Repeated Keyword Auctions Played by Finite Automata

Wenkui Ding[1], Tie-Yan Liu[2], Tao Qin[2], and Pingzhong Tang[1]

[1] Tsinghua University, Beijing, 100084, P.R. China
dingwenkui@gmail.com,kenshinping@gmail.com
[2] Microsoft Research Asia, No.5, Danling Street, Haidian District, Beijing, 100080, P. R. China
{taoqin,tyliu}@microsoft.com

Abstract. While our understanding of keyword auctions, and generalized second-price auctions (GSP) in particular, as single-shot games has been advanced immensely over the past decade, there are perversely few works that consider them as repeated games, given their nature of repeated execution. One reason is that repeated games are generally hard to analyze when the underlying stage games are complex. Existing work of this sort largely focuses on the so-called Myopic best response strategies where agents always best respond to the plays of the previous round. Best response strategies have clear intuitions and could sometimes accidentally explain certain bidding behaviors risen in real data, unfortunately, they are vulnerable to manipulation and fail to form equilibria. The second reason is that, Folk theorem demystifies equilibria in repeated games, by saying that every sensible utility profile can be achieved by a Nash equilibrium, leaving little space for exploration on the game theoretical front. However, merely looking at the payoffs can hardly help us understand the rationality of the advertisers' bidding strategies or some very interesting interactions they develop during the course of repeated play.

In this paper, we restrict attention to the set of strategies that can be described by Mealy machines, a type of finite automata that generalizes Moore machine, which is widely used to model repeated prisoner's dilemma (PD). The set of automation strategies are rich enough to implement a wide variety of outcomes: we prove a new version of Folk theorem by explicitly constructing a pair of two-state Mealy machines that can form sub-game perfect equilibrium (SPE). This is in sharp contrast to previous work on myopic best responses, which cannot form any equilibrium at all. The new Folk theorem subsumes a previous Folk theorem that uses two-state Moore machines. The construction also facilitates optimizations over the set of outcomes for certain desirable objectives. The set of automation strategies are also descriptive enough to capture interesting interactions between advertisers, such as collusions, threats and punishments. This is also in sharp contrast to the line of work on Folk theorem, which says little on the rationality and intuitions of strategies. Specifically, we find that strategies such as Tit-for-Tat, Grim Trigger as well as Punisher, all of which mainly used for PD, also have their counterparts in repeated keyword auctions. Moreover, we obtain sufficient conditions under which these strategies form SPE. Our proof of SPE is based on a novel generalization of single-deviation principle, which might be of independent interests.

1 Introduction

Keyword auctions is the main format for selling ad slots to advertisers in sponsored search and the major monetization approach for search engine companies. One important feature that distinguishes keyword auction from any one-shot auction, is its repetition over time. During the course of repeated interactions, advertisers can adjust their bids dynamically and condition future play on past observations. As a result, advertisers can develop sophisticated and interesting strategies that cannot be explained by restricting attentions to the underlying one-shot game, and this is indeed the case evidenced from realistic data [4, 12]. In this paper, we deviate from the standard treatment of keyword auctions as single-shot game [5, 11] and analyze keyword auctions from the perspective of repeated game.

A major difficulty in analyzing such a repeated game is the prohibitively large strategy space. Since a repeated game is an extensive-form game with infinite horizon, one must, if possible, specify an action for each history on an infinite game tree. It is therefore sensible to look at subsets of strategies with compact representations. One important subset, which also serves as a start point of our work, is the class of myopic best responses (aka. greedy bidding strategies) [2, 6]. When using a myopic best response strategy, an advertiser places a bid that best responds to the others' bids of the previous round. For standard keyword auctions such as generalized-second price auction, there are typically a range of bids that are best responses, as long as they all guarantee the advertiser with her favorite slot. Depending on how the advertiser's attitudes towards others (e.g., cooperative or spiteful), there are different meaningful best responses strategies studied in this literature [2, 6], including the so-called Altruistic Bidding (AB), Competitor Busting (CB).

However, sticking to a myopic best response strategy is rigid and vulnerable to manipulations. Imagine all the other agents use myopic best responses, instead of best responding to the bids in the *previous round* as myopic best response would suggest, one could first compute others' myopic bids and achieve a best response to the others' bids of the *current round*. Indeed, to our best knowledge, none of these native myopic best responses constitute an equilibrium and empirically, manipulations has been observed from Yahoo! auction dataset [12].

