus pairs are helpful, and we reflect this by setting $\delta < 0$. We also set $\beta > 0$. Clearly $\psi(X^*) = 0$ for any values of β, γ, δ since $X^*(i) = 1$ for all i, and $\mathcal{R}(X^*) = \emptyset$. For ψ to be a well-defined potential function, we must also show that $\psi(X) > 0$ whenever $X \neq X^*$. This is achieved if $-2 < \delta < 0$, since there can be at most half as many pairs in X as there are minuses.

3.1 The analysis

We now analyse the HCR process using the potential function ψ. Let $X_0 \in \{1, -1\}^V$ be fixed. Clearly if $X_0 = X^*$ there is nothing to prove. So, suppose that X_0 contains at least one minus. Let X_1 be the result of performing one step of the dynamics from starting point X_0. We will find an upper bound for $\mathbf{E}\left[\psi(X_1) - \psi(X_0)\right]$.

Note that each edge overlaps at most one run in $\mathcal{R}(X_0)$, and that there are $\ell + 1$ edges which overlap a given ℓ-run. Specifically, if $r = [i, j]$ then these $\ell(r) + 1$ edges are

$$\{i-1, i\}, \ldots, \{j, j+1\}.$$

Let $L(X_0)$ be defined by

$$L(X_0) = \sum_{r \in \mathcal{R}(X_0)} (\ell(r) + 1).$$

Then $L(X_0)$ equals the number of edges which overlap some run in X_0. Denote by $\mathbf{E}\left[\psi(X_1) - \psi(X_0) \mid e\right]$ the value of $\psi(X_1) - \psi(X_0)$ given that the edge e has been chosen by the HCR process in step 1. Let r be an ℓ-run and let

$$\sigma(r) = \sum_{e \text{ overlaps } r} \mathbf{E}\left[\psi(X_1) - \psi(X_0) \mid e\right].$$

By definition we have

$$\mathbf{E}\left[\psi(X_1) - \psi(X_0)\right] = \frac{1}{n} \sum_{e \in E} \mathbf{E}\left[\psi(X_1) - \psi(X_0) \mid e\right],$$

since there are n edges in G. When X_0 contains both pluses and minuses we can also state that

$$\mathbf{E}\left[\psi(X_1) - \psi(X_0)\right] = \frac{1}{n} \sum_{r \in \mathcal{R}(X_0)} \sigma(r),$$

since runs are disjoint and an edge which does not overlap a run makes no change to X_0. Let $M = M(X_0)$ be defined by

$$M = \max\left\{ \frac{\sigma(r)}{\ell(r) + 1} \mid r \in \mathcal{R}(X_0) \right\}.$$

That is, M is the maximum over all runs of the average contribution of each edge in that run. The way in which M will be used is described below. We ignore two configurations: the all-pluses configuration X^*, and the all-minuses configuration. The latter is treated separately in Section 3.2.

Lemma 1 *Suppose that X_0 contains both pluses and minuses. With M, ψ and L defined as above, we have*

$$\mathbf{E}\left[\psi(X_1)\right] \le \left(1 + \frac{ML(X_0)}{\psi(X_0)n}\right)\psi(X_0).$$

Proof. From above, we have

$$\begin{aligned}\mathbf{E}\left[\psi(X_1) - \psi(X_0)\right] &= \frac{1}{n}\sum_{r\in\mathcal{R}(X_0)} \sigma(r)\\ &\le \frac{1}{n}\sum_{r\in\mathcal{R}(X_0)} (\ell(r)+1)M\\ &= \frac{ML(X_0)}{n}\end{aligned}$$

We can rearrange this inequality to give

$$\mathbf{E}\left[\psi(X_1)\right] \le \psi(X_0) + \frac{ML(X_0)}{n} = \left(1 + \frac{ML(X_0)}{\psi(X_0)n}\right)\psi(X_0),$$

as stated. □

Suppose that the values of β, γ, δ could be set to ensure that $M < 0$. Then, by Lemma 1, the value of ψ *decreases* in expectation at every step. This will be used in Section 3.2 to calculate an upper bound for the absorbtion time of the HCR process.

Let $r = [i, j]$ be a run. Then there are two *outer rim* edges associated with r, namely $\{i-1, i\}$ and $\{j, j+1\}$. If r has length at least 3 then there are also two *inner rim* edges associated with r, namely $\{i, i+1\}$ and $\{j-1, j\}$. If r is a singleton then there are no inner rim edges, while if r is a pair $[i, i+1]$ then there is a unique inner rim edge $\{i, i+1\}$. All other edges which overlap r are strictly inside the interval $[i, j]$, and we call these edges *internal* edges.

