

Choose Early or Choose Wisely - A Chinese Restaurant Game Approach

Fang-Li Kung

Research Center for Information Technology Innovation,
Academia Sinica, Taipei, Taiwan
marykv1@citi.sinica.edu.tw

Chih-Yu Wang

Research Center for Information Technology Innovation,
Academia Sinica, Taipei, Taiwan
cywang@citi.sinica.edu.tw

ABSTRACT

Agents in a network often face situations requiring them to make decisions without sufficient information. In such situations, they may postpone their decisions in order to observe and collect more information through learning from other agents. In this paper, we discuss the advantages of the postponement strategy from a game-theoretic perspective. We propose an extension to Chinese Restaurant Game, a general framework for social learning. In the proposed extension, rational agents may change their decision order at will. We find that two important elements in Chinese Restaurant Game, social learning and negative network externality, still dominate agents' decision process and the postponement strategy. We study a two-player case in detail. Through simulations, we find that the signal quality and table size ratio greatly influence whether a rational agent will apply the postponement strategy or not. In some cases, rational agents may postpone their decisions in response to some, but not all, signals they received. We observe that such a strategy is informative, which also helps other agents improve their strategies accordingly.

Categories and Subject Descriptors

J.4 [Social and Behavioral Sciences]: Sociology

General Terms

Algorithms, Theory.

Keywords

Chinese Restaurant Game, Game Theory, Social Learning

1. INTRODUCTION

Rational agents in a social network may encounter incidents that require them to make decision without complete information. In such occasions, they will learn the required knowledge from external information sources. Signals revealed by other agents, such as chats, reviews, rating, etc., are the main sources of information for agents to learn from [1-4]. Sometimes agents require a period of time to collect and analyze these signals, which force them to postpone their decisions. The action of postponing the decisions, or "wait and see" strategy, is widely seen in the real world especially when the outcome is highly uncertain. Nevertheless, the action could be costly to the agents when they are in a competition with others. How rational agents balance the accuracy and delay of the decision is the main focus of this paper.

Chinese restaurant game is a general framework for modeling the individual decision in a network with negative network externality [5-7]. The original framework addresses two different effects:

advantages of learning from other agents' experiences and disadvantages of competition from others. Nevertheless, the decision order of agents in legacy Chinese restaurant game is fixed, which means that agents cannot postpone their decisions at will. In light of this, we extend the legacy Chinese restaurant game to support the action of postponing, that is, agents or customers may postpone their decisions at will in exchange for collecting more informative signals. Such an extension significantly expands the degree of freedom of agent's strategy and makes the model more suitable for real world applications.

In this paper, we introduce an extension to Chinese Restaurant Game to allow customers to postpone their decisions. We present a detailed discussion on the two-player case where two customers and two tables in the restaurant are considered. Specifically, we illustrate how the Nash equilibrium of the proposed game can be classified into four types. We propose an iterative-update algorithm to identify the pure strategy Nash equilibrium. We demonstrate the critical influence on the decision order and how rational agents may choose when the order can be altered at will. We discuss the influences through simulations, where we show how the table size and signal quality affect the final decision of each agent in the game.

2. SYSTEM MODEL

Let us consider a Chinese restaurant with a set of tables and customers. Each table has infinite seats in different size, and customers each request a table for having a meal in a sequential order. We model the table selection problem as a game, which is similar to the original Chinese Restaurant Game [6]. Specifically, the table set is $X = \{1, \dots, K\}$, which is also the table selection action set of each customer. The $x_i \in X$ is the choice of customer i . Then, let the utility function of customer i be $U(R_{x_i}(\theta), n_{x_i})$, where n_{x_i} is the number of customers choosing table x_i at the end of the game, and $R_{x_i}(\theta)$ is the size of the x_i table. We let the state θ be an objective parameter, which means the state θ changes when the restaurant is remodeled, and the table size function $R_j(\theta)$ be fixed. Finally, we denote the number of customers in table i as n_i and $n = \{n_1, n_2, \dots, n_k\}$ as the grouping, i.e., the grouping of customers in the restaurant [5].