To overcome the limitation of these greedy bidding strategies, we propose the use of finite automata to model advertisers' bidding strategies. A finite automaton consists of a finite set of states, an initial state, an action function and a state transition function. Given the state and the actions of the current round, the action function of the automata specifies an action for the current state, and the transition function of the automation determines the next state based on the current state and action profiles. Finite automata have been widely used to model repeated prisoner's dilemma and repeated games in general. [10] studied properties of equilibrium in these games where the players are restricted to automation strategies and applied the results to the repeated prisoner's dilemma. [1] studied a similar setting and provided necessary conditions on the form of Nash equilibrium strategies and payoffs. To our best knowledge, our work is the first to consider automation strategies in the context of repeated keyword auctions.

By the use of finite automata, we can construct complex bidding strategies using simple bidding strategies as building blocks (such as myopic strategies, and equilibrium strategies of one-shot game). In our construction, we use one state to record a simple bid strategy. For example, we could have a state called "myopic" such that whenever the advertiser is in this state, she will play a myopic best response for this round. A key observation here is that, to determine the action for the current round, one will need to know the "state (bidding strategy)" as well as the bids of previous rounds. Standard *Moore machine* fails for this purpose: in a Moore machine, action played in this round is solely determined by the current state. For our purpose, it is natural to consider *Mealy machine* for keyword auctions. A Mealy machine differs from Moore machine in that the action played depends both on the current state and actions of the previous round[3].

We now give a preview of our main contribution of this paper. Our contribution is two-fold.

We first prove a new version of Folk theorem of SPE that strictly generalizes the one developed in [9]. Our proof of the theorem is constructive via explicitly use of Mealy machine.

- The use of Mealy machine allows us to reach a strictly larger set of SPE outcomes than the set defined in [9].
- Via the constructive proof, we are able to set up a simple formulation of optimizing an objective function over these SPE outcomes. The general problem turns out to be NP-hard and we give a fully polynomial-time approximation scheme (FPTAS).

We then investigate several specific SPEs, that are constructed from automation strategies, for repeated keyword auctions.

- We establish a new version of single deviation principle that allows us to check whether a profile of simple automaton strategies constitutes a SPE.
- By making use of this single deviation principle, we are able to obtain the conditions under which, commonly seen automation strategies, such as Tit-for-Tat, Grim Trigger and Punisher, can form SPE.

[3] In fact, one can be easily construct an equivalent Moore machine from any Mealy machine by renaming any bids-state pair in the Mealy machine as a state in the resulting Moore machine. However, such a Moore machine will have an enormous amount of states that preclude standard SPE analysis techniques, such as single-deviation, that requires exhaustive enumeration of state profiles

The obtained results have other interesting implications: 1) by modeling advertiser strategies as finite automata, advertisers interact with each other not only on the bid level but also on the state level. Highly rewarded but unstable bidding profile sustain via credible threat. 2) cyclical bid change occurs in the equilibrium, which coincides with empirical findings in [12].

The remaining of the paper is organized as follows. Mealy machines are introduced in section 2 and a new version of Folk theorem is proved in section 3. In section 4, we proposed several automata bidding strategies. We characterize the structure of the sub-game perfect equilibrium (SPE) in repeated auctions and obtain sufficient conditions for SPE in section 5. The conclusion and future work are given in the last section.

2 Mealy Machines

As stated in the introduction, myopic best response strategies are vulnerable to manipulations. We consider a richer class of strategies, called automation strategies. In the literature, Moore machines have been used [10, 1] to model players' strategies in repeated games. However, Moore machines are not a ideal choice for repeated keyword auctions since to make a sensible decision for the current round, an advertiser needs to know both his/her strategy and the bids of other advertisers in previous rounds, while in a Moore machine, the action for this round can only depend on the current state. In contrast, Mealy machines fit well into our problems and in this work we use Mealy machines as finite automata. To adapt the idea of automaton strategy to the setting of keyword auctions, we consider an alternate automata called Mealy machine.

The formal definition of Mealy machines is given as follows.

Definition 1. *A Mealy machine M_i is a finite state automaton defined by a tuple $\langle Q_i, q_i^0, \alpha_i, \beta_i, \sigma_i, \lambda_i \rangle$ where*

- *Q_i is a finite set of states for player i*
- *$q_i^0 \in Q_i$ is an initial state*
- *α_i is the input space.*
- *β_i is the action space.*
- *The output function $\sigma_i : Q_i \times \alpha_i \rightarrow \beta_i$ chooses an action based on the current state and the current input.*
- *Given the current state and the input, the transition function $\lambda_i : Q_i \times \alpha_i \rightarrow Q_i$ determines the next state.*