Suppose that there are two runs in $\mathcal{R}(X_0)$ which are only separated by a single plus, i.e. $[i, j]$ and $[j+2, k]$ for some i, j, k. Then there are two edge choices $\{j, j+1\}$ and $\{j+1, j+2\}$ which cause the two runs to merge (note that these edges are both outer rim edges for the runs which they overlap). For simplicity, we will first assume that there are no edge choices which cause runs to merge. That is, in Lemma 2 we assume that all adjacent runs in $\mathcal{R}(X_0)$ are separated by at least two pluses. By carefully choosing values for β, γ and δ in this case, we show that M is negative: specifically $M = -1/14$. In Lemma 3 we return to configurations which contain adjacent runs separated by a single plus.

Before presenting Lemma 2, we make a few general remarks. When all adjacent runs are separated by at least two pluses, choosing an outer rim edge will always cause r to increase in length by 1, introducing an extra minus. Similarly, choosing an inner rim edge will always cause r to decrease in length by 2, changing two minuses to pluses. When the length of r is small there might be additional effects from these four edges, as we shall see. When any internal edge is chosen, the run r is *split* into two runs which are separated by two pluses. If the two runs have length k and ℓ we say that this edge choice produces a (k, ℓ)-*split*.

We can now prove that M is negative for certain fixed values of β, γ and δ, when X_0 contains both pluses and minuses and all adjacent runs are separated by at least two pluses.

Lemma 2 *Let X_0 contain both pluses and minuses, and suppose that adjacent runs in X_0 are separated by at least two pluses. Then setting $\beta = 27/14$, $\gamma = 4/7$ and $\delta = -4/7$ we obtain $M = -1/14$.*

Proof. We will consider runs r of different lengths in turn, and calculate $\sigma(r)/(\ell(r)+1)$ in each case. Then M is the maximum of these values.

A 1-run. Let r be a 1-run $[i, i]$. The only edges which overlap r are the outer rim edges $\{i-1, i\}$ and $\{i, i+1\}$. When either of these edges are chosen, a vertex adjacent to the 1-run changes from a plus to a minus. This introduces an extra minus and changes a 1-run (singleton) to a 2-run (a pair), without changing the total number of runs. Therefore

$$\frac{\sigma(r)}{2} = 1 - \gamma + \delta = -\frac{1}{7}. \tag{1}$$

A 2-run. Suppose that $r = [i, i+1]$. There are 3 edges which overlap r. When either of the outer rim edges $\{i-1, i\}$ or $\{i+1, i+2\}$ are chosen the 2-run becomes a 3-run, introducing an extra minus and deleting a pair. There is only one inner rim edge, the edge $\{i, i+1\}$. When this edge is chosen, both minuses in the pair become pluses. Here we lose two minuses and delete a pair, decreasing the number of runs by 1. Adding these contributions and dividing by 3 we find that

$$\frac{\sigma(r)}{3} = \frac{2(1-\delta) - (2+\beta+\delta)}{3} = -\frac{\beta + 3\delta}{3} = -\frac{1}{14}. \tag{2}$$

A 3-run. Suppose that $r = [i, i+2]$. There are 4 edges which overlap r, namely the two outer rim edges and the two inner rim edges. Choosing an outer rim edge turns the 3-run into a 4-run, introducing an extra minus. Choosing an inner rim edge turns the 3-run into a 1-run. Hence

$$\frac{\sigma(r)}{4} = \frac{2 + 2(-2+\gamma)}{4} = \frac{-1+\gamma}{2} = -\frac{3}{14}. \tag{3}$$

A 4-run. Suppose that $r = [i, i+3]$ for some i. There are 5 edges which overlap r. Choosing an outer rim edge causes r to increase in length by 1, introducing a new minus. Choosing an inner rim edge causes the length of r to decrease by 2: in this case this introduces a new pair. Finally, there is one internal edge $\{i+1, i+2\}$. Choosing this edge produces a $(1,1)$-split. This introduces two singletons and increases the total number of runs by 1, while removing two minuses. Adding these contributions together and dividing by 5, we obtain

$$\frac{\sigma(r)}{5} = \frac{2 + 2(-2+\delta) + (-2+\beta+2\gamma)}{5} = \frac{-4+\beta+2\gamma+2\delta}{5} = -\frac{29}{70}. \tag{4}$$

