Improving Quantal Cognitive Hierarchy Model Through Iterative Population Learning

Extended Abstract

Yuhong Xu
School of Computing and Information Systems
Singapore Management University
Singapore
yuhongxu@smu.edu.sg

Shih-Fen Cheng
School of Computing and Information Systems
Singapore Management University
Singapore
sfcheng@smu.edu.sg

Xinyu Chen
School of Computing and Information Systems
Singapore Management University
Singapore
xychen@smu.edu.sg

ABSTRACT

In this paper, we propose to enhance the state-of-the-art quantal cognitive hierarchy (QCH) model with iterative population learning (IPL) to estimate the empirical distribution of agents' reasoning levels and fit human agents' behavioral data. We apply our approach to a real-world dataset from the Swedish lowest unique positive integer (LUPI) game and show that our proposed approach outperforms the theoretical Poisson Nash equilibrium predictions and the QCH approach by 49.8% and 46.6% in Wasserstein distance respectively. Our approach also allows us to explicitly measure an agent's reasoning level distribution, which is not previously possible.

KEYWORDS

behavioral game theory; cognitive hierarchy model; quantal cognitive hierarchy model; lowest unique positive integer game

ACM Reference Format:
Yuhong Xu, Shih-Fen Cheng, and Xinyu Chen. 2023. Improving Quantal Cognitive Hierarchy Model Through Iterative Population Learning: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

When a strategic scenario is modeled using game theory, Nash equilibrium is a well-known solution concept for predicting how agents would behave. However, when agents are only partially rational, Nash equilibrium would provide poor outcome predictions [2]. To address this issue, the field of behavioral game theory aims to explicitly model human agents' limited rationality, and come up with refined equilibrium concepts that would be more suitable for games with human agents.

One popular behavioral game-theoretic model is the "cognitive hierarchy" (CH) framework introduced by Camerer et al. [1], which allows us to explicitly specify different rationality levels for agents in a game. In the CH framework, non-strategic agents are regarded as level-0, and their strategies are generated irrespective of other agents (e.g., uniformly randomly or greedily). For strategic agents at level-k ($k \geq 1$), they would assume that other agents would be behaving at level-j, where $j < k$, and compute best responses. Camerer et al. [1] assume that an agent with level-k believes that its opponent's reasoning levels would follow a Poisson distribution with parameter τ between 0 and $k-1$. Based on the CH model, the quantal CH (QCH) model allows human decision errors to be captured by a soft-max function and is shown to be the state of the art in matching data involving human subjects [5, 6].

In this paper, we further improve the QCH model by introducing Iterative Population Learning (IPL) for learning the reasoning level distribution (previously assumed to be a Poisson distribution), which plays a vital role in computing quantal best responses. We formally define the determination of the reasoning level distribution as a fixed point-seeking problem and prove that a fixed point exists in the population dynamics that we define. We then propose an iterative process that could efficiently identify a stable population reasoning level distribution. Finally, we test our approach on the lowest unique positive integer game dataset collected from a series of laboratory experiments. We demonstrate that (1) Our model can fit agents' behavioral traces much better (in Wasserstein distance, our approach outperforms the QCH model by close to 50%). (2) Our approach is capable of estimating individual agents' reasoning level distributions. (3) Combined with personal-level observations, our approach is shown to be more accurate in measuring the strategic reasoning level of the population.

2 ITERATIVE POPULATION LEARNING

Let $L = \{0, 1, \ldots, m\}$ be the set of reasoning levels to be considered. Let $p = (p_0, \ldots, p_m) = \Delta(L)$ be the reasoning level estimation at the population level, and $p_{l,i}$ be the probability that agent i belonging to level l. Naturally, $\sum_{l \in L} p_{l,i} = 1, \forall i \in N$. The QCH-IPL approach is defined by the iterative steps below:

(1) **Initialization**: Let $t \leftarrow 1$ be the current iteration. Initialize $(p_0^t, \ldots, p_m^t)$ uniformly randomly.
(2) **Compute QBRs for all levels**: Assume that the game is symmetric; let A_0 be the common action space and π_k^t be the QBR for level-k agent in iteration t.