Throughout the paper, we consider the repeated auctions for a keyword with two advertiser competing for two ad slots. We use $b_{i,t}$ to denote the bid of advertiser i at time step t and $b_{-i,t}$ to denote the bid of advertiser i's competitor at time step t. Figure 2 shows an example of a bidding automaton for player i where A is the set $\{(b_{t-2}, b_{t-1}) \in R_+^2 \times R_+^2 | b_{-i,t-1} = b_{i,t-2} + \epsilon\}$ and ϵ is a small positive constant. In this automaton,

$$\sigma_i(q_0, (b_{t-2}, b_{t-1})) = b_{-i,t-1} + \epsilon,$$

$$\sigma_i(q_1, (b_{t-2}, b_{t-1})) = v_i,$$
$$\forall q \in \{q_0, q_1\}, \lambda_i(q, (b_{t-2}, b_{t-1})) = q_0, \text{ if } (b_{t-2}, b_{t-1}) \in A,$$
$$\forall q \in \{q_0, q_1\}, \lambda_i(q, (b_{t-2}, b_{t-1})) = q_1, \text{ if } (b_{t-2}, b_{t-1}) \notin A.$$

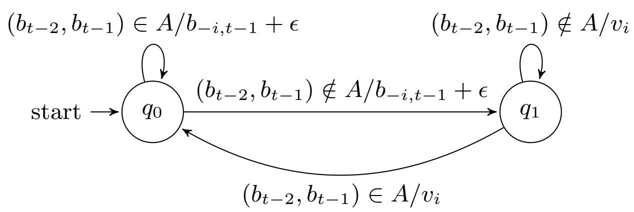

Fig. 1. An example of bidding automata

3 A New Version of Folk Theorem

In this section we prove a new version of Folk theorem to the one described in [9]. Our proof is constructive, thus provides an explicit means to implement the payoff profiles. In the following theorem, we use $u_i(\cdot)$ to denote the payoff function of player i.

Proposition 1 (Folk Theorem). *Given a one-shot game with a set of feasible action profiles* $(b^1, b^2, \cdots, b^T)$ *and a Nash equilibrium* $(b_1^*, b_2^*, \cdots, b_N^*)$ *with the payoff profile* $(e_1, e_2, \cdots, e_N)$ *in which*

$$\forall i, \quad \frac{1}{T}\sum_{j=1}^{T} u_i(b^j) \geq e_i, \tag{1}$$

there exists a subgame-perfect equilibrium of the infinitely repeated game, that provides the average utility v_i *to every player* i*.*

Proof. To prove our theorem, we use a simple two-state Mealy machine shown in Figure 3. When all players stay in state q_0, let the acution profile sequence $(b^1, b^2, \cdots, b^T)$ be the path that players go through. In state q_1, player i always play b_i^* where $(b_1^*, b_2^*, \cdots, b_N^*)$ is a Nash equilibrium of the single-shot game and player i obtains payoff e_i. μ is the trasition function and λ is the output function.

To obtain the condition to ensure this strategy profile is a sub-game perfect Nash equilibrium, we apply the single deviation principle and consider the following two cases.

1. No player has deviated in any previous period. By the single deviation principle, it is sufficient to check that, for all i

$$\frac{1}{T}\sum_{j=1}^{T} u_i(b_j) \geq e_i.$$

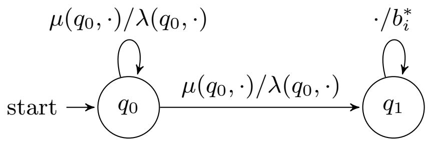

Fig. 2. A two-state bidding automata

The left-hand side is the average utility to player i from following the prescribed strategy. The right-hand side is the average utility to player i from deviation.

2. One or more players have deviated from the state profile q_0 in the past. In subsequent periods, all players play an open-loop strategy b_i^*. Because (b_1^*, b_2^*) is a Nash equilibrium of the single-shot keyword auction, nobody wants to deviate in any single round.

Remind that some existing best response strategies (such as AB) do not form an equilibrium, and others (such as BB) achieves the revenue equal to the revenue of a static equilibrium (such as the VCG equilibrium of GSP). The above result shows even a simple automata can do better: 1) automata constitutes a SPE; 2)the payoff to players can be greater than a static equilibrium.

With the new Folk theorem in mind, we are able to optimize over outcomes that can be implemented by SPE via our construction. The objective can be any function on a set of outcomes, including social welfare. We can write the following efficiency maximization problem

$$\begin{aligned} &\max_{T,b^1,b^2,\cdots,b^T} \tfrac{1}{T}\textstyle\sum_{j=1}^{T} W(b^j) \\ &\text{s.t.} \quad \forall i, \textstyle\sum_{j=1}^{T} u_i(b^j) \geq T e_i \\ &\forall j, b^j \in \{a^1, a^2, \cdots, a^K\} \end{aligned} \tag{2}$$

where $\{a^1, a^2, \cdots, a^K\}$ is all the action profiles, or its equivalent form

$$\begin{aligned} &\max_x \frac{\sum_{j=1}^{K}\sum_{i=1}^{2} u_i(a^j)x_j}{\sum_{j=1}^{K} x_j} \\ &\text{s.t.} \quad \forall i, \textstyle\sum_{j=1}^{K} u_i(a^j)x_j \geq \sum_{j=1}^{K} x_j e_i \\ &\forall j, x_j \in \{0,1\} \end{aligned} \tag{3}$$

The objective function of the problem is rational which turns out to be NP-hard [3]. We use the techniques in [3, 8, 7] to tackele the problem.