A 5-run. Let $r = [i, i+4]$ for some i. There are six edges which overlap r. Choosing an outer rim edge causes r to increase in length by 1. Choosing an inner rim edge causes the length or r to decrease by 2. There are two internal edges, $\{i+1, i+2\}$ and $\{i+2, i+3\}$. Choosing either of these edges produces a $(1,2)$-split, deleting two minuses, introducing a singleton and a pair, as well as increasing the number of runs by 1. Adding the contributions from all of these edges together, and dividing by 6, we obtain

$$\frac{\sigma(r)}{6} = \frac{2 - 4 + 2(-2+\beta+\gamma+\delta)}{6} = \frac{-3+\beta+\gamma+\delta}{3} = -\frac{5}{14}. \tag{5}$$

A 6-run. Let $r = [i, i+5]$. There are 7 edges which overlap r. If an outer rim edge is chosen then r increases in length by 1. If an inner rim edge is chosen then r decreases in length by 2. There are 3 internal edges. Choosing $\{i+1, i+2\}$ or $\{i+3, i+4\}$ produces a $(1,3)$-split, decreasing the number of minuses by 2 while increasing the number of singletons and the number of runs by 1. Finally, choosing the edge $\{i+2, i+3\}$ produces a $(2,2)$-split, decreasing the number of minuses by 2, increasing the number of runs by 1 and the number of pairs by 2. Combining this information we find that

$$\frac{\sigma(r)}{7} = \frac{2 - 4 + 2(-2+\beta+\gamma) + (-2+\beta+2\delta)}{7} = \frac{-8+3\beta+2\gamma+2\delta}{7} = -\frac{31}{98}. \tag{6}$$

An ℓ-run, where $\ell \geq 7$. Now suppose that $r = [i, j]$ is an ℓ-run for some $\ell \geq 7$. Choosing either of the two outer rim edges causes r to increase in length by 1. Choosing either of the two inner rim edges causes r to decrease in length by 2. There are 4 internal edges which need careful analysis. Choosing either $\{i+1, i+2\}$ or $\{j-2, j-1\}$ produces a $(1, \ell-3)$-split, introducing a singleton and increasing the number of runs by 1, while decreasing the number of minuses by 2. Similarly, choosing either $\{i+2, i+3\}$ or $\{j-3, j-2\}$ produces a $(2, \ell-4)$-split, introducing a pair and increasing the number of runs by 1, while decreasing the number of minuses by 2. There are $\ell - 7$ other internal edges which split r into pairs of runs, each of length at least 3. In each case, the number of minuses decreases by 2 while the number of runs increases by 1, but the numbers of singletons and pairs are unchanged. We obtain

$$\begin{aligned}
\frac{\sigma(r)}{\ell+1} &= \frac{2-4+2(-2+\beta+\gamma)+2(-2+\beta+\delta)+(\ell-7)(-2+\beta)}{\ell+1} \\
&= \beta - 2 - \frac{4\beta - 6 - 2\gamma - 2\delta}{\ell+1} \\
&= -\frac{1}{14} - \frac{12}{7(\ell+1)}. \qquad (7)
\end{aligned}$$

Now M is equal to the maximum of the right hand sides of (1)–(7). It is easy to verify that the maximum is $-1/14$, as stated. □

For the remainder of this section, the values of $\beta = 27/14$, $\gamma = 4/7$ and $\delta = -4/7$ are fixed. These values were chosen without explanation for use in the proof of Lemma 2 above. They were originally derived by setting $\beta = 2 - \eta$, $\gamma = 1/2 + \eta$ and $\delta = -\gamma$, and choosing η to minimize M. The interested reader can easily verify that $\eta = 1/14$ is the optimal choice.

We now show that the value $M = -1/14$ can still be used in Lemma 1 even when the initial configuration has adjacent runs which are separated by a single plus.