We assume that customers do not know the exact state θ , and they need to collect information, such as reviews and recommendations, to help them identify the true state of the restaurant. The information that customers gather in process is called signals. We assume that all customers know the prior distribution of the state information θ , which is denoted as $g_0 = \{g_{0,l} | g_{0,l} = \Pr(\theta = l, \forall l \in \Theta)\}$; the predefined distribution $f(s|\theta)$ produces the signal $s_i \in S$ each customer received.

A rational customer can estimate the system state θ by his belief [6]. We assume all customers choose tables sequentially. Each customer reveals his signal to others after he chooses a table. Let

us denote the signals customer i learned, excluding his own signal s_i , as $h_i = \{s\}$. Customer i can estimate the current system state in probability with the belief being defined as

$$g_i(h_i, s_i) = \{g_{i,l}(h_i, s_i) = \Pr(\theta = l | h_i, s_i, g_0) | \forall l \in \Theta\}$$

According to the above definition, $g_{i,l}(h_i, s_i)$ represents the probability that system state θ is equal to l conditioning on the collected signals h_i , received signal s_i , the prior probability g_0 , and the conditional distribution $f(s|\theta)$. Since we assume that customers are fully rational and follow the Bayesian rule, they will update their beliefs by the Bayesian rule as follows [6,7]:

$$g_{i,l}(h_i, s_i) = \frac{g_{0,l} \Pr(h_i, s_i | \theta = l)}{\sum_{l' \in \Theta} g_{0,l'} \Pr(h_i, s_i | \theta = l')}$$

In the paper, we propose an extension to Chinese Restaurant Game by giving each customer in the line a chance to change their order of making decision. Specifically, a customer, except the customer N , may voluntarily leave and rejoin the line as the last customer **after** he received his signal and **right before** he is asked to select a table. In that case, he gives up the chance to choose a table earlier in exchange for more signals revealed by other customers. Mathematically speaking, the action of customer i is denoted by $a_i = (o_i, x_i) \in \{0,1\} \times X$, $i < N$, where $o_i \in \{0,1\}$ represents whether customer i chooses to postpone his decision and rejoin the line when he is asked to select a table. The observed signals h_i will also change accordingly.

A rational customer will predict the expected utility he may gain if he chooses to postpone his decision given his received signals h_i , his own signal s_i , the decisions of previous customers, and the **expected reactions** of other customers. He will choose to make decision later if he believes that making decision later will lead him to choose the right table and maximize his utility. We assume that the customers who make decision after him have full knowledge about the decision the customer made before them.

3. EQUILIBRIUM ANALYSIS

As mentioned above, the customers are rational and follow Bayesian learning rule. Considering the information they collect, customers will choose tables and decide **whether to postpone their decisions** so that they can maximize their own expected utility. We first represent $n_i = (n_{i,1}, n_{i,2}, \dots, n_{i,j})$ as the current grouping observed by customer i before making decision, where $n_{i,j}$ is the number of customers choosing table j before customer i . Then, let $h_i = \{s_1, s_2, \dots, s_i\}$ be the history of revealed signals before customer i . In such a case, the best response of customer i can be written as follows [6]:

$$BE_i(n_i, h_i, s_i) = \arg \max_{\{o_i, x_i\}} E[U(R_{x_i}(\theta), n_{x_i}) | n_i, h_i, s_i]. \quad (1)$$

In [6], a recursive approach is proposed to compute the expected utility (1) under the assumption that the decision order is intact, which is not applicable here since customers may choose to postpone their decisions and rejoin the line if it is beneficial.

Due to page limitation, a two-player case is demonstrated. We will show that even under this simple setting we can derive interesting insights from the game. We consider a two-player game where there are two customers A and B, and two tables $\{1,2\}$ in a Chinese restaurant under a system state θ . Without losing generality, customer A chooses first and customer B chooses later. Nevertheless, customer A has a chance to postpone his decision after receiving his own signal. In this regards, the action of customer A and B are $\{o_A, x_A\}$ and $\{x_B\}$, respectively. Both

customers aim at choosing a table that maximizes their own utility. The game ends when both customers choose a table. Then, the utility of the customer i choosing table j is given by $U(R_j(\theta), n_j)$, where n_j is the number of customers choosing table j at the end.

