

A Deterministic Protocol for Sequential Asymptotic Learning

Yu Cheng
Duke University
Durham, NC, USA
Email: yucheng@cs.duke.edu

Wade Hann-Caruthers and Omer Tamuz
California Institute of Technology
Pasadena, CA, USC
Email: {whanncar, omer tamuz}@gmail.com

Abstract—In the classic herding model, agents receive private signals about an underlying binary state of nature, and act sequentially to choose one of two possible actions, after observing the actions of their predecessors.

We investigate what types of behaviors lead to asymptotic learning, where agents will eventually converge to the right action in probability. It is known that for rational agents and bounded signals, there will not be asymptotic learning.

Does it help if the agents can be cooperative rather than act selfishly? This is simple to achieve if the agents are allowed to use randomized protocols. In this paper, we provide the first *deterministic* protocol under which asymptotic learning occurs. In addition, our protocol has the advantage of being much simpler than previous protocols.

I. INTRODUCTION

When making decisions, we often have some amount of information which we have learned on our own, but we also look to what choices other people have made in the past. In sequential learning models, there is a sequence of agents $\mathbb{N} = \{1, 2, \dots\}$ who are interested in learning an unknown state of the world $\theta \in \{0, 1\}$. Each agent i receives a signal s_i that depends on the state, takes an action $a_i \in \{0, 1\}$, and receives a utility of 1 if the action matches the state, and utility 0 otherwise. Each agent also observes the actions of her predecessors before choosing her own.

The classical result of [1], [2] is that from some point on, rational agents will disregard their own private signal and emulate the actions of their predecessors, resulting in a *herd* in which almost all the information is lost. In particular, *asymptotic learning* does not occur: the probability that the i^{th} agent chooses the correct action does not tend to one, as it would if the agents could observe the signals (rather than the actions) of their predecessors.

A natural question is whether non-selfish agents can coordinate on a protocol in which they do not always choose the action that is optimal given their information, but where asymptotic learning does occur. Achieving this with probabilistic protocols is straightforward [3], [4]: for example [3], one could have the i^{th} agent choose an action that reveals her private signal with probability $1/i$, and choose her optimal action with probability $1 - 1/i$.

In this paper, we show that asymptotic learning can also be achieved using a deterministic protocol.

Theorem I.1. *There is a deterministic protocol that achieves asymptotic learning.*

At the heart of the problem is a trade-off between optimization and communication. In order for asymptotic learning to occur, agents must rely more and more on the information provided by their predecessors' actions, since their own private signals have a uniformly bounded amount of information. However, if agents rely too much on previous observations, they will pass too little information on through their own actions. The question of whether there are any deterministic protocols under which asymptotic learning occurs is essentially the following: are there any deterministic protocols which can strike the right balance between agents taking optimal actions and agents using their actions to communicate their private information to their successors?

Our work raises some potentially interesting questions: in what reasonable economic settings do selfish agents naturally coordinate and achieve asymptotic learning in equilibrium? And how could one incentivize them to do so in other settings?

A. Related Work

The classical sequential learning papers [1], [2] have been followed by a large literature in Economics (see, e.g., [3], [5]–[7], as well as Chamley's book [8]) and engineering (see, e.g., [4], [9]). Smith and Sørensen [5] showed that when private signals are unbounded (i.e., there is no limit to how strong an indication a private signal can give) then agents do learn the state in the sequential setting; this observation has motivated us—and others—to understand under what other conditions learning can be achieved. As far as we know, [4] is the only other paper that studies related algorithmic questions. In particular, in [4] it is shown how agents with bounded memory who observe only some of their predecessors can achieve asymptotic learning using a randomized protocol.

Finally, in [10], the authors study a particular family of protocols involving randomness for achieving asymptotic learning, and demonstrate that the fastest rate of learning possible in this family is $O(1/n)$. These protocols differ in two important ways from the protocol presented here. First, the protocol itself is not deterministic; each agent uses a randomization device to decide whether or not to reveal her signal. Second, the devices used by different individuals are independent, so that whether

illustrated in Figure 1. We define the strategy of each agent depending on if she is revealing or not. For a given i , let k be the unique integer with $2^{k-1} \leq i < 2^k$. Then

$$f_i(a_1, \dots, a_{i-1}, s_i) = \begin{cases} g_k(a_{t_1}, \dots, a_{t_{k-1}}, s_i) & \text{if } i \neq t_k, \\ s_i & \text{if } i = t_k \end{cases}$$

where

$$g_k(x_1, \dots, x_k) = \begin{cases} 0 & \text{if } \frac{1}{k} \sum_{j=1}^k x_j \leq \frac{q_0 + q_1}{2}, \\ 1 & \text{if } \frac{1}{k} \sum_{j=1}^k x_j > \frac{q_0 + q_1}{2}. \end{cases}$$

Theorem I.1 is a direct corollary of Lemma III.1.

