Best-Reply Mechanisms
(extended abstract)

Noam Nisan[*] Michael Schapira[†] Gregory Valiant[‡] Aviv Zohar [§]

Abstract

In many real-world settings, strategic agents are instructed to follow *best-reply dynamics*. Indeed, many computational protocols are based on such repeated *greedy* interactions. Such settings give rise to a natural question, that has received very little attention: *Is it in the best interest of the strategic agents to follow best-reply dynamics*? That is, is it true that a player cannot improve his final outcome by not behaving *myopically*? Surprisingly, we find that in many interesting cases the answer to this question is *Yes*. This enables us to design *incentive-compatible* algorithms for many environments, both old and new.

We initiate the study of best-reply dynamics in the context of *mechanism design*. We first describe a structural property of games that implies that best-reply dynamics not only converge to a pure Nash equilibrium, but are also incentive-compatible. We prove that *all* the following examples have this helpful structure: Internet protocols (TCP-inspired congestion-control settings, and interdomain routing environments), classic cost-sharing problems, and several economic stable matching problems. Finally, we consider two important auction settings that do not have this special structure: iterative auctions with unit-demand bidders, and adword auctions. Using inherently different techniques, we prove that best-reply dynamics converge and are incentive-compatible in these settings.

[*]The School of Computer Science and Engineering, The Hebrew University of Jerusalem, Israel. Google, Tel Aviv, Israel. Supported by grants from the Israel Science Foundation. E-mail: noam@cs.huji.ac.il

[†]Department of Computer Science, Yale University, CT, USA, and Computer Science Division University of California at Berkeley, CA, 94720 USA. Supported by NSF grant 0331548. E-mail: michael.schapira@yale.edu.

[‡]Computer Science Division University of California at Berkeley, CA, 94720 USA. Work supported under a National Science Foundation Graduate Research Fellowship. E-mail: gvaliant@cs.berkeley.edu

[§]The School of Computer Science and Engineering, The Hebrew University of Jerusalem, Israel, and Microsoft Israel Research and Development Center. Supported by grants from the Israel Science Foundation. E-mail: avivz@cs.huji.ac.il

1 Introduction

Perhaps the most natural approach for finding a (pure) Nash-equilibrium of a given game is executing *best reply dynamics*: start with an arbitrary strategy profile of the players, and in each step let some player switch his strategy to be the best reply to the current strategies of the others. If this process converges, then we have reached a pure Nash equilibrium, and moreover, a highly justifiable one. The many qualities of best-reply dynamics (simplicity, being distributed in nature, low costs in terms of memory and space, and many more) make it very attractive from an algorithmic perspective. Indeed, many protocols used in practice are essentially executions of best-reply dynamics (Internet protocols, auctions, economic markets, and more). As observed in [28], a noteworthy example for this is the *Border Gateway Protocol (BGP)*, that handles *interdomain routing* in today's Internet.

Best-reply dynamics are the subject of a large body of literature in game-theory and computer-science. Often, in such contexts, best-replying is adopted without any strategic justification ("myopic play"). That is, it is assumed that players obediently follow best-reply dynamics. However, in many cases (e.g., interdomain routing [28]) this assumption is very problematic. In this paper we present and address a natural question that has received little, if any, attention: "*When are best-reply dynamics strategically justified*?". That is, when are best-reply dynamics incentive-compatible, in the sense that a player cannot better his final outcome by unilaterally deviating from them. To answer this fundamental question, we consider best-reply dynamics in the context of *mechanism design*.

Our framework is the standard one of mechanism design with private values, where each player's private value is his own utility function. In many such scenarios a natural objective is to implement a Nash equilibrium of the full-information game induced by players' utilities. An indirect mechanism that we may hope can implement this outcome is a *"repeated-reply mechanism"* where players repeatedly announce a reply to other players' most recent announcements. An appealing property of best-reply dynamics is that the actions of players are "*uncoupled*": each player's best-reply depends on his own utility function and not on the utilities of other players. Hence, best-reply dynamics are feasible strategies in games with private information. We refer to repeated-reply mechanisms in which the prescribed behaviour for players is best-replying as "best-reply mechanisms". We want best-reply dynamics to be in some kind of equilibrium, in the sense that as long as all other players are continuously best-replying, the best course of action for a player is to do the same.

Unfortunately, simple examples show that best-replying in repeated-reply mechanisms is not necessarily incentive-compatible, even in very simple games. Yet, surprisingly, we are able to design incentive-compatible algorithms for various important and well-studied settings, and for several new and interesting ones. We point out several realistic environments for which we show (1) convergence of best-reply dynamics to a pure Nash equilibrium, and (2) that best-reply strategies are incentive compatible in ex-post-Nash equilibrium. Our results show that best-reply mechanisms are a fairly general tool for the design of incentive-compatible mechanisms. We note that other than the celebrated VCG framework, which poses many serious obstacles [35], almost no such techniques are known (see [16, 20, 26] for *randomization-based* techniques to obtain truthfulness in interesting cases).

Many of our results are based on identifying a subclass of max-solvable games (defined in [36]), which we term "*universally-max-solvable*", for which we show that best-reply dynamics are guaranteed to converge to a pure Nash equilibrium in an incentive-compatible manner. In fact, the structure of these games even implies that best-reply dynamics (1) are collusion-proof (resilient to manipulations by *coalitions* of players of *any* size) (2) converge to a Pareto-optimal solution (3) converge even in asynchronous settings (in which players may update their strategies *simultane-*

ously and learn of each others' actions via update messages that can be delayed or lost) [36], and (4) converge in polynomial time (in the total size of the strategy spaces of the players) [36].

At first sight, the definition of universally-max-solvable games might appear unreasonably restrictive. However, we prove that *all* the following examples can be formulated as universally-max-solvable games: Internet protocols (TCP-inspired congestion-control settings, and interdomain routing environments), classic cost-sharing problems, and several economic stable matching problems. This simplifies, strengthens, and unifies known results, and leads to new ones. In particular, our results generalize, and provide greatly simplified proofs for the main theorem in [28], as well as fundamental results regarding BGP convergence [22, 18]. Still, many games *do not* fall into the restricted category of universally-max-solvable games. For such games, proving that best-reply dynamics converge and are incentive-compatible (if that is indeed the case) requires inherently different techniques.