3.1 Iterative Solution for Two-Player Case

We propose an iterative-update method to derive the Nash equilibrium in the proposed game model. The key concept is to update the best response function of each customer iteratively. The Nash equilibrium is derived when the algorithm terminates.

Initialization: We first initialize the best response function of each customer using the recursive method proposed in [6] with customer A as the first customer while the customer B as the second customer. Here we denote the best response function of the first and second customer as $BE_1(s_1)$ and $BE_2(n_2, h_2, s_2)$, respectively. Notice that it is not necessary for the first customer to be customer A, and vice versa. Since we have only two customers, the best response function of the second customer can be simplified as $BE_2(x_1, s_1, s_2)$. We then have $BE_{1,A}(s_1) = BE_{1,B}(s_1) = BE_1(s_1)$, where the first two notations denote the best response of customer A and B if they are the first customer to choose a table. Finally, we denote $O(s_A)$ as the postponement strategy. In each iteration, two stages are applied.

1st Stage: In the first stage, customer A compares the expected utility if he remains the order and the expected utility if he changes to the position of customer B under the assumption that the strategy applied by customer B remains $BE_{1,B}(s_B)$ conditioning on the signal s_A he received. The expected utility of both customers can be derived by the recursive approach in [6] as follows. The expected utility of customer A is

$$E[U_A | s_A] = \sum_{s_B \in S} \Pr(s_B | s_A) E[U(R_j(\theta), n_j) | n_2, s_A, s_B] \Big|_{j=BE_{1,A}(s_A)}$$

The expected utility of customer A postponing his decision under the assumption that customer B follows the same table selection strategy as him, conditioning on his received signal s_A , is

$$\begin{aligned} E[U_A^* | s_A] \\ = \sum_{s_B \in S} \Pr(s_B | s_A) E[U(R_j(\theta), n_j) | n_2, s_A, s_B] \Big|_{\substack{j=BE_2(x_B, s_B, s_A) \\ x_B=BE_{1,B}(s_B)}} \end{aligned}$$

where n_B^* is the grouping after customer B selects a table according to $BE_1(s_B)$.

Notice here that the best response of customer A and B interchange when their order exchanges. In such a case, customer A receives both the signals s_A and s_B due to the postponement. In sum, the postponement strategy of customer A would be:

$$O(s_A) = \begin{cases} 1, & \text{if } E[U_A | s_A] \leq E[U_A^* | s_A] \\ 0, & \text{otherwise.} \end{cases}$$

We denote $S_o = \{s_A | O(s_A) = 1\}$ as the signals of postponing.

2nd Stage: In this stage, customer B has a chance to alter his strategy if he is forced to select a table first. In the previous stage, customer A might choose to make decision later when receiving any or a subset of signals. There are three possible cases:

Case 1: customer A does not postpone his decision given any signal. In such a case, customer B follows his original best response function.

Case 2: customer A postpones his decision given any signal; the postponement reveals no new information to customer B. In such