Lemma 3 *Suppose that* $X_0 \in \{1, -1\}$ *contains both pluses and minuses. Then the conclusion of* Lemma 1 *holds with*

$$M = -\frac{1}{14}.$$

Proof. By Lemma 2, we have $M = -1/14$ whenever no two adjacent runs in X_0 are separated by a single plus. So now suppose that there are exactly s distinct values $i \in \{0, \ldots, n-1\}$ such that $X_0(i-1) = -1$, $X_0(i) = 1$ and $X_0(i+1) = -1$, where $s \geq 1$. We will call such an i a *rim vertex.* Define a new cycle $G' = (V', E')$ from G by splitting the vertex i into two new vertices, i' and i'', for each rim vertex i. Thus G' is a graph on $n + s$ vertices. Let the edges of G' be obtained from the edges of G by deleting the edges $\{i-1, i\}$, $\{i, i+1\}$ and adding the edges $\{i-1, i'\}$, $\{i', i''\}$, $\{i', i+1\}$, for each rim vertex i. Thus G' forms a cycle on $n + s$ vertices. Construct the configuration ${X_0}' \in \{1, -1\}^{V'}$ from X_0 by replacing the single plus at i by two pluses on i', i'', for each rim vertex i. That is, let

$${X_0}'(j) = \begin{cases} X_0(j) & \text{if } j \text{ is unprimed,} \\ 1 & \text{otherwise.} \end{cases}$$

By definition, ${X_0}'$ has no two adjacent runs separated by a single plus. Note also that $L({X_0}') = L(X_0)$. Let ${X_1}'$ be the result of running the HCR process for one step from ${X_0}'$. Combining Lemma 1 and Lemma 2, we see that

$$\mathbf{E}\left[\psi({X_1}') - \psi({X_0}')\right] \leq -\frac{L({X_0}')}{14(n+s)}.$$

Suppose that we could show that

$$\sum_{e \in E(G')} \mathbf{E}\left[\psi({X_1}') - \psi({X_0}') \mid e\right] \geq \sum_{e \in E(G)} \mathbf{E}\left[\psi(X_1) - \psi(X_0) \mid e\right]. \tag{8}$$

Then we would have

$$\begin{aligned}
\mathbf{E}\left[\psi({X_1}') - \psi({X_0}')\right] &= \frac{1}{n+s} \sum_{e \in E(G')} \mathbf{E}\left[\psi({X_1}') - \psi({X_0}') \mid e\right] \\
&\geq \frac{1}{n+s} \sum_{e \in E(G)} \mathbf{E}\left[\psi(X_1) - \psi(X_0) \mid e\right] \\
&= \frac{n}{n+s} \mathbf{E}\left[\psi(X_1) - \psi(X_0)\right].
\end{aligned}$$

From this we could conclude that

$$\begin{aligned}
\frac{n}{n+s} \mathbf{E}\left[\psi(X_1) - \psi(X_0)\right] &\leq \mathbf{E}\left[\psi({X_1}') - \psi({X_0}')\right] \\
&\leq -\frac{L({X_0}')}{14(n+s)} \\
&= -\frac{L(X_0)}{14(n+s)}.
\end{aligned}$$

Multiplying this inequality through by $(n+s)/n$ proves the lemma. Hence it suffices to establish (8).

It is not difficult to see that any edge which does not overlap a rim vertex in X_0 makes the same contribution in both the primed and unprimed settings. For these edges e belong to both $E(G)$ and $E(G')$, and

$$\mathbf{E}\left[\psi({X_1}') - \psi({X_0}') \mid e\right] = \mathbf{E}\left[\psi(X_1) - \psi(X_0) \mid e\right].$$

Therefore, to prove (8) it suffices to prove that $Y' > Y$ for all rim vertices i, where

$$Y = \mathbf{E}\left[\psi(X_1) - \psi(X_0) \mid \{i-1, i\}\right] + \mathbf{E}\left[\psi(X_1) - \psi(X_0) \mid \{i, i+1\}\right]$$

and

$$Y' = \mathbf{E}\left[\psi({X_1}') - \psi({X_0}') \mid \{i-1, i'\}\right] + \mathbf{E}\left[\psi({X_1}') - \psi({X_0}') \mid \{i'', i+1\}\right].$$

(Clearly the edge $\{i', i''\}$ makes no contribution to $E\left[\psi({X_1}') - \psi({X_0}')\right]$.) Let r_1 and r_2 be the two runs which are separated by i in X_0, and define a and b by

$$a = |\{j \in \{1,2\} \mid r_j \text{ is a singleton}\}| \qquad \text{and} \qquad b = |\{j \in \{1,2\} \mid r_j \text{ is a pair}\}|.$$