We provide two examples in the space of games that lie beyond universal-max-solvability, for which best-reply dynamics converge in an incentive-compatible fashion. We regard these examples as a first step towards finding more general structures than universal-max-solvability that lead to incentive-compatible convergence of best-reply dynamics. Both of our examples are famous auction settings: auctions with unit-demand bidders, and adword auctions. Unlike before, our proofs now decouple the question of whether best-reply dynamics converge to a pure Nash equilibrium, and the question of whether they are incentive-compatible, and address these separately via different methods. Our proofs for both cases are based on classic results in auction theory. In particular, our proof for the unit-demand case relies on the famous links between this environment and primal-dual algorithms [5], and specifically Kuhn's Hungarian method for finding weighted matchings in bipartite graphs [25].

1.1 Examples of Best-Reply Mechanisms

- **Internet Protocols:** Our first two examples are inspired by two of the most famous protocols that make up the Internet infrastructure.

  - **Congestion-Control Games:** These games are inspired by the Transmission Control Protocol (TCP).[1] The setting consists of a directed graph, with the nodes being non-strategic routers. Strategic players ("flows") want to transmit packets along fixed paths in the network. The routers drop packets if the network is congested, using Weighted Fair Queuing [9]. Flows must decide on their transmission rates, and wish to maximize their throughput. We consider a natural distributed congestion-control protocol that simulates best-reply dynamics.

    Theorem: *In congestion-control games, best-reply dynamics are incentive-compatible and converge to a stable outcome even in asynchronous Internet-like settings. Furthermore, the resulting stable state maximizes the max-min fairness.*

  - **Routing Games:** These games are a model of interdomain routing in the Internet, where best-reply dynamics model the Border Gateway Protocol (BGP). The setting assumes strategic nodes in a directed graph, each wishing to establish a path in the network from itself to a given fixed target node d. Each node has its own preference order over all possible paths from itself to d (a preference that may depend on various network parameters, as well as on various business parameters). Each node's action is choosing a single outgoing edge. The collection of these outgoing edges determines

[1] The study of TCP from a game-theory/mechanism design perspective is a long-standing agenda [37, 41, 2].

a unique path (or none) from each node to d. The networking community has spent considerable effort into understanding the convergence of BGP and has come up with the well-known Gao-Rexford commercial constraints [18]. As a corollary of our results for routing games we get the following:

Theorem: *Under the Gao-Rexford constraints, best-reply dynamics (BGP) are incentive-compatible and converge to a stable outcome even in asynchronous Internet-like settings.*

Our proof greatly simplifies, and unifies, the proofs of the recent results of Levin et al. [28], and of well-known results regarding BGP convergence in networking literature [22, 18].

- **Cost-Sharing Games:** Cost-sharing problems arise in situations in which we wish to distribute the cost of some public service (e.g., building a bridge) between self-interested users that will benefit from this service in different extents. A cost-sharing problem is defined by a cost function C, that specifies, for every subset of users, the cost of providing the public service to these users. Every user is assumed to have a private value for being serviced. A cost-sharing mechanism outputs, for every possible combination of private values of the users, a set of users that may benefit from the public good, and the way that the cost will be divided between these users. A well-known family of truthful cost-sharing mechanisms are the Moulin mechanisms [33, 34].In [29], a more general family of truthful cost-sharing mechanisms, called "acyclic mechanisms" is presented. These mechanisms have various advantages over Moulin mechanisms. In particular, they follow in a generic way from off-the-shelf primal-dual algorithms. We wish to compute these famous cost-sharing mechanisms in a distributed, incentive-compatible fashion (by the users themselves). We show that the distributed protocol based on best-reply dynamics does exactly that:

  Theorem: *In cost-sharing games, best-reply dynamics are incentive-compatible and implement the outcome of acyclic mechanisms.*

- **Stable-Roommates Games:** This well known example [17] considers how to pair students (say, for the purpose of sharing a dorm room), when each student has his own preference order over possible roommates. The classic goal is to find a "stable matching" where no two students prefer each other over their assigned roommates. A natural mechanism stipulates that students repeatedly propose to their most preferred roommate among those that do not immediately reject them, and immediately reject all proposers except for their most preferred one so far. Is this natural mechanism incentive-compatible? We show that this setting can be formalized as a game, in which best-reply dynamics correspond to this protocol, such that:

  Theorem: *For the two well-known cases described below, best-reply dynamics converge to a stable matching in an incentive-compatible manner*[2].

  - **Intern-Assignment Games:** This example considers how to assign interns to hospitals. Each intern has his own preference order over possible hospitals. The hospitals have a common way to rank interns (say, according to their GPA).
  - **Correlated Two-Sided Markets:** These games are inspired by real-life one-sided market games in which players have preferences about a set of markets, and the preferences of markets are correlated with the preferences of players [1] (e.g., market sharing games,

[2]Unlike in the related stable-marriage problem [17], in this unisex variant it is known that stable matchings need not exist in general.

and distributed caching games). The game is represented by a weighted bipartite graph in which vertices on one side represent players (buyers), and vertices on the other side represent markets (sellers). Each buyer has a preference order over sellers, such that a seller A is ranked higher than another seller B by that buyer if the edge connecting that buyer to A has a bigger weight than the edge connecting the buyer to B. Similarly, sellers prefer buyers that are connected to them by heavier edges.