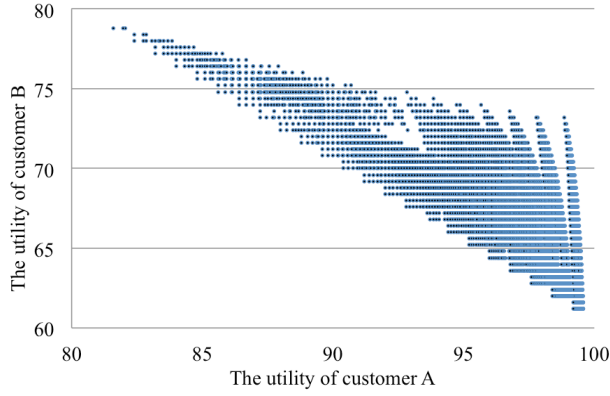


Figure 1: Utility of customers at $r = 0.6$

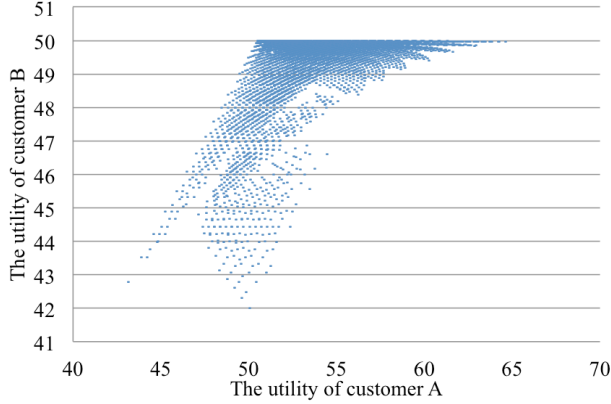


Figure 2: Utility of customers at $r = 0$

a case, customer B has no choice but to follow the same best response function he originally applies.

Case 3: customer A postpones his decision only when a subset of signals is received. In such a case, the action of postponing the decision itself reveals new information to customer B: the customer A must receive a subset of signals. In such a case, customer B may make use of this information to alter his strategy when he is forced to choose first.

For customer B, the knowledge that customer A postponed only when receiving a certain subset of signals can be used to improve 1) his belief on the system state, and 2) his prediction on customer A's choice on a table. Specifically, the postponement event suggests that customer A received signal $s_A \in S_o$, which reduces the signal space the customer B must be aware of. Although the customer B may not know the exact signal customer A received, he can use Bayes' theorem to estimate the possibility of each signal the customer A may receive and update his expected utility:

$$E[U(R_j(\theta), n_j) | s_B] = \sum_{l \in \theta} \sum_{s_A \in S_o} g_{1,l}(s_A, s_B) U(R_j(l), m_j + 1)$$

where $m_j = 1$ if $BE_2(j, s_A, s_B) = j$, and $m_j = 0$ otherwise.

The best response of customer B can then be updated as follows:

$$BE_{1,B}(s_B) = \arg \max_j E[U(R_j(\theta), n_j) | s_B]$$

Termination: The proposed method terminates when both $BE_1(s_1)$ and $BE_2(n_2, h_2, s_2)$ remain unchanged after an iteration. When this condition is met, a Nash equilibrium is reached. We output $BE_1(s_1)$, $BE_2(n_2, h_2, s_2)$, and the postponement strategy.

4. SIMULATION RESULTS

We now study the proposed two-player case through simulations. In this section, we consider two customers A and B, and two tables $\{1,2\}$ in a Chinese restaurant with two possible states $\theta \in \{1,2\}$. The size of tables is given by $R_i(\theta)$, where i is the number of table and θ is the state of the restaurant; $R_i(\theta)$ is unknown to all customers. When $\theta = 1$, the size of table 1 is $R_1(1) = 100$ and the size of table 2 is $R_2(1) = 100r$, where r is the ratio of table size and $r \in [0,1]$. When $\theta = 2$, $R_1(2) = 100r$ and $R_2(2) = 100$. The prior distribution of the state information is randomly chosen with $Pr(\theta = 1) = Pr(\theta = 2) = 0.5$. Customers each receive a randomly generated signal $s_i \in S = \{-2, -1, 1, 2\}$. The signal distribution $f(s|\theta)$ is given by

$$Pr(s = 2|\theta = 1) = Pr(s = -2|\theta = 2) = p_1$$

$$Pr(s = 1|\theta = 1) = Pr(s = -1|\theta = 2) = p_2$$

$$Pr(s = -1|\theta = 1) = Pr(s = 1|\theta = 2) = p_3$$

$$Pr(s = -2|\theta = 1) = Pr(s = 2|\theta = 2) = p_4$$

where $p_1 + p_2 + p_3 + p_4 = 1$ and $1 > p_1 > p_2 > p_3 > p_4 > 0$ can be regarded as the quality of signals. When the quality of signals is higher, the corresponding signal is more likely to reflect the true state θ . The utility of the customer i choosing table j is given by $(R_j(\theta), n_j) = R_j/n_j$, where n_j is the number of customers choosing table j at the end of the game. We investigate how the quality of signals influences the decision order and the utility of customers, and how the ratio of table size affects the decision order and the utility of customers.