Then $0 \leq a + b \leq 2$. Consider choosing either $\{i-1, i'\}$ or $\{i'', i+1\}$ for ${X_0}'$. Clearly either choice will cause a minus to be introduced. For a of these choices a singleton is removed and a pair is created, while for b of these choices a pair is removed. Therefore

$$Y' = 2 - a\gamma + a\delta - b\delta.$$

Now consider choosing either $\{i-1, i\}$ or $\{i, i+1\}$ in X_0. The expected change of ψ is identical for either choice. Choosing either of these edges introduces a minus, decreases the number of runs by 1, and deletes all singletons or pairs which are present in X_0. The merged run which is created has length $\ell(r_1) + \ell(r_2) + 1 \geq 3$, so no singletons or pairs are created. Therefore

$$Y = 2(1 - \beta - a\gamma - b\delta).$$

Hence, using the values of β, γ and δ we obtain

$$Y' - Y = 2\beta + a(\gamma + \delta) + b\delta \geq 2(\beta + \delta) > 0,$$

proving the lemma. $\square$

3.2 Bounding the absorbtion time

We have fixed $\beta = 27/14$, $\gamma = 4/7$, $\delta = -4/7$. Recall that $L(X_0) = \sum_{r \in \mathcal{R}(X_0)} (\ell(r) + 1)$. Combining Lemmas 1, 2, 3 we obtain

$$\mathbf{E}\left[\psi(X_1)\right] \leq \left(1 - \frac{L(X_0)}{14\,\psi(X_0) n}\right) \psi(X_0), \tag{9}$$

for all X_0 which contain both pluses and minuses. We need the following result.

Lemma 4 *Let $X \in \{1,-1\}^V$ and let ψ, L be as defined above. Then*

$$\psi(X) \leq \frac{7L(X)}{4}$$

if X contains both pluses and minuses, while

$$\frac{47}{14} \leq \psi(X) \leq \frac{7n}{4} \tag{10}$$

for all $X \neq X^*$.

Proof. First suppose that X contains both pluses and minuses. Let $\psi(r)$ denote the potential of the run r, for all $r \in \mathcal{R}(X)$. That is,

$$\psi(r) = \begin{cases} 1+\beta+\gamma & \text{if } r \text{ is a singleton,} \\ 2+\beta+\delta & \text{if } r \text{ is a pair,} \\ \ell(r)+\beta & \text{otherwise.} \end{cases}$$

Clearly

$$\psi(X) = \sum_{r\in\mathcal{R}(X)} \psi(r).$$

It is not difficult to check that the inequality

$$\frac{\psi(r)}{\ell(r)+1} \leq \frac{7}{4}$$

holds, with equality if and only if r is a singleton. Hence

$$\psi(X) = \sum_{r\in\mathcal{R}(X)} \psi(r) \leq \sum_{r\in\mathcal{R}(X)} \frac{7(\ell(r)+1)}{4} = \frac{7L(X)}{4},$$

as stated. Now $L(X)$ denotes the number of edges which overlap some run in X. Since there are at exactly n edges in G, it follows that

$$\psi(X) \leq \frac{7n}{4}$$

whenever X contains both pluses and minuses. Since the all-minuses configuration has potential n, this proves the upper bound in (10). Finally, note that

$$\psi(r) \geq \frac{47}{14},$$

with equality if and only if r is a pair. Therefore the lowest potential of all configurations with both minuses and pluses is obtained on any configuration which contains a unique run, this unique run being a pair. The all-minuses configuration has potential n, but we have assumed that n is at least 4. This proves the lower bound in (10). □

Proof of Theorem 1. Combining (9) and the upper bound given in (10), we can conclude that

$$\mathbf{E}\left[\psi(X_1)\right] \leq \left(1 - \frac{2}{49n}\right)\psi(X_0) \tag{11}$$

for all X_0 which contain both pluses and minuses. However, (11) also holds for the all-minuses configuration, as follows. Let $\tilde{X}$ be the all-minuses configuration, defined by $\tilde{X}_0(i) = -1$ for all i. Let $\tilde{X}_1$ be the result of running the HCR proces for one step from $\tilde{X}_0$. No matter which edge is chosen, the number of minuses decreases by 2 and the number of runs increases from 0 to 1. Therefore $\mathbf{E}\left[\psi(\tilde{X}_1) - \psi(\tilde{X}_0)\right] = \beta - 2 = -1/14$. Since $\psi(\tilde{X}_0) = n$, we conclude that

$$\mathbf{E}\left[\psi(\tilde{X}_1)\right] = \left(1 - \frac{1}{14n}\right)\psi(\tilde{X}_0) < \left(1 - \frac{2}{49n}\right)\psi(\tilde{X}_0),$$

as claimed.