- **Iterative Auctions:**
  - **Iterative Auctions with Unit-Demand Bidders:** In this classic setting, a set of items is sold to a group of bidders (with discretized bids). Each bidder has a private value for each item and is only interested in receiving one item. The objective is to partition the items between the bidders so as to maximize the social welfare (the sum of bidders' values for their allocated items). This setting is the subject of extensive study in economics literature. The canonical one-shot VCG mechanism asks each bidder for his values, allocates items optimally, and charges each bidder a (VCG) payment so that his net utility is his marginal product. Many researchers seek to implement the VCG outcome in dynamic settings, using an iterative auction procedure. In particular, it is well-known that the widely-used ascending English auction [24], in which the prices of items continuously rise until only all bidders but one drop out of the competition for each item, converges to the VCG allocation and prices and is incentive-compatible.

    We formulate a dynamic game in which bidders repeatedly change their bid in response to the other bids, until no further changes are desired. In this game a bidder that is allocated an item pays his bid for the item (as in a 1st-price auction). The ascending auction can be regarded as a special case of best-reply dynamics in our game.

    Theorem: *Best-reply dynamics (with specific, carefully chosen, tie-breaking rules) are incentive-compatible and always converge to the VCG allocation and prices.*

    Our proof technique relies on the famous connection [5] between these auctions and primal-dual-based algorithms for computing maximal-weighted-matchings. Specifically, we show that best-reply dynamics can, in some sense, mimic Kuhn's Hungarian Method [25]. Our results imply a natural family of iterative incentive-compatible auctions that can be computed in a distributed fashion. This generalizes well-known auctions like the ascending English auction and the descending-price Dutch auction [30].
  - **Adword Auctions:** Sponsored search advertising is a major source of revenue for Internet search engines (close to 98 percent of Googles total revenue of 6 billion for the year 2005 came from sponsored search advertisements). In the adword auctions setting [10, 42] there are k slots that are to be sold to different advertisers. Each slot has a clique-through-rate (CTR) (higher slots are preferred to lower ones). Each bidder (advertiser) has a private value per click. The auction rule used by companies like Google and Yahoo! is the *generalized second-price auction* (GSP). GSP assigns the k slots to the k highest bidders (the highest bidder gets the highest slot, etc.) and charges each winning bidder a cost per-click that equals the bid of the bidder that was assigned the slot below his.

    Adword auctions are dynamic in nature – the advertisers are allowed to change their bids quite frequently. It is known that a version of the ascending English auction always reaches a specific, economically reasonable, allocation of slots in an incentive-compatible manner [10]. The generalized English auction can be regarded as a special case of best-reply dynamics. However, the more general (and arguably more realistic) case is much more problematic; it was shown by Cary et al. [8] that best-reply dynamics in this setting are not even always guaranteed to converge. However, [8] shows that this problem can be overcome by using randomness. We prove that this randomized convergence of best-reply dynamics is also incentive-compatible.

Theorem: *The randomized best-reply mechanism in [8] is incentive-compatible.*

1.2 Related Work

Best-reply dynamics are a subject of a vast body of literature in economics and computer science. Many of these works are concerned with the convergence of best-reply dynamics to pure Nash equilibria (see Rosenthal [39] and Monderer and Shapley [31], and also [12, 19, 1]).

Unlike these previous works, our primary concern is the incentive-compatibility of best-reply dynamics. This problem has so far been considered in very few isolated contexts. In particular, it has been observed by Fabrikant and Papdimitriou [11] and by Levin et al. [28] that BGP can be regarded as best-reply dynamics, and its incentive-compatibility was studied in a series of papers by Feigenbuam et al. (see [13, 14]) and recently in [28].

To the best of our knowledge, our work is a first attempt at analyzing the underlying structure of games for which best-reply dynamics are strategically justified. The works closest in telos to ours are those about "learning equilibria", presented by Brafman and Tennenholtz in [6, 7], and later studied in [4, 3, 32]. However, the learning equilibrium framework put forth in these papers is different from ours in that it does not consider best-reply dynamics, and takes a repeated-games approach.

1.3 Organization of the Paper

In Section 2 we present our formal framework. In Section 3 we illustrate our framework via the routing games example, and define universally-max-solvable games. In Section 4 we present a non-universally-max-solvable example of best-reply mechanisms in auction settings. Finally, in Section 5 we present open problems and directions for future research.

2 Best-Reply Dynamics and Mechanism Design

In this section we present our formal framework of best-reply mechanisms (in Appendix A we explain how our framework can be extended to incorporate randomness, and to handle realistic asynchronous environments).

Let G be a *game with private information*, in which players have *action sets* $A_1, \ldots, A_n$. Each player i has a utility function u_i that assigns a value to each action-profile, and is his private information. $u = (u_1, \ldots, u_n)$ belongs to a set of utility-function profiles U. An action a_i is said to be a best-reply of i to the actions of the other players a_{-i} if it maximizes u_i, given a_{-i}.

Our results will be that, for certain private-information games, best-replying is incentive compatible. This turns out to be true in quite general settings, but for clarity of presentation we will start by considering a simple centralized and synchronous model. Our model can be extended to distributed settings (with no computational centre) with enough rounds of interaction (even asynchronous) as well as to settings where non-convergence is (or can be made) undesired by the players (e.g., just because of time-value). The only significant properties of this model for our results are:

1. That there be a sufficient number of "rounds" for best-reply dynamics to converge to a pure Nash equilibrium.

2. Ensuring that "non-convergence" is not in the best interest of any player.

Our model captures the first point by simply considering a synchronous round-robin protocol, and the second point by enforcing a fixed "penalty" outcome if convergence is not reached by a predefined deadline.