The basic idea is to create an approximate Pareto-optimal front in pseudo-polynomial time and return the best solution among the approximate Pareto-optimal front. It is have proved in [7] that at least one solution in the approximate Pareto-optimal front is an approximation solution. In our problem, it is sufficient to solve the following problem in pseudo-polynomial time: if there exist x such that

$$\begin{aligned} & \textstyle\sum_{j=1}^{K} c_j x_j = r_1, \\ & \textstyle\sum_{j=1}^{K} x_j = r_2, \\ \text{s.t.} \quad & \forall i, \textstyle\sum_{j=1}^{K} a_{i,j} x_j \geq 0, \\ & x_j \in \{0,1\} \end{aligned} \tag{4}$$

where $\forall j, c_j, r_1, r_2 \in Z_+$. This is a generalization of subset sum problem which can be solved by dynamic programming in pseudo-polynomial time.

The algorithm is given in Table 1.

Algorithm 1 A fully polynomial-time approximation scheme

Require: The optimization problem π and $\epsilon > 0$

Initialization: a set $P_\epsilon(\pi) = \emptyset$

Consider the box of possible function values ($\sum_{j=1}^{K} \sum_{i=1}^{2} u_i(a^j) x_j$ and $\sum_{j=1}^{K} x_j$). Divide this box into smaller boxes, such that in each dimension, the ratio of successive divisions is equal to $(1+\epsilon)^{\frac{1}{2}}$.

for each corner point (r_1, r_2) of all smaller boxes **do**

Define $r'_1 = [K/\epsilon'_1]$ where $\epsilon'_1 = (1+\epsilon)^{1/2} - 1$ and $r'_2 = [K/\epsilon'_2]$ where $\epsilon'_2 = 1 - (1+\epsilon)^{-1/2}$. Define a function $f_1(x) = \sum_{j=1}^{K} a'_j x_j$ where $a'_j = \min\{r'_1, [\sum_{i=1}^{2} u_i(a^j) r'_1 / r_1]\}$ and a function $f_2(x) = \sum_{j=1}^{K} b'_j x_j$ where $b'_j = [r'_2/r_2]$. Find a solution to the problem

$$\begin{aligned} & \textstyle\sum_{j=1}^{K} c_j x_j = r'_1, \\ & \textstyle\sum_{j=1}^{K} x_j = r'_2, \\ \text{s.t.} \quad & \forall i, \textstyle\sum_{j=1}^{K} u_i(a^j) x_j \geq 0, \\ & x_j \in \{0,1\} \end{aligned} \tag{5}$$

and add it to the set $P_\epsilon(\pi)$, or assert that no solution exists.

end for

Among all the solutions in $P_\epsilon(\pi)$, return the one with the maximum function value.

Proposition 2. *The algorithm in Table 1 provides an FPTAS for solving problem 3.*

4 Automaton Strategies for Repeated Keyword Auctions

In this section, we describe several automaton strategies for repeated keyword auctions built up on simple bidding strategies.

4.1 Simple Bidding Strategies in Keyword Auctions

The most well-known bidding strategies are greedy bidding strategies proposed in [2]. The greedy strategies are the myopic best response for an advertiser: when the advertiser needs to set a bid for the current round, he/she assumes that other advertisers will not change their bids and will use the same bid as in the previous round, and chooses a bid to maximize his/her utility. Usually there is a range of optimal bids that can maximize his/her utility, and different greedy bidding strategies corresponds to different choices of a single bid in the optimal range. Altruistic Bidding (AB) strategy uses the smallest bid necessary to get the optimal ad slot and to make other advertisers pay as little as possible. Competitor Busting (CB) uses the largest bid to get the optimal ad slot and to make competitors pay as much as possible. For the case of two advertisers competing for two slots, AB can be written as

$$b_{i,t+1} = \begin{cases} b_{-i,t} + \epsilon, & \text{if } c_1(v_i - b_{-i,t}) > c_2(v_i - r), \\ r, & \text{otherwise}, \end{cases}$$

where c_i is the click-through rate of slot i, v_i is the per-click valuation of advertiser i, r is the reserve price set by the search engine and ϵ is a small positive constant. Similarly, CB can be written as

$$b_{i,t+1} = \begin{cases} b_{-i,t} - \epsilon, & \text{if } c_1(v_i - b_{-i,t}) < c_2(v_i - r) \text{ and } b_{-i,t} - \epsilon < v_i, \\ v_i, & \text{otherwise}. \end{cases}$$

In general, these strategies do not form equilibria in the repeated game.