So now let $X_0 \in \{1, -1\}^V$ satisfy $X_0 \neq X^*$. Starting from X_0, run the HCR process for t steps and let the resulting state be X_t. By applying (11) iteratively t times we obtain

$$\mathbf{E}\left[\psi(X_t)\right] \leq \left(1 - \frac{2}{49n}\right)^t \psi(X_0) \leq \left(1 - \frac{2}{49n}\right)^t \frac{7n}{4},$$

using the first statement of Lemma 4 for the last inequality. Let $\varepsilon > 0$ be given. Whenever

$$t \geq \frac{49}{2} n \log\left(\frac{7n}{4\varepsilon}\right)$$

we have $\mathbf{E}\left[\psi(X_t)\right] \leq \varepsilon$. Using (10), any nonzero value of ψ must be at least 47/14. Applying Markov's Lemma, we have

$$\text{Prob}\left[\psi(X_t) \neq 0\right] = \text{Prob}\left[\psi(X_t) \geq \frac{47}{14}\right] \leq \frac{14\varepsilon}{47}.$$

This completes the proof. □

4 Exponential absorbtion on the complete graph

In this section we prove Theorem 2, showing that the absortion time of the HCR process is exponential on the complete graph K_n.

We will use the notation $X_t \in \{1, -1\}^n$ to refer to a configuration after t steps of the HCR process. Let N_t be the number of nodes in X_t with label 1. Clearly X_t is equal to the all-pluses absorbing state if and only if $N_t = n$. We will assume that $N_0 \leq 0.61n$. The basis of our proof is the observation that the process N_t is simple to analyse, even if the process X_t is not. Let p_t, q_t denote the labels of the two nodes chosen at stage t. Then the transition probabilities of N_t are given by the following rule:

$$N_{t+1} = \begin{cases} N_t - 1 & \text{if } p_t q_t = -1 \text{ (probability } N_t(n - N_t)/\binom{n}{2}\text{)}, \\ N_t & \text{if } p_t = q_t = 1 \text{ (probability } \binom{N_t}{2}/\binom{n}{2}\text{)}, \\ N_t + 2 & \text{if } p_t = q_t = -1 \text{ (probability } \binom{n-N_t}{2}/\binom{n}{2}\text{)}. \end{cases}$$

Let I denote the interval $[0.61n, 0.7n]$.

Lemma 5 *Suppose that* $N_\tau \in I$. *Let* $T = \min\{t > \tau \mid N_\tau \notin I\}$. *Then the process* $N_\tau, \ldots, N_T$ *is stochastically dominated by the process* $Q_\tau, \ldots, Q_T$ *where* Q_t *is the simple Markov chain in which* $Q_\tau = N_\tau$ *and for* $t \geq \tau$,

$$Q_{t+1} = \begin{cases} Q_t - 1 & \textit{with probability } 0.35, \\ Q_t & \textit{with probability } 0.49, \\ Q_t + 2 & \textit{with probability } 0.16. \end{cases}$$

Proof. For $t \in [\tau, T-1]$, the pair (N_{t+1}, Q_{t+1}) can be chosen from the following joint distribution which satisfies $N_t \leq Q_t$.

$$(N_{t+1}, Q_{t+1}) = \begin{cases} (N_t - 1, Q_t - 1) & \text{with probability } 0.35, \\ (N_t - 1, Q_t) & \text{with probability } N_t(n - N_t)/\binom{n}{2} - 0.35, \\ (N_t, Q_t) & \text{with probability } [\binom{N_t}{2} + \binom{n-N_t}{2}]/\binom{n}{2} - 0.16, \\ (N_t, Q_t + 2) & \text{with probability } 0.16 - \binom{n-N_t}{2}/\binom{n}{2}, \\ (N_t + 2, Q_t + 2) & \text{with probability } \binom{n-N_t}{2}/\binom{n}{2}. \end{cases}$$

Note that the probabilities are all between 0 and 1 since $N_t \in I$. □

To finish, we just need one more lemma. We show that if Q_t is in the lower half of the interval I, then it is very likely to exit I by dropping below $0.61n$ (rather than by rising above $0.7n$).