Lemma III.1. *Let q_0, q_1 denote the parameters of the Bernoulli distributions D_0 and D_1 respectively. We assume that $\epsilon < q_0 < q_1 < 1 - \epsilon$ and $q_1 - q_0 > 2\epsilon$ for some constant $\epsilon > 0$. The protocol $\{f_i\}_{i \in \mathbb{N}}$ defined above satisfies $p_n \geq 1 - 2n^{-\epsilon^2}$ for all $n \in \mathbb{N}$.*

Lemma III.1 follows immediately from Lemmas III.2 and III.3. We defer the proof of Lemma III.1 to the end of this section.

Lemma III.2. *The probability of agent n being a revealing agent is at most $n^{-\epsilon}$.*

Proof. Fix an index i with $2^{k-1} \leq i < 2^k$, i.e., agent i is on the k th level of the binary tree. If agent i is selected as a revealing agent, we must have $t_k = i$. This happens if and only if the private signals of the first $k-1$ revealing agents form the binary representation of the number $(i - 2^{k-1})$. In other words, there is a unique path from the root of the tree to agent i and we must always take the correct edges to reach i . Observe that the private signals of the first $k-1$ revealing agents are drawn i.i.d. from D_θ . Since $\epsilon < q_0 < q_1 < 1 - \epsilon$, each private signal matches the binary representation of i with probability at most $1 - \epsilon$. Therefore, we conclude that the probability agent n with $n \geq 2^{k-1}$ is on the paths is at most

$$(1-\epsilon)^{k-1} \leq e^{-\epsilon(k-1)} = (2^{k-1})^{-\epsilon \log_2(e)} \leq n^{-\epsilon \log_2(e)} \leq n^{-\epsilon}.$$

where first step uses that $1 - x \leq e^{-x}$ for all $x \in \mathbb{R}$. \square

Lemma III.3. *When agent n is not a revealing agent, she answers correctly with probability at least $1 - n^{-\epsilon^2}$.*

Proof. Fix an index n with $2^{k-1} \leq n < 2^k$. Agent n has access to the actions of the revealing agents t_1, \dots, t_{k-1} since $t_k < 2^{k-1}$. Because these agents reveal their private signals, agent i has at least k i.i.d. samples from D_θ (including her own signal). She can simply take the average of these samples and check if the empirical mean is closer to q_0 or q_1 to guess the hidden state θ . Note that most of the agents are not revealing, and they are acting rationally.

Let $\bar{q} = \frac{q_0 + q_1}{2}$. Since we assumed that $q_1 - q_0 \geq 2\epsilon$, we have that $|\bar{q} - q_\theta| \geq \epsilon$. Therefore, as long as the empirical mean of the k samples has additive error less than ϵ , agent i

will be able to guess θ correctly. By standard application of Chernoff-Hoeffding bounds, we have

$$\Pr \left[\left| \frac{1}{k} \left(\sum_{j=1}^{k-1} a_{t_j} + s_i \right) - q_\theta \right| \geq \epsilon \right] \leq e^{-2k\epsilon^2} \leq n^{-\epsilon^2}. \quad \square$$

We are now ready to prove Lemma III.1.

Proof of Lemma III.1. We prove the protocol $\{f_i\}_{i \in \mathbb{N}}$ defined as above satisfies $p_n \geq 1 - 2n^{-\epsilon^2}$. The claim then follows from the assumption that $\epsilon > 0$ is a constant.

Agent n can pick the wrong action due to one of the two reasons:

- 1) Agent n is revealing, and her private signal is different from θ . The probability of this event is upper bounded by the probability that agent n is revealing, which is $n^{-\epsilon}$ by Lemma III.2.
- 2) The agent is not revealing, but the previous information leads to the wrong conclusion on θ . By Lemma III.3, this happens with probability at most $n^{-\epsilon^2}$.

We conclude the proof by taking a union bound over these two cases. \square

IV. AN IMPROVED PROTOCOL

In this section we describe a slightly more complicated deterministic protocol which achieves the theoretically optimal learning rate for random protocols from [10]. This protocol shares the essential structure with the previous protocol, but is slightly modified to remove some of the inefficiencies in the previous protocol and allow for an improvement on the asymptotic rate.

The new protocol will differ from the previous protocol in the following ways.

First, rather than a single agent revealing from each group, k agents from the interval $2^{k-1}, \dots, 2^k - 1$ will reveal. The revealing agents will be $t_k, t_k + 1, \dots, t_k + k - 1$ if $t_k + k - 1 < 2^k$ and $t_k, \dots, 2^k - 1, 2^{k-1}, \dots, 2^{k-1} + k - 1 - (2^k - t_k)$ otherwise. Intuitively, the revealing agents will be the block of k agents starting with t_k if the agents $2^{k-1}, \dots, 2^k - 1$ are arranged cyclically.