Besides AB and CB, we also use another pair of simple greedy bidding strategies to show other types of cooperation that cannot be explained by static analysis. The strategy S1 is

$$b_{i,t+1} = \min\{\max\{b_{-i,t} - \epsilon, r\}, v_i\}$$

and the strategy S2 is

$$b_{i,t+1} = v_i.$$

These strategies may not be best responses. While leaving the top position to the opponent, S1 also pushes the price paid by the opponent. In practice, once the opponent runs out of budget, the top position can be obtained with a very low price. S2 punishes the opponent by bidding the true value and it is similar to the defection in PD. These strategies are similar to those proposed in [12].

4.2 Finite Automata Built on Greedy Bidding Strategies

An implicit assumption in existing work is that each advertiser adopts a fixed greedy bidding strategy and does not change it in repeated auctions. However, in practice, we often observe that advertisers change their bidding strategies in the game. Thus, more complex models are needed to describe advertisers bidding behaviors. Our proposal is to use finite automata introduced in Section 2 to express more complex bidding strategies of advertisers.

Now we make the following discussion based on greedy bidding strategies AB and CB. However, it can be easily generalized to other simple bidding strategies as well.

As a bidding strategy, the input space of a Mealy machine includes all possible bid profiles in the previous two rounds and the output space includes all possible bids could be submitted by advertiser i. In each state $q_i \in Q_i$, advertiser i bids according to a simple bidding strategy (e.g. AB or CB). The output function $\sigma_i : Q_i \times \alpha_i \to \beta_i$ choose the bid based on the current state and the bid profile in the last round. Given the current state and the input, the transition function $\lambda_i : Q_i \times \alpha_i \to P(Q_i)$ determines the next state by a distribution over all possible states. We choose transition function so that state transitions only depends on the opponent's bidding strategies. By comparing the bids of the other player in the last two rounds, a player comes to a conclusion whether a simple bidding strategy (AB or CB) is adopted by the other player in the previous round or a different bid other than the bid generated by a simple bidding strategy is adopted. If a simple bidding strategy is adopted by the opponent, the state of player i transits to a deterministic state denoted by $\lambda_i(q_i, AB)$ or $\lambda_i(q_i, CB)$ with abuse of the notation λ. We use b^{AB} to denote the bid generated by AB and b^{CB} to denote the bid generated by CB. If neither AB nor CB is observed, there are three cases:

- if the bid of the opponent in the last round is smaller than b^{AB}, the state of player transit to $\lambda_i(q_i, AB)$;
- if the bid of the opponent in the last round is larger than b^{CB}, the state of player transit to $\lambda_i(q_i, CB)$
- if the bid of the opponent is between b^{AB} and b^{CB}, the state of player transit to $\lambda_i(q_i, AB)$ with the probability $f_i(b; b^{CB}, b^{AB})$ and to $\lambda_i(q_i, CB)$ with the probability $1 - f_i(b; b^{CB}, b^{AB})$.

The above definition of the transition functions avoids the complex state transition in a general Mealy machine.

We introduce several automata strategies here and analyze the conditions under with they form SPE.

The diagram of the Tit-for-Tat strategy is shown in Figure 4.2. In this strategy, every player follows the bidding pattern of the opponent in the previous round.

The diagram of the Grim Trigger strategy is shown in Figure 4.2. In this strategy, each player falls into the state CB if the other plays CB once.

The diagram of the Punisher strategy is shown in Figure 4.2.

We will also consider the above three automata with AB replaced by S1 and CB replaced by S2.