Lemma 6 *Suppose that* $0.61n \leq Q_t \leq 0.65n$ *for some* t. *Define* T *by*

$$T = \min\{t' > t \mid Q_{t'} \notin I\}.$$

There exists a constant $d > 1$ *such that*

$$\text{Prob}\left[Q_T < 0.61n\right] \geq 1 - d^{-n}.$$

Proof.

Let $M = 3n$. For every $m \in [1, \ldots, M]$, let $x_m = Q_{t+m} - Q_{t+m-1}$ and $S_m = \sum_{i=1}^{m} x_i$. The random variables $\{x_m\}$ are independent of each other since, for all m, the variable x_m is fully determined by the random choice made by the Markov chain Q at time $t + m$. By definition, $Q_{t+m} = Q_t + S_m$. Note that $\mathbf{E}[S_m] = -0.03m$. Using the value of $\mathbf{E}[S_M]$, we find that $\text{Prob}(S_M \geq -0.04n)$ is equal to $\text{Prob}(S_M \geq \mathbf{E}[S_M] + 0.05n)$. By a Chernoff-Hoeffding bound, this is at most $\exp(-2(0.05n)^2/(9M))$, which is less than $(1/2)d^{-n}$ if d is chosen to be sufficiently close to 1. Similarly,

$$\begin{aligned} \sum_{m=1}^{M} \text{Prob}(S_m \geq 0.05n) &\leq \sum_{m=1}^{M} \text{Prob}(S_m \geq \mathbf{E}[S_m] + 0.05n) \\ &\leq \sum_{m=1}^{M} \exp(-2(.05n)^2/(9m)) \\ &< (1/2)d^{-n}, \end{aligned}$$

by choosing d even closer to 1, if necessary. Thus, with probability at least $1 - d^{-n}$, for every $m \in [1, \ldots, M]$ we have

$$Q_{t+m} = Q_t + S_m < 0.65n + 0.05n = 0.7n$$

and

$$Q_{t+M} = Q_t + S_M < 0.65n - 0.04n = 0.61n.$$

$\square$

To complete the proof of Theorem 2, note that every time the chain enters the interval I from below, the probability that it exits out the top of the region (rather than the bottom) is at most d^{-n}. Thus, the probability that the chain reaches absorbtion in as few as $((d+1)/2)^n$ visits to the region is at most

$$\left(\frac{d+1}{2d}\right)^n = o(1).$$

$\square$

5 Other topics

We have started the rigorous study of absorbtion time for the HCR rule of Shoham and Tennenholtz on classes of graphs and obtained rigorous confirmation for the experimental results in [4]. Several issues remain for further study. A natural extension of our results would be to investigate our dynamics for other classes of graphs. Two other cases seem particularly interesting, degree-bounded trees and random graphs. Another possible ingredient to our model is random noise (or player mistakes). The importance of this parameter has been previously recognized in evolutionary game theory, for example in [5].

References

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

[2] R. Axelrod, *The Complexity of Cooperation. Agent-Based Models of Competition and Cooperation*, Princeton Studies in Complexity, Princeton University Press, 1997.

[3] D. Fudenberg and D. K. Levine, *The Theory of Learning in Games*, M.I.T. Press, 1999.

[4] J. E. Kittock, *Emergent conventions and the structure of multi-agent systems*, in L. Nadel and D. Stein, eds., *1993 Lectures in Complex systems: the proceedings of the 1993 Complex systems summer school*, Santa Fe Institute Studies in the Sciences of Complexity Lecture Volume VI, Santa Fe Institute, Addison-Wesley, 1995.

[5] M. Kandori, G. J. Mailath and R. Rob, *Learning, Mutation and Long Run Equilibria in Games*, Econometrica 63 (1993), pp. 29-56.

[6] M. Nowak and K. Sigmund, *A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game*, Nature 364 (1993), pp. 56-58.

[7] T. Sandholm and R. Crites, *Multiagent Reinforcement Learning in the Iterated Prisoner's Dilemma.* Biosystems 37 (1995), pp. 147-166, Special Issue on the Prisoner's Dilemma.

[8] Y. Shoham and M. Tennenholtz, *Co-learning and the evolution of social activity*, Tech. Rep. CS-TR-94-1511, Computer Science Department, Stanford University, (1994).

[9] Y. Shoham and M. Tennenholtz, *On the emergence of social conventions: modelling, analysis and simulations*, Artificial Intelligence, 94 (1997), pp. 139–166.