$$\pi_0^{t+1}(a_0) = |A_0|^{-1}, \forall a_0 \in A_0;\ \pi_{0:k}^{t+1} = \frac{\sum_{l=0}^{k} p_l^t \pi_l^{t+1}}{\sum_{l'=0}^{k} p_{l'}^t}, \forall k \in L \quad (1)$$

$$\pi_k^{t+1} = QBR\left(\pi_{-i,0:k-1}^{t+1}; \lambda\right), \forall k = 1, \ldots, m. \quad (2)$$

(3) **Fit each agent's behaviors to QBRs**:

$$\left(p_{0,i}^{t+1}, \ldots, p_{m,i}^{t+1}\right) = CLR\left((\pi_0^{t+1}, \ldots, \pi_m^{t+1}), tr_i\right), \forall i \in N, \quad (3)$$

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

where agent i's behavioral traces are denoted as tr_i and CLR refers to *constrained linear regression*. CLR is a regular linear regression with tr_i as the dependent variable, $\pi_0^{t+1}, \ldots, \pi_m^{t+1}$ as independent variables, but with constraints that all coefficients are non-negative and sum to 1.

(4) **Aggregation**: We compute reasoning level estimation for the whole agent population by simple average:

$$p_l^{t+1} = \sum_{i \in N} p_{l,i}^{t+1}/n, \forall l \in L. \tag{4}$$

(5) **Checking for convergence**: If

$$\left\|(p_0^{t+1}, \ldots, p_m^{t+1}) - (p_0^t, \ldots, p_m^t)\right\|_2 < \epsilon,$$

terminate; otherwise, $t \longleftarrow t+1$, repeat from 2.

By combining all the steps above, we have:

$$(\pi_0, \ldots, \pi_m) = F\Big(G\Big(\big(CLR\left((\pi_0, \ldots, \pi_m), tr_i\right)\big)_{i \in N}\Big)\Big), \tag{5}$$

where $G(\cdot)$ is the aggregation function in Step (4), and $F(\cdot)$ is the computation of QBR $(\pi_0, \ldots, \pi_m)$ in Step (2). From above we can see that $(\pi_0, \ldots, \pi_m)$ is a fixed point of (5). We can further prove that a fixed point always exists for (5).

3 NUMERICAL RESULTS

We test our QCH-IPL approach in a lab LUPI game designed and executed by Östling et al. [5]. The lab LUPI is a Poisson game [3, 4] with the following rules: (1) the number of active players in each round follows a Poisson distribution with a mean of 26.9, (2) players (both active and inactive) are required to choose an integer between 1 and 99, (3) the smallest number that is uniquely chosen wins the game round. Participants are compensated to play 49 rounds (each round is considered a *day*, and we group every 7 days as a *week*). For each round, the winner (if any) earns an extra bonus. The winning number is made known to all participants after each round concludes, but the winner is notified privately. There are in total 152 participants in the lab LUPI game.

We compare our model with two baselines: the Poisson-Nash Equilibrium (PNE) and the QCH model, both described in detail by Östling et al. [5]. The QCH model is constructed with the assumption that the agent's reasoning levels follow a Poisson distribution with parameter τ estimated from a similar field game, and Östling et al. [5] propose to search for λ using 100,000 search grids, with the maximum likelihood estimation (MLE) as the selection criterion.

When setting up our QCH-IPL approach, we also search for λ using MLE, and we execute the main algorithm with up to 100 iterations. We return the model with the best log-likelihood value at the end of our search.

To evaluate the performance of competing models, we look at: (1) χ^2 goodness of fit test, on whether we can reject the null hypothesis that the target distribution follows the given distribution (significance is bad), (2) the log-likelihood (the larger the better), (3) proportion below, which measures the percentage of target distribution that is covered by the given distribution (the larger the better), and (4) the Wasserstein distance, which is a popular metric that measures the effort to transform the target distribution into the given distribution (the smaller the better).