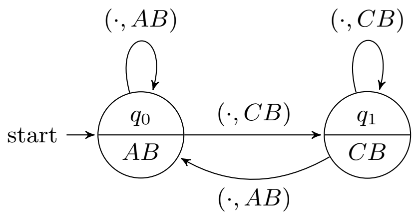

Fig. 3. The Tit-for-Tat strategy

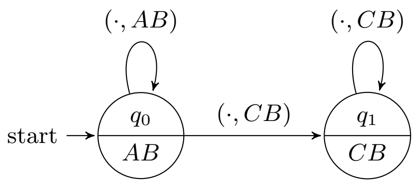

Fig. 4. The Grim Trigger strategy

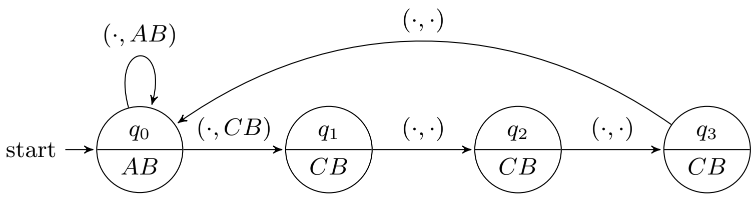

Fig. 5. The Punisher strategy

5 SPE Formed by Automata

5.1 Single Deviation Principle

Lemma 1. *A pair of Mealy machine constitutes a subgame-perfect equilibrium if and only if every player cannot increase his/her payoff by deviating in a single round for any state profile and input in any round.*

The above lemma results from single deviation principle directly and holds for any Mealy machine. If we use simple bidding strategies such AB and CB to construct Mealy machines, stronger results can be obtained.

Lemma 2. *A pair of Mealy machine constitutes a subgame-perfect equilibrium if and only if every player cannot increase his/her payoff by deviating in a single round to b^{CB} or b^{AB} for any state profile in any round and any bid profile in the last two rounds; a pair of Mealy machine constitutes a subgame-perfect equilibrium if and only if every player cannot increase his/her payoff by deviating in a single round to b^{S1} or b^{S2} for any state profile in any round and any bid profile in the last two rounds.*

Despite the non-deterministic nature of the Mealy machine in the original model, this lemma tells us that we only need to check deterministic deviations.

Proof. Without loss of generality, we assume player i deviate from a bid b^{EQ} to a bid between b^{CB} and b^{AB} and the deviation to b^{CB} or b^{AB} is not better. Then the state of his opponent will be CB with probability $f_{-i}(b; b^{CB}, b^{AB})$ and AB with probability $1 - f_{-i}(b; b^{CB}, b^{AB})$. After this single deviation, the payoff of player i is

$$\begin{aligned}
& f_{-i}(b^{EQ}; b^{CB}, b^{AB})u(b^{CB}) + (1 - f_{-i}(b^{EQ}; b^{CB}, b^{AB}))u(b^{AB}) \\
\leq\ & f_{-i}(b^{EQ}; b^{CB}, b^{AB})u(b^{EQ}) + (1 - f_{-i}(b^{EQ}; b^{CB}, b^{AB}))u(b^{EQ}) \\
=\ & u(b^{EQ})
\end{aligned}$$

where we have assumed $u(b^{CB}) \leq u(b^{EQ})$ and $u(b^{AB}) \leq u(b^{EQ})$. Apply the single deviation principle, we obtained the desired result. The proof for the Mealy machines constructed by strategies S1 and S2 is very similar.

Based on the above lemma, we have the following proposition.

Proposition 3. *A pair of Mealy machine constitutes a subgame-perfect equilibrium if and only if every player cannot increase his/her payoff by deviating to b^{AB} or b^{CB} in a single round for any state profile and any bid profile in the previous round; a pair of Mealy machine constitutes a subgame-perfect equilibrium if and only if every player cannot increase his/her payoff by deviating to b^{S1} or b^{S2} in a single round for any state profile.*

5.2 Sufficient Conditions for SPE

Based on the result in the last subsection, we have the following theorem.

Theorem 1. *If AB and CB are used as building block in automata strategies,*

a.1 a pair of Tit-for-Tat strategies cannot form a sub-game perfect equilibrium.
a.2 a pair of Grim Trigger strategies form a sub-game perfect equilibrium if and only if

$$\frac{v_2 - r + 4\epsilon}{2v_2 - 2r + 4\epsilon}(v_1 - r) - \frac{v_2 - r}{4\epsilon} > v_1 - v_2; \tag{6}$$

a.3 a pair of Punisher strategies cannot form a sub-game perfect equilibrium;

If $S1$ and $S2$ are used as building block in automata strategies,

b.1 A pair of Tit-for-Tat strategies cannot form a sub-game perfect equilibrium.
b.2 A pair of Grim Trigger strategies form a sub-game perfect equilibrium if and only if

$$c_1(kv_1 - \frac{k}{2}(v_2 + r - \epsilon) + c_2(k+2)(v_1 - r) \geq 2(k+1)c_1(v_1 - v_2) \tag{7}$$

where $k = \frac{v_2 - r - \epsilon}{2\epsilon}$.
b.3 A pair of Punisher strategies cannot form a sub-game perfect equilibrium.

Proof. We give the proofs for cases *a.1*, *a.2* and *b.2* here. Other cases are similar to the case *a.1* and we proved in the full version of the paper.