Definition 1 *A penalty in a game with private-information G is an action-profile a such that for all $u = (u_1, ..., u_n) \in U$, for every player i, and for every action profile a', it hold that $u_i(a) \leq u_i(a')$.*

Definition 2 *A repeated-reply mechanism for a game with private information G, is an extensive form game with perfect recall, that has the same players as G, and consists of at most M steps. In each step of the game, a single player announces an element of A_i. Players play in round-robin order (for some fixed order on players). This goes on until all players "pass" in n consecutive steps (we say that a player "passed" in step i if he announced the same action in step i and in step $i-n$). In this case the mechanism enforces the the action profile of the most recently announced actions of the players. The payoff of a player in this case is his utility for this action-profile in G. If M steps go by and this does not occur, then the mechanism penalizes the players.*

Definition 3 *The best-reply strategy of player i in a repeated-reply mechanism (starting from an initial action profile $a = (a_1, \ldots, a_n)$) is to repeatedly announce the best-reply to the most recently announced actions of the other players (or to a in the first $n-1$ steps).*

Best-reply dynamics is the strategy-profile in which all players play best-reply strategies. Best-reply dynamics are said to be incentive-compatible in ex-post Nash if it is an ex-post Nash equilibrium [38] in a repeated-reply mechanism. Ex-post Nash in this case roughly means that if all other players are playing best-reply strategies, then a player cannot gain by not playing a best-reply strategy (this is true for *every* possible realization of the private-information of the players). Best-reply dynamics are collusion-proof in ex-post Nash [15] if no coalition of players of any size can deviate from it and strictly improve the outcome of one member in the coalition without strictly harming another. We shall refer to a repeated-reply mechanism as a "best-reply mechanism" if the prescribed strategy-profile [38] (that the players are instructed to follow) is best-reply dynamics.

3 The Universal-Max-Solvability of Routing Games (and More)

We shall now present in detail the routing games example discussed in the Introduction, and prove that it belongs to the class of universally-max-solvable games. This immediately implies classic convergence results in networking literature [22, 18], as well as the incentive-compatibility results in [28]. In Appendix C we exploit universal-max-solvability to prove similar results for many other games.

3.1 Illustrative Example: Routing Games

The Game. A *routing game* [28], is played over a graph G with vertex set V, and an edge set E. The players are the vertices, and each vertex has a private full order of *strict* preferences over all the simple paths in the graph between itself and a common fixed target node d. Each vertex must choose an outgoing edge representing a choice of neighbouring node to forward traffic to. Vertices are selfish and wish to be assigned as highly ranked routes as possible (given their preferences).

We consider the well-known *"No Dispute Wheel"* [22] condition, which is known to generalize the Gao-Rexford setting [18] (see [28] and references therein).

The Mechanism. We go over the vertices in round-robin order. In each step, a single vertex chooses an outgoing edge. This continues until all vertices do not switch their outgoing edges (i.e., repeat their previous edge-choices). Then, the outputted route assignment consists of the final outgoing edges. If after $|V|^3 - 2|V|^2 + |V|$ rounds this does not occur, then no vertex is assigned an outgoing edge (no traffic is sent at all).

The Prescribed Best-Reply Strategies. In each step, a vertex a should choose a single outgoing edge such that the route from a to d induced by this choice (given current choices of other vertices) is most preferred by a (if no such route is induced then a does not choose an outgoing edge at all).

Consider the following theorem, due to Levin et al. [28]:

Theorem 3.1 *[28] If No Dispute Wheel holds, then the prescribed best-reply strategies are incentive-compatible and converge to a unique pure Nash equilibrium.*

In fact, Theorem 3.1 can be shown to hold even in a realistic Internet-like asynchronous setting in which vertices can be activated simultaneously and in any order, vertices learn of each others' action via update messages that can be delayed or lost, and there is no central entity that can penalize vertices in case of non-convergence. (We present it in its current form to simplify the exposition of our approach).

Next, we discuss the underlying structure of routing games that implies the convergence and incentive-compatibility of best-reply dynamics. We exhibit a subclass of max-solvable games [36] of which routing games are a special case. This generalizes the recent results of Levin et al. [28] and places them in a much broader context (as well as provides a much simpler proof for the main theorem in [28]). It also generalizes the known convergence proofs for BGP presented in [22, 18].

3.2 Universally-Max-Solvable Games

We now present the class of universally-max-solvable games and its connection to best-reply mechanisms. In Appendix B we present the proofs for the statements in this section.

Full-information games. We define a subclass of *full-information games* called *universally max-solvable games.*

Definition 4 *Let S_i be the set of player i's strategies. We say that a set $T \subset S_i$ is universally max-dominated if $T \neq S_i$, $T \neq \emptyset$, and $max_{s_{-i},s_i \in T} u_i(s_i, s_{-i}) < min_{s_{-i},s'_i \in S_i \setminus T} u_i(s'_i, s_{-i})$.*

Intuitively, a subset of player i's strategies is universally max-dominated if the maximal utility player i can derive from playing a strategy in it is less than the minimal utility he can derive from playing a strategy that is not in it.

Definition 5 *A game G is said to be universally max-solvable if there is a sequence of games $G_0, \ldots, G_r$ such that:*

- $G_0 = G$
- *For every $k \in \{0, \ldots, = r-1\}$, G_{k+1} is a subgame of G_k achieved by removing a universally max-dominated strategy-set from the strategy space of one player in G_k.*

- *The strategy space of each player in G_r is of size* 1.

We shall refer to an elimination order of universally-max-dominated strategy-sets, that results in a single strategy-profile as an *elimination sequence* of a universally-max-solvable game.

Definition 6 *A strategy profile $s^* = (s_1^*, \ldots, s_n^*)$ is said to be collusion-proof if the following holds: For any subset of players T, there is no pure strategy-profile $s = (s_1, \ldots, s_n)$ such that for all $i \notin T$ $s_i = s_i^*$, for some $j \in T$ $u_i(s) > u_i(s^*)$, and for all $i \neq j \in T$ $u_i(s) \geq u_i(s^*)$.*

Proposition 3.2 *A universally max-solvable game has a unique pure Nash equilibrium. This pure Nash equilibrium is collusion-proof.*

In particular, collusion-proofness holds for the coalition of all players. Hence, we get the following corollary:

Corollary 3.3 *The unique pure Nash equilibrium of a universally-max-solvable game is Pareto optimal.*

Private-information games. A game with private-information G is said to be universally-max-solvable if for every $u \in U$ the full-information game induced by players' actual utilities is universally-max-solvable.