For each *week* of the lab experiment, we have an independent performance evaluation. In the interest of space, we only list the results from weeks 1, 4, and 7 in Table 1. On "χ^2 test", PNE is rejected as a good fit for all weeks, QCH is rejected in weeks 1, 2, and 7, while only week 1 is rejected for QCH-IPL. On "log-likelihood", QCH-IPL outperforms QCH in all weeks. On "proportion below", QCH-IPL outperforms both PNE and QCH over all weeks, and on average the advantage margins are 7% and 4.6% respectively. On the "W-distance", QCH-IPL also achieves the best (lowest) values for all weeks, and on average the advantage margins against PNE and QCH are 49.8% and 46.6% respectively.

Table 1: Goodness-of-fit for the lab LUPI game: Comparing Poisson-Nash equilibrium, QCH model, and QCH-IPL model.

Week	(1)	(4)	(7)
Poisson-Nash equilibrium			
χ^2 (for average frequency)	24.7***	21.8**	21.4***
Proportion below (percent)	82.25	88.64	87.06
Wasserstein distance	3.362	1.323	0.896
QCH Model			
Log-likelihood	−210.4	−88.7	−99.4
χ^2 (for average frequency)	24.3***	4.6	10.1*
Proportion below (percent)	84.62	92.54	91.07
Wasserstein distance	4.972	0.688	0.590
QCH-IPL Model			
Log-likelihood	-185.0	-78.2	-85.3
χ^2 (for average frequency)	16.7***	0.6	3.3
Proportion below (percent)	86.58	95.19	93.70
Wasserstein distance	2.87	0.432	0.435

Note: Significance level: * 1%. ** 5%. * 10%.**

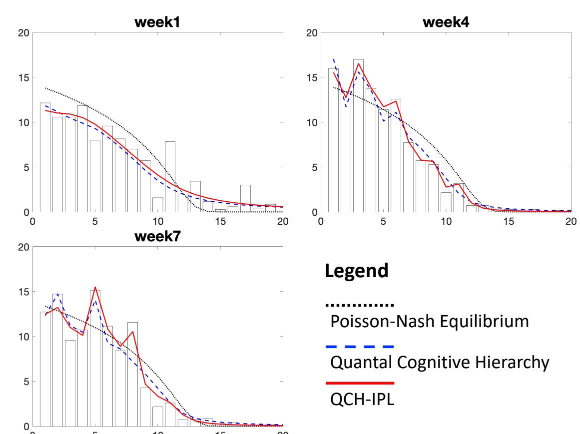

Figure 1: Behavioral traces against PNE, QCH, and QCH-IPL.

Finally, we also visualize players' actual action frequencies and all behavioral models from selected weeks in Figure 1 (as numbers above 20 have low usage frequencies, we cut the x-axis off at 20). From Figure 1 we can see that the PNE is mostly smooth, thus not able to capture discontinuities in players' action frequencies. Both QCH and QCH-IPL models respond to discontinuities, but QCH-IPL offers visibly better fits than the QCH model.

ACKNOWLEDGEMENT

This research is supported by the Ministry of Education, Singapore, under its Social Science Research Thematic Grant (Grant Number MOE2020-SSRTG-018).

REFERENCES

[1] Colin F Camerer, Teck-Hua Ho, and Juin-Kuan Chong. 2004. A cognitive hierarchy model of games. *The Quarterly Journal of Economics* 119, 3 (2004), 861–898.

[2] Jacob K Goeree and Charles A Holt. 2001. Ten little treasures of game theory and ten intuitive contradictions. *American Economic Review* 91, 5 (2001), 1402–1422.

[3] Roger B Myerson. 1998. Population uncertainty and Poisson games. *International Journal of Game Theory* 27, 3 (1998), 375–392.

[4] Roger B Myerson. 2000. Large Poisson games. *Journal of Economic Theory* 94, 1 (2000), 7–45.

[5] Robert Östling, Joseph Tao-yi Wang, Eileen Y Chou, and Colin F Camerer. 2011. Testing game theory in the field: Swedish LUPI lottery games. *American Economic Journal: Microeconomics* 3, 3 (2011), 1–33.

[6] James R Wright and Kevin Leyton-Brown. 2017. Predicting human behavior in unrepeated, simultaneous-move games. *Games and Economic Behavior* 106 (2017), 16–37.