First, we show the results of two settings: table size ratio $r = 0.6$ and $r = 0$. When $r = 0.6$, the second table size is 60. The utility of customers under the different qualities of signals is illustrated in Figure 1. When customer A always decides to choose first whatever the signals he receives, he has significantly higher expected utility than customer B has. In this case, the network externality effect gives the first customer more advantage than the learning process gives.

When $r = 0$, the second table size is 0, which means that both of customers want to choose the right table. The result is different from the result of original Chinese Restaurant Game, where the first customer should have lower utility. From Figure 2, we can see that the utility customer A still has higher utility than customer B has. It shows that offering the option to postpone his strategy changes the outcome and gives the first customer significantly larger advantage.

In detail, we list the Nash equilibrium of the game given the received signals when $r = 0$ in Figure 3(a) and 3(b). There are four possible pure Nash Equilibrium. Type 1 and 4 represent individually the case that customer A always chooses first and always chooses later. Type 2 and Type 3 represent that customer A may postpone his decision only when some kind of signals are received, but the response of customer B may be different, respectively. Figure 3(a) demonstrates an example of Type 1, that is, customer A always chooses to make decision first. When signal quality p_1 is low ($p_1 < 0.3$), customer A follows the signal he received to make decision first and choose the expected bigger table. Customer B, as the second customer, will choose the table that is still empty. In this case, the low quality of the signal leads customer B to believe that the expected utility of both tables is similar. Given that the signal quality is low, the benefit for choosing later is very limited, which causes customer A to always choose first.