Theorem 3.4 *If G is a universally-max-solvable game with private information, then best-reply dynamics are incentive-compatible in ex-post Nash (and even collusion-proof in ex-post Nash) for any best-reply mechanism that consists of at least $n(\Sigma_i m_i)$ steps, where n is the number of players, and m_i is the size of the action space of player i. The outcome of best-reply dynamics is the unique pure Nash equilibrium of the full-information game induced by G.*

The proof of this theorem can easily be extended to prove analogous results for randomized best-reply mechanisms and for asynchronous best-reply mechanisms. Observe that the definition of universal-max-domination involved (for ease of exposition) strict inequalities. However, all of our definitions easily carry over to the case of weak inequalities, as long as the prescribed best-reply strategies break ties in a way that is consistent with the order of an elimination sequence. In this case, there might be more than one pure Nash equilibrium, yet best-reply dynamics will still converge to a specific one. Also, in some cases (as is shown in Appendix C) the definition of universally-max-solvable games can be extended to games in which the strategy-space of the players is not finite.

3.3 Applications

Theorem 3.1 regarding routing games follows immediately from the following lemma (proven in Appendix C):

Lemma 1 *Routing games are universally-max-solvable.*

The proof of Lemma 1 explicitly constructs an elimination sequence of universally-max-dominated strategies of the appropriate size.

In Appendix C we present many more incentive-compatible mechanisms for interesting settings using these techniques, specifically we design mechanisms for:

- Iterative single-item auctions.
- Congestion-control games (inspired by TCP).
- Cost-sharing problems.
- Two stable-roommates games.

4 Beyond Universal-Max-Solvability

Many games *do not* fall into the category of universally-max-solvable games. We present an illustrative example of such games in the form of auctions with unit-demand bidders, and prove that despite its non-universal-max-solvability, best-reply dynamics converge and are incentive-compatible in these games. This necessitates inherently different proof techniques. Our result for auctions with unit-demand bidders generalizes classic results in economics regarding ascending and descending auctions by exhibiting a natural family of incentive-compatible iterative auctions. In Appendix D we present another example (discussed in the Introduction) – adword auctions.

4.1 Illustrative Example: Iterative Auctions with Unit-Demand Bidders

The Game. There are m items and n bidders. Each bidder i has a value $v_{i,j}$ for every item j, and is interested in getting only one item. Each bidder submits a vector of nonnegative bids, one for each item, where each bid is chosen from some fixed discrete set of bids, say, multiples of some arbitrarily small $\epsilon > 0$. Each bidder must also specify a single item that shall be referred to as his "*target* item". To explain the rule for allocating (single) items to the different bidders we present a few classic definitions that capture the intuitive notion of a set of items that are actively being competed over.

Definition 7 *Given a vector of item prices $\overrightarrow{p} = (p_1, \ldots, p_m)$, a bidder i's* most-demanded *set of items with respect to $\overrightarrow{p}$ is the set $D_i(p) = \{j |\ j \in argmax_{j'}\ v_{i,j'} - p_{j'}\}$.*

Definition 8 *Given a price vector p^*, an* overdemanded set *is a set of items X such that (at prices p^*) there is a set of bidders Y with $D_i(p^*) \subseteq X$ for all $i \in Y$ and $|Y| > |X|$.*

Namely, a set of items X is overdemanded if there are more than $|X|$ players each of whose set of most-demanded items lies entirely in X.

Definition 9 *We say that such a set of items X is a* minimal *overdemanded set if no subset of X is overdemanded.*

Definition 10 *The vector of prices induced by an n-tuple of bid vectors is the price vector in which the i's coordinate is the maximal bid for item i (over all bidders).*

We can now define the allocation-rule: if there is no "overdemanded set", given the vector of prices induced by the current bids (in that more than k bidders all have highest bids among some set of k bids), then there is an assignment in which each bidder with a highest bid is allocated an item for which he bid highest. Among allocations of the above form, the allocation of items chosen will be such that the number of bidders receiving their target items is minimized. Each bidder is then charged the amount he bid for the item he got. If there is no such allocation, then the bidders/items in minimal overdemanded sets are left unassigned, and all other bidders are assigned items as above.

The Mechanism. In some prescribed round-robin order, bidders repeatedly announce their bid vectors, together with a *target item* and are allocated items (temporarily) as defined above, until all bidders "pass" (repeat the same vector of bids), at which point the final bids and corresponding item-allocation are used. If there is no agreement after $nm \max_{i,j} v_{i,j}/\epsilon$ rounds then no items are allocated.

The Prescribed Best-Reply Strategies. For bidder i's turn, let b_j^* denote the highest bid for item j provided by one of the other bidders (and recall that $v_{i,j}$ denotes bidder i's private value for item j). Bidder i is instructed to bid as follows:

- If, for all j, $v_{i,j} \leq b_j^* + \epsilon$, then we say that bidder i has no demanded items at prices b^*, and bidder i places bids of $v_{i,j} - \epsilon$.
- If there is an item j that is in bidder i's most-demanded set such that with bidder i bidding b_j^* item j is not in a minimal overdemanded set, then bidder i bids b_j^* for item j, and for all items $j' \neq j, \max\{v_{i,j'} - v_{i,j} + b_j^* - \epsilon, 0\}$. Additionally, bidder i specifies item j as his target item.
- Finally, if for each demanded item, by bidding b_j^* item j will be in a minimal overdemanded set, i bids $b_j^* + \epsilon$ for a demanded item, and $\max\{v_{i,j'} - v_{i,j} + b_j^* + \epsilon, 0\}$ for all other items, and specifies item j as its target item.

Theorem 4.1 *The prescribed best-reply strategies are incentive compatible and converge to the pure Nash equilibrium in which items are allocated as in the maximal-weighted-matching and prices are the corresponding VCG prices.*

The proof of Thereom 4.1 can be found in Appendix D. Informally speaking, we prove that the prescribed best-reply strategies (with our carefully chosen tie-breaking rules) converge to the VCG allocation and prices (from *all* possible initial bids) by interpreting the prescribed best-reply strategies as a form of Kuhn's Hungarian Method for computing optimal weighted matchings in a bipartite graph [25].