Similar to the previous protocol, t_k will be calculated from t_{k-1} according to the recursive formula

$$t_k = t_{k-1} - 2^{(k-1)-1} + b_{k-1} + 2^{k-1}$$

where b_k is the parity of the number of revealing agents in the k^{th} block that take action 1.

As before, all agents that are not revealing can recover which agents before them were revealing, and will take the action corresponding to whichever signal has appeared more often among their own signal and the signals of the revealing agents.

Proposition IV.1. *Under this protocol, $p_n = O(\log n)/n$.*

Proof. As in the proof of Lemma III.1, we will bound the probability of taking the wrong action by separately bounding the probability of not being revealing and taking the wrong action and the probability of being revealing and taking the wrong action.

Fix ϵ as in Lemma III.1.

Consider agent n , where $2^{k-1} \leq n < 2^k$. Suppose this agent is not revealing. Then she will, at a minimum, see the signals of the $1 + \dots + k - 2 = \frac{(k-1)(k-2)}{2}$ previous revealing agents, in addition to her own signal. Hence, by a Chernoff-Hoeffding bound (as in the proof of Lemma III.1), the probability that she takes the wrong action will be bounded above (for sufficiently large k) by:

$$\begin{aligned} \exp(-(k-1)(k-2)\epsilon^2) &\leq \exp\left(-\frac{1}{4}k^2\epsilon^2\right) \\ &\leq \exp\left(-\frac{1}{8}(\log(n))^2\epsilon^2\right) \\ &= 2^{(\log(\epsilon))(-\frac{1}{8}(\log(n))^2\epsilon^2)} \\ &= \left(\frac{1}{n}\right)^{\left(\frac{1}{8}\epsilon^2 \log(\epsilon)\right) \log(n)}. \end{aligned}$$

Hence, this gives an upper bound for the probability that agent n is not revealing and makes a mistake. Now, to bound the probability that agent n is revealing and makes a mistake, it will be sufficient to bound the probability that agent n is revealing.

First, note that the probability that b_k is 1 converges exponentially (in k) to $\frac{1}{2}$. Hence, a simple calculation shows that the probability of that $t_k = n$ is bounded above by d/n for some fixed constant d . Hence, the probability that agent n is revealing is bounded above by $kd/n = d \log(n)/n$.

Thus, the probability that agent n takes the wrong action is bounded above by

$$\left(\frac{1}{n}\right)^{\left(\frac{1}{8}\epsilon^2 \log(\epsilon)\right) \log(n)} + d \log(n)/n$$

which is bounded above by

$$2d \log(n)/n$$

for all sufficiently large n . \square

ACKNOWLEDGMENT

Part of this work was done while Yu Cheng was a student at the University of Southern California. Yu Cheng was supported in part by Shang-Hua Teng's Simons Investigator Award. Omer Tamuz was supported in part by grant #419427 from the Simons Foundation.

REFERENCES

- [1] A. V. Banerjee, "A simple model of herd behavior," *The Quarterly Journal of Economics*, pp. 797–817, 1992.
- [2] S. Bikhchandani, D. Hirshleifer, and I. Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of political Economy*, pp. 992–1026, 1992.
- [3] D. Acemoglu, M. A. Dahleh, I. Lobel, and A. Ozdaglar, "Bayesian learning in social networks," *The Review of Economic Studies*, vol. 78, no. 4, pp. 1201–1236, 2011.
- [4] K. Drakopoulos, A. Ozdaglar, and J. N. Tsitsiklis, "On learning with finite memory," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6859–6872, 2013.
- [5] L. Smith and P. Sørensen, "Pathological outcomes of observational learning," *Econometrica*, vol. 68, no. 2, pp. 371–398, 2000.
- [6] E. Eyster and M. Rabin, "Naive herding in rich-information settings," *American economic journal: microeconomics*, vol. 2, no. 4, pp. 221–243, 2010.
- [7] W. Hann-Caruthers, V. V. Martynov, and O. Tamuz, "The speed of sequential asymptotic learning," 2017.
- [8] C. Chamley, *Rational herds: Economic models of social learning*. Cambridge University Press, 2004.
- [9] I. Lobel, D. Acemoglu, M. Dahleh, and A. Ozdaglar, "Rate of convergence of learning in social networks," in *Proceedings of the American Control Conference*, 2009.
- [10] Y. Peres, M. Z. Racz, A. Sly, and I. Stuhl, "How fragile are information cascades?" *arXiv preprint arXiv:1711.04024*, 2017.