Define a threshold value $\theta_i = (1 - \frac{c_2}{c_1})v_i + \frac{c_2}{c_1}r$ for advertiser i. It can be verified that the optimal position for advertiser i is the top position if and only if $b_{-i} < \theta_i$.

a.1 We first consider the case that starts with the state profile (q_1, q_1). If both players follow the Tit-for-Tat strategy, they will both stay in the state q_1 in subsequent rounds and bid according to CB. If player 1 deviates to AB, the state profiles in subsequent rounds has a cycle $(q_1, q_0), (q_0, q_1)$. Consider another case that starts with the state profile (q_0, q_1). If both players follow the prescribed strategy, the subsequent state profiles has a cycle $(q_1, q_0), (q_0, q_1)$. If player 1 deviates, both players will be in the state q_1. Put the above two cases together, it is clear that player 1 prefers to deviate in at least one case, unless the utility for player 1 does not change after the deviation in both cases.

We then analyse the average utility that player 1 could obtain in both cases. Denote the starting bid profile by (h_1, h_2). For the first case, all of the greedy bidding strategies used by both advertisers are shown in Fig. 5.2 with the top nodes belong to player 1 and the bottom nodes belong player 2. Consider a sequence of greedy bidding strategies on the path denoted in the figure. If $v_2 > \theta_1$, it can be shown that a cycle of the corresponding bid sequence is $v_1, v_2, v_2 - \epsilon, v_2 - 2\epsilon, \cdots, v_2 - k\epsilon$ where $k = \lfloor \frac{v_2 - \theta_1}{\epsilon} \rfloor + 1$ and v_1 is submitted by player 1; otherwise, a cycle is v_1, v_2 and v_1 is submitted by player 1. Similar result can be obtained for the nodes out of the path in the figure. For the second case, the nodes on the path use AB instead of CB and the other nodes are not affected. A cycle of the bid sequence is $r, r+\epsilon, r+2\epsilon, \cdots, r+l\epsilon$ where $l = \frac{\theta_2 - r}{\epsilon}$.

Combine these results, we have:

– If $v_2 > \theta_1$, the average utility of player 1 is $c_1(v_1 - v_2)$ before deviation and $\frac{c_1}{l+1}\sum_{i=0}^{\frac{l-1}{2}}(v_1 - r - 2i\epsilon) + \frac{c_2}{2}(v_1 - r)$ after the deviation. So it must hold that $c_1(v_1 - v_2) = \frac{c_1+c_2}{2}(v_1 - r) - \frac{c_1(l-1)}{4}\epsilon$.
– Similarly, change the role of player 1 and 2, we get the average utility of player 1 is 0 before deviation and a non-negligible amount after the deviation. So the Tit-for-Tat strategies are not a sub-game perfect equilibrium.

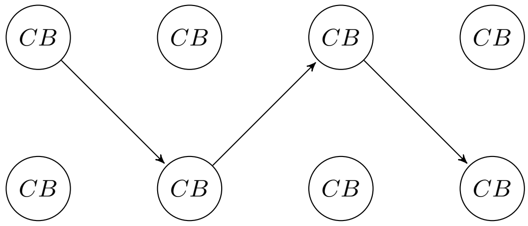

Fig. 6. State profiles

a.2 First, we consider a case that two advertisers both adopt q_0 at the beginning of the game. Then according to the bidding rule of the Grim Trigger, after a long term running, the bid profiles have a cycle $(r+\epsilon, r), (r+\epsilon, r+2\epsilon), (r+3\epsilon, r+2\epsilon), \cdots, (r+2k-1, r+2k), (r+2k+1, r+2k), (r+2k+1, r)$ where $k = \frac{v_2 - r}{2\epsilon}$. Thus The payoff of player 1 from the top slot is

$$\begin{aligned}&\frac{1}{2k+2}\left((v_1 - r) + (v_1 - r - 2\epsilon) + \cdots + (v_1 - r - 2k\epsilon) + (v_1 - r)\right)\\ &= \frac{k+2}{2(k+1)}(v_1 - r) - \frac{k}{2}\epsilon\end{aligned}$$

and the payoff of player 2 from the top slot is

$$\begin{aligned}&\frac{1}{2k+2}\left((v_2 - r - \epsilon) + (v_1 - r - 3\epsilon) + \cdots + (v_1 - r - (2k-1)\epsilon)\right)\\ &= \frac{k}{2(k+1)}(v_2 - r) - \frac{k^2}{2(k+1)}\epsilon.\end{aligned}$$