5 Discussion and Open Questions

We have presented a class of private-information games, namely "universally-max-solvable games", for which best-reply dynamics are guaranteed to converge in an incentive-compatible manner. We have shown that a hoard of interesting examples fall within this restricted category of games. We believe that many other settings can be formulated as universally-max-solvable games (thus leading to new incentive-compatible algorithms).

We also exhibit two auction settings that are not in this class of games but for which best-reply dynamics converge in an incentive-compatible manner. The obvious open question is identifying structures more general than universal-max-solvability that lead to such results. We note that the incentive-compatibility of both examples is fundamentally intertwined with the properties of VCG and competitive equilibria. Unraveling this connection for broader environments might be a first step in this direction.

The study of game dynamics from a mechanism design perspective is a new and challenging research direction. Questions similar to the ones addressed in this paper can be asked regarding better-reply dynamics, and other types of dynamics.

Acknowledgements

We thank Tim Roughgarden for helpful discussions and for pointing our attention to the connection between our work and cost-sharing mechanisms.

References

[1] H. Ackermann, P. Goldberg, V. Mirrokni, H. Roeglin, and B. Voecking. Uncoordinated two-sided markets. In *EC*, 2008.

[2] Aditya Akella, Srinivasan Seshan, Richard Karp, Scott Shenker, and Christos Papadimitriou. Selfish behavior and stability of the internet:: a game-theoretic analysis of tcp. In *SIGCOMM '02: Proceedings of the 2002 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 117–130, New York, NY, USA, 2002. ACM.

[3] Itai Ashlagi, Dov Monderer, and Moshe Tennenholtz. Learning equilibrium in resource selection games. In *AAAI*, pages 18–23, 2007.

[4] Itai Ashlagi, Moshe Tennenholtz, and Dov Monderer. Robust learning equilibrium. In *UAI*, 2006.

[5] Sushil Bikhchandani, Sven de Vries, James Schummer, and Rakesh Vohra. Ascending auctions and linear programming. Discussion Papers 1368, Northwestern University, Center for Mathematical Studies in Economics and Management Science, April 2003.

[6] Ronen I. Brafman and Moshe Tennenholtz. Efficient learning equilibrium. *Artif. Intell.*, 159(1-2):27–47, 2004.

[7] Ronen I. Brafman and Moshe Tennenholtz. Optimal efficient learning equilibrium: Imperfect monitoring in symmetric games. In *In Proceedings of the National Conference on Artificial Intelligence (AAAI*, pages 726–731. Press, 2005.

[8] Matthew Cary, Aparna Das, Benjamin Edelman, Ioannis Giotis, Kurtis Heimerl, Anna R. Karlin, Claire Mathieu, and Michael Schwarz. Greedy bidding strategies for keyword auctions. In *ACM Conference on Electronic Commerce*, pages 262–271, 2007.

[9] A. Demers, S. Keshav, and S. Shenker. Analysis and simulation of a fair queueing algorithm. In *SIGCOMM '89: Symposium proceedings on Communications architectures & protocols*, 1989.

[10] Benjamin Edelman, Michael Ostrovsky, and Michael Schwarz. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American Economic Review*, 97(1):242–259, March 2007.

[11] Alex Fabrikant and Christos Papadimitriou. The complexity of game dynamics: BGP oscillations, sink equlibria, and beyond. In *Proceedings of SODA 2008.*

[12] Alex Fabrikant, Christos Papadimitriou, and Kunal Talwar. The complexity of pure nash equilibria. In *Proceedings of STOC 04'*, 2006.

[13] Joan Feigenbaum, Christos H. Papadimitriou, Rahul Sami, and Scott Shenker. A BGP-based mechanism for lowest-cost routing. *Distributed Computing*, 18(1):61–72, 2005.

[14] Joan Feigenbaum, Vijay Ramachandran, and Michael Schapira. Incentive-compatible interdomain routing. In *7th Conference on Electronic Commerce*, pages 130–139, New York, 2006. ACM.

[15] Joan Feigenbaum, Michael Schapira, and Scott Shenker. *Distributed Algorithmic Mechanism Design. A chapter in 'Algorithmic Game Theory"*. Cambridge University Press, 2007.

[16] Amos Fiat, Andrew V. Goldberg, Jason D. Hartline, and Anna R. Karlin. Competitive generalized auctions. In *34th Symposium on Theory of Computing*, pages 72–81, New York, 2002. ACM.

[17] D. Gale and L.S. Shapley. College admissions and stability of marriage. *Amer. Math. Monthly*, (69):9–15, 1962.

[18] Lixin Gao and Jennifer Rexford. Stable Internet routing without global coordination. *IEEE/ACM Transactions on Networking*, 9(6):681–692, 2001.

[19] Michel Goemans, Vahab Mirrokni, and Adrian Vetta. Sink equilibria and convergence. In *FOCS '05: Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 142–154, Washington, DC, USA, 2005. IEEE Computer Society.

[20] A. V. Goldberg, J. D. Hartline, A. R. Karlin, M. Saks, and A. Wright. Competitive auctions. *Games and Economic Behavior*, 55(2):242 – 269, 2006.

[21] Timothy G. Griffin, F. Bruce Shepherd, and Gordon Wilfong. Policy disputes in path-vector protocols. In *7th International Conference on Network Protocols*, pages 21–30, Los Alamitos, 1999. IEEE Computer Society.

[22] Timothy G. Griffin, F. Bruce Shepherd, and Gordon Wilfong. The stable paths problem and interdomain routing. *IEEE/ACM Transactions on Networking*, 10(2):232–243, April 2002.

[23] Faruk Gul and Ennio Stacchetti. Walrasian equilibrium with gross substitutes. *Journal of Economic Theory*, 87(1):95–124, July 1999.

[24] Faruk Gul and Ennio Stacchetti. The english auction with differentiated commodities. *Journal of Economic Theory*, 92(1):66–95, May 2000.