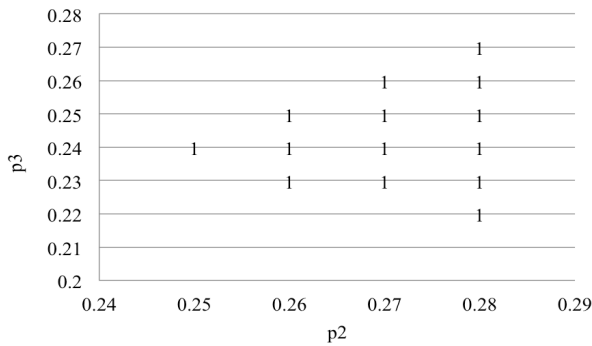


Figure 3(a): Types at $p_1 = 0.29$ and $r = 0$

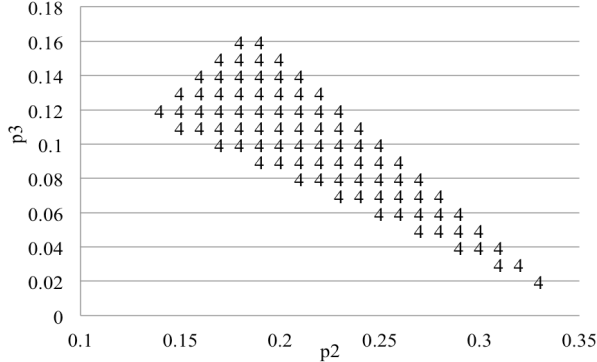


Figure 3(b): Types at $p_1 = 0.65$ and $r = 0$

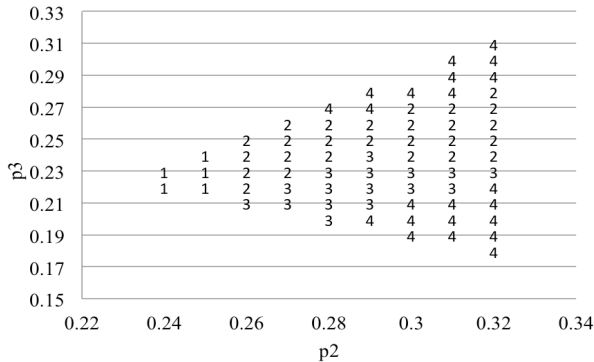


Figure 4: Types at $p_1 = 0.33$ and $r = 0$

However, in Figure 3(b), we can see Type 4, where customer A always decides to make decision later after he received any kind of signals. When p_1 is high ($p_1 > 0.57$), the best response of customer A is opposite, i.e., he will decide to make decision later and choose the table that is indicated as the larger one by the signal he received. The postponement strategy shows its value under this setting: when the imperfect but high quality signals are given, the uncertainty on the table size is low. It leads to clearly different expected table size for both tables. In such a case, customer A prefers to make decision later in order to collect and identify the larger table because it can provide a higher expected utility, compared with choosing the smaller table alone. In this case, learning from previous signals enables the latter customer to have a significant advantage. Therefore, customer A will choose to postpone his strategy to gain the advantages from learning.

Under some specific settings, we observe Type 2 and Type 3. In Type 2 case, customer A decides to make decision later only when he receives a subset of signals; customer B becomes the first customer and makes decision first. Type 3 is that customer A

chooses to make decision later only when receiving a subset of signals. Then customer B becomes the first customer but changes his strategy. At $r = 0$, when $0.29 < p_1 < 0.58$, the quality of other signals may affect the strategy applied by customer A. For instance, in Figure 4, we can see that when $p_1 = 0.33$, there are four types of Nash equilibrium. When p_2 is not closer to p_1 and p_3 is not closer to p_4 , signal 2 and -2 is more informative than 1 and -1. In such a case, customer A chooses to postpone his decision only when receiving signal 2 or -2. The reason of customer A's strategy is that when customer A received those informative signals, he can safely assume that he already finds the larger table and predicts that either both customers choose the same large table, or customer B will choose another table. In such a case, it doesn't harm to choose a table later, and he can collect more signals to identify the large table more accurately if he postpones his decision. In sum, customer A will choose to postpone his decision only when receiving a more informative signal. For customer B, when he receives the signal 1 and -1, the signals provide different belief in the state from the signals 2 and -2. Therefore, customer B will change his strategy only when the signals 1, and -1 are stronger enough to give him a confidence that the other table is the right table. In such a case, Type 3 outcome is reached. Otherwise, Type 2 outcome is expected.

5. CONCLUSION

In this paper, we study the influence of decision order and how agent may benefit for the postponement strategy from a game-theoretical perspective. We first present a general model. Then, a two-player case is studied. By analyzing the model, we classify the Nash equilibrium into four types according to the execution of the postponement strategy and its influences on the opponent's strategy. We propose an iterative-update algorithm to derive the Nash equilibrium of the proposed game. From the simulations, we show that rational agents will strategically postpone their decisions when the advantage of learning is significant enough to cover the loss in negative network externality.

6. ACKNOWLEDGMENTS

This work is supported by Ministry of Science and Technology under Grants 103-2218-E-001-002-MY2.

7. REFERENCES

- [1] V. Bala and S. Goyal, "Learning from neighbours," *The Review of Economic Studies*, vol. 65, no. 3, p. 595, 1998.
- [2] D. Acemoglu, M. Dahleh, I. Lobel, and A. Ozdaglar, "Bayesian Learning in Social Networks," *LIDS report 2780*, *Review of Economic Studies*, Jan. 2011.
- [3] D. Acemoglu and A. Ozdaglar, "Opinion dynamics and learning in social networks," *Dynamic Games and Applications*, vol. 1, pp. 3–49, 2011.
- [4] B. Golub and M. O. Jackson, "Naive Learning in Social Networks and the Wisdom of Crowds," *American Economic Journal: Microeconomics*, vol. 2, no. 1, pp. 112–149, 2010.
- [5] C.-Y. Wang, Y. Chen, and K. J. R. Liu, "Chinese restaurant game," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 898–901, Dec. 2012.
- [6] —, "Sequential Chinese Restaurant Game," *IEEE Transactions on Signal Processing*, Vol. 61, Issue 3, pp. 571–584, Feb. 2013.
- [7] C. Jiang, Y. Chen, Y.-H. Yang, C.-Y. Wang, and K. J. R. Liu, "Dynamic chinese restaurant game: Theory and application to cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 4, pp. 1960–1973, April 2014.