If one player deviates to CB, then these payoffs are $v_1 - v_2$ and 0 respectively. The equilibrium condition requires that

$$\frac{\lfloor\frac{v_2 - r}{2}\rfloor + 2}{2(\lfloor\frac{v_2 - r}{2}\rfloor + 1)}(v_1 - r) - \frac{\lfloor\frac{v_2 - r}{2}\rfloor}{2} > v_1 - v_2.$$

In the other cases, at least one player adopt q_1 at the beginning of the game. After a long term running, the state profiles have a cycle (q_1, q_1) so both player bids according to CB not matter how a single player deviates.

b.2 If both advertisers adopt q_0 at the beginning of the game and follow the Grim Trigger strategy thereafter, a cycle of the state profile is (q_0, q_0). After a long term running, the bid profiles have a cycle $(v_2-\epsilon, v_2), (v_2-\epsilon, v_2-3\epsilon), (v_2-3\epsilon, v_2-2\epsilon), \cdots, (r, r-\epsilon), (r, v_2)$. The average utility of player 1 is $\frac{1}{2k+2}\{c_1(v_1-(v_2-2\epsilon)) + c_1(v_1-(v_2-4\epsilon)) + \cdots + c_1(v_1-(r+\epsilon)) + (k+2)c_2(v_1-r)\}$ where $k = \frac{v_2-r-\epsilon}{2\epsilon}$.

If player 1 deviates to CB, the cycle of the bid profile will be (v_1, v_2) which means that every player try to punisher his/her opponent by submitting his/her true value. The utility of player 1 is $c_1(v_1-v_2)$.

By the single deviation principle, the equilibrium condition requires that

$$c_1(kv_1 - \frac{k}{2}(v_2 + r - \epsilon) + c_2(k+2)(v_1 - r) \geq 2(k+1)c_1(v_1 - v_2).$$

In the other cases, at least one player adopt q_1 and bids according to CB at the beginning of the game. If a player deviate unilaterally, the cycle of the state profiles remains the same.

6 Conclusions and Future Work

We have studied the set of strategies described by Mealy machines. We proved a new version of Folk theorem by explicitly constructing a pair of two-state Mealy machines that can form sub-game perfect equilibrium. The construction also facilitates optimizations over the set of outcomes for certain desirable objectives. The set of automation strategies are also descriptive enough to capture inter- esting interactions between advertisers, such as collusions, threats and punishments. We find that strategies such as Tit-for-Tat, Grim Trigger as well as Punisher, all of which mainly used for PD, also have their counterparts in repeated keyword auctions. We obtained sufficient conditions under which these strategies form SPE.

For future work, we plan to verify the connections between automation strategies in SPE and real data generated from sponsored search. It is also interesting to look at strategies described by more complex automations such as push down automata.

References

1. 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.
2. Matthew Cary, Aparna Das, Ben Edelman, Ioannis Giotis, Kurtis Heimerl, Anna R Karlin, Claire Mathieu, and Michael Schwarz. Greedy bidding strategies for keyword auctions. In *Proceedings of the 8th ACM conference on Electronic commerce*, pages 262–271. ACM, 2007.

3. José R Correa, Cristina G Fernandes, and Yoshiko Wakabayashi. Approximating a class of combinatorial problems with rational objective function. *Mathematical programming*, 124(1-2):255–269, 2010.
4. Benjamin Edelman and Michael Ostrovsky. Strategic bidder behavior in sponsored search auctions. *Decision support systems*, 43(1):192–198, 2007.
5. Benjamin Edelman, Michael Ostrovsky, and Michael Schwarz. Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords. Technical report, National Bureau of Economic Research, 2005.
6. Patrick Maillé, Evangelos Markakis, Maurizio Naldi, George D Stamoulis, and Bruno Tuffin. Sponsored search auctions: an overview of research with emphasis on game theoretic aspects. *Electronic Commerce Research*, 12(3):265–300, 2012.
7. Shashi Mittal and Andreas S Schulz. A general framework for designing approximation schemes for combinatorial optimization problems with many objectives combined into one. *Operations Research*, 61(2):386–397, 2013.
8. Christos H Papadimitriou and Mihalis Yannakakis. On the approximability of trade-offs and optimal access of web sources. In *Foundations of Computer Science, 2000. Proceedings. 41st Annual Symposium on*, pages 86–92. IEEE, 2000.
9. David C. Parkes and Seuken Sven. Economics and computation:a design approach. Technical report, 2013.
10. Ariel Rubinstein. Finite automata play the repeated prisoner's dilemma. *Journal of economic theory*, 39(1):83–96, 1986.
11. Hal R Varian. Position auctions. *International Journal of Industrial Organization*, 25(6):1163–1178, 2007.
12. Xiaoquan Michael Zhang and Juan Feng. Cyclical bid adjustments in search-engine advertising. *Management Science*, 57(9):1703–1719, 2011.