[25] Harold W. Kuhn. The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2:83–97, 1955.

[26] Ron Lavi and Chaitanya Swamy. Truthful and near-optimal mechanism design via linear programming. In *FOCS*, pages 595–604, 2005.

[27] Herman B Leonard. Elicitation of honest preferences for the assignment of individuals to positions. *Journal of Political Economy*, 91(3):461–79, June 1983.

[28] Hagay Levin, Michael Schapira, and Aviv Zohar. Interdomain routing and games. In *STOC*, pages 57–66, 2008.

[29] Aranyak Mehta, Tim Roughgarden, and Mukund Sundararajan. Beyond moulin mechanisms. In *EC '07: Proceedings of the 8th ACM conference on Electronic commerce*, pages 1–10, New York, NY, USA, 2007. ACM.

[30] Debasis Mishra and David C. Parkes. Multi-item Vickrey-Dutch auctions. *Games and Economic Behavior*, 2007. Submitted, under review.

[31] Dov Monderer and L.S. Shapley. Potential games. *Games and Economic Behavior*, (14):124–143, 1996.

[32] Dov Monderer and Moshe Tennenholtz. Learning equilibrium as a generalization of learning to optimize. *Artif. Intell.*, 171(7):448–452, 2007.

[33] Herv Moulin. Incremental cost sharing: Characterization by coalition strategy-proofness. *Social Choice and Welfare*, 16(2):279–320, 1999.

[34] Herv Moulin and Scott Shenker. research article : Strategyproof sharing of submodular costs:budget balance versus efficiency. *Economic Theory*, 18(3):511–533, 2001.

[35] Noam Nisan and Amir Ronen. Algorithmic mechanism design. *Games and Economic Behavior*, 35(1):166–196, 2001.

[36] Noam Nisan, Michael Schapira, and Aviv Zohar. Asynchronous best-reply dynamics. In *Proceedings of WINE 2008*.

[37] Christos H. Papadimitriou. Algorithms, games and the internet. In *In STOC*, pages 749–753. ACM Press, 2001.

[38] David C. Parkes and Jeffrey Shneidman. Specification faithfulness in networks with rational nodes. In *23rd Symposium on Principles of Distributed Computing*, pages 88–97, New York, 2004. ACM.

[39] R. W. Rosenthal. A class of games possessing pure-strategy nash equilibria. *Int. J. Game Theory*, (2):65–67, 1973.

[40] Lloyd S. Shapley and Martin Shubik. The assignment game I: The core. *International Journal of Game Theory*, 1(2):111130, 1972.

[41] Scott Shenker. Open problems in distributed algorithmic mechanism design. A presentation (cowritten with Joan Feigenbaum), DIMACS Workshop of Computational Issues in Game Theory and Mechanism Design, 2001.

[42] Hal R. Varian. Position auctions. *International Journal of Industrial Organization*, 25(6):1163–1178, December 2007.

A Best-Reply Mechanisms – Extensions

A.1 Randomized Best-Reply Mechanisms

The definition of a randomized best-reply mechanism is almost identical to that of a best-reply mechanism. The only difference is that while in best-reply mechanisms players are chosen to play in some round-robin order, in a randomized best-reply mechanism one player chosen uniformly at random may update his action in each step.

A.2 Asynchronous Best-Reply Mechanisms

So far, our model considered best-reply dynamics in settings in which one can assume the presence of a central entity that synchronizes players' actions (make sure that players play in round-robin order, or choose one player to play in each step at random). However, in some of the settings considered in this paper (inspired by the Internet, or by market dynamics) no such entity exists (asynchronous environments). In such settings, we can think of the players as residing in a computer network, where their best-replies are transmitted to other players and serve as the basis for the other players' best-replies (as is often the case in the Internet [21]). In this asynchronous environment, several players may choose their best-replies simultaneously. More generally, messages may be delayed or lost and players may choose their best-replies based on information that is out of date and unsynchronized with others.

To analyze best-reply dynamics in asynchronous settings, we change our definition of a repeated-reply mechanism: An *asynchronous repeated-reply mechanism* for a game with private information G, is still defined as an extensive form game with perfect recall, that has the same players as G. However, now the repeated-reply mechanism has an *infinite* number of steps. In each step one *or more* players are chosen to participate by an adversarial entity called the *Scheduler* [28]. A player i chosen to participate can make one or more of the following moves:

- Receive update messages from other players.
- Choose an element of A_i (this represents the action i actually chooses).
- Send an update message to another player announcing an element of A_i (this represents the action i announces to the other player).

The scheduler is in charge of making choosing the initial action profile of the players $a = (a_1, \ldots, a_n)$, determining which players participate in each step, and determining when sent update messages reach their destinations. The scheduler is restricted not to indefinitely starve any player from best-replying (that is, each player participates in infinitely many rounds). We name all choices made by the scheduler a *schedule.*

A strategy of a player in an asynchronous repeated-reply mechanism specifies the moves that a player makes in each step (as a function of his private information and the history of moves). The best-reply strategy of a player i is to repeatedly:

- Read update messages from all players.
- Choose the best-reply a_i to the most recent actions of the other players *that were announced to i.*
- Announce a_i to all other players.

The payoff of a player i in an asynchronous repeated-reply mechanism (given a specific schedule) is defined to be $u_i(a)$, where $a = (a_1, \ldots, a_n)$ is the profile of *actual* players' actions that appears in infinitely many rounds, and is assigned the highest value by u_i (this liberal definition of a player's payoff in the case that there was no convergence to a single action profile only strengthens our results).

B Proofs for Section 2

Proposition B.1 *A universally max-solvable game has a unique pure Nash equilibrium. This pure Nash equilibrium is collusion-proof.*

Proof: By contradiction. Fix an elimination sequence of universally-max-dominated strategy-sets for the game (by definition, one exists). Let s^* be the strategy-profile reached after eliminating all the universally-max-dominated strategies in the elimination sequence. Let s be a strategy-profile such that s can be reached from s^* by changing the strategies of of some subset of the players T, for every $i \in T$ $u_i(s) \geq u_i(s^*)$ and for at least one player in T this inequality is strict. Let s_i be the strategy in s that is eliminated first (given this elimination sequence) of a player $i \in T$. Consider the elimination step in which s_i is eliminated. The strategy-set containing s_i is universally-max-dominated. This means that any strategy outside this set *always* guarantees i a *strictly* higher utility (given that all previous strategy-sets in the elimination sequence are removed). In particular, this is true for s_i^*. However, observe that both s and s^* do not contain any strategies that were removed in the elimination sequence prior to this stage. This contradicts the fact that $u_i(s) \geq u_i(s^*)$. □

Theorem B.2 *If G is a universally-max-solvable game with private information, then best-reply dynamics are incentive-compatible in ex-post Nash (and even collusion-proof in ex-post Nash) for any best-reply mechanism that consists of at least $n(\Sigma_i m_i)$ steps, where n is the number of players, and m_i is the size of the action space of player i. The outcome of best-reply dynamics is the unique pure Nash equilibrium of the full-information game induced by G.*

Proof: We begin with a definition of a *phase* – a period of time in which every player has been activated and allowed to play at least once (In our current setting, as players are activated in round-robin order, a phase is a period of exactly n rounds. However, in asynchronous settings a phase can be a different period of time).

Let us examine the outcome that is reached if all players act according to a best reply strategy. Since the game is universally max-solvable, there exists a sequence of eliminations of universally-max-dominated actions that eventually leaves us with a single action profile. Let a_1 be the first action that is eliminated, and let i be the player to whom this action belongs. Player i is activated during the first phase, and changes his action to be the best-reply to the current action profile of the other players. Note, that no matter what the other players have chosen, the action a_1 is never the best-reply. By definition, there is always a different action (that may depend on what the other players are playing) that yields a higher payoff. Therefore, the strategy a_1 will no longer be played by player i after the first phase. From this point onwards, we are effectively playing in the game G_1 (in which strategy 1 has been eliminated). In this game, there is a universally-max-dominated strategy a_2 (that may belong to a different player) that is eliminated by the second phase for similar reasons.

More generally, in phase j, given that action $a_1, a_2, \ldots a_{j-1}$ are no longer played, the j'th action in the elimination order a_j will be considered by its player and will not be chosen as a best reply from that moment on (since other previous strategies are no longer played, no situation will ever arise in which a_j is the best-response). After $\sum_i (m_i - 1)$ phases, all of the actions in the elimination order are no longer played. The surviving action profile is then the Nash equilibrium of the original game, and no player deviates from it. This ends the game and sets this action profile as the outcome (and no penalty is activated). Since each phase lasts n rounds, we are assured of convergence within $(n-1)\sum_i m_i$ rounds.

Now we shall consider the case in which some group of players $I \subseteq [n]$ decides not to follow the best reply strategy. Since the mechanism gives out a penalty to the players if they did not converge after enough rounds of play, and because this penalty is strictly worse for any player than the payoff if gets if all follow best-reply, it cannot be that the group I gains from a situation in which there is no convergence. Therefore, we assume from this point on that the dynamics converged to an action profile $a = (a_1, a_2, \ldots, a_n)$. Specifically, the players in $[n] \setminus I$ are stable in the action profile a, and their action is already a best reply to the actions of the others. If a is the Nash equilibrium action profile, then the players in I have not gained from deviating. If it is not that action profile, then there are actions being played by some players in a that appear in the elimination sequence. let j be the player whose action is the earliest to be eliminated among all actions in a. Since no action before a_j in the elimination sequence is played by any of the players, a_j cannot be a best response to the current action profile. Therefore, player j is one of the members of the group I (all other players are stable). However, because a_j is universally-max-dominated, its best payment (after the elimination of all actions before it in the elimination sequence) is strictly lower than the payment of all other remaining strategies for player j, which specifically include the payment player j would get in the Nash equilibrium profile. Therefore player j loses from deviating from the best-reply strategy. This shows that any group of players that deviates must contain a player that loses by not playing best-reply, and so the best reply strategy profile is stable. □

C Examples of Universally-Max-Solvable Games

In this section we rely on the previous section to provide various examples of best-reply mechanisms by exhibiting interesting games that are universally-max-solvable.

C.1 Internet Protocols

C.1.1 Congestion-Control Games

The Game. There is a directed graph $G =< V, E >$ in which the vertices represent non-strategic routers, and the edges represent communication links between them. There is a capacity function c that assigns every $e \in E$ a capacity $c(e) \in (0, 1]$. The players are n *flows*. Each flow is represented by a route r_i between a pair of nodes in V. Each flow i must choose a transmission rate $a_i \in [0, d_i]$, where d_i is a private value that represents i's maximal transmission rate.

The routers share the capacity of the edges between the flows in the following way: Let $e = (a, b)$ be an edge that appears on the routes of k flows. If the a_is of all k flows are at least $\frac{c(e)}{k}$ then each of the k flows will get an equal capacity-share of $\frac{c(e)}{k}$. Otherwise, the flow with the minimal transmission rate a_j gets a share of a_j and the remaining capacity of $c(e) - a_j$ is shared between the $k - 1$ remaining flows in a similar manner (with $c(e)$ substituted by $c(e) - a_j$ and k substituted by $k - 1$). This corresponds to what is known as "Weighted Fair Queueing" in networking literature.

The *throughput* of a flow is the value of its minimal capacity-share, taken over all the edges its route traverses. Flows are selfish and wish to maximize their throughput.

The Mechanism: Flows repeatedly, one by one and in round-robin order, choose transmission rates. This goes on until all flows do not change their transmission rates (i.e., repeat their previous choices of transmission rates), at which point the final transmission-rates are used. If after $2n(n + |E|) + n$ rounds this does not occur then no flow is allowed to transmit at all.

The Prescribed Best-Reply Strategies: In each step, flow i chooses the *lowest* transmission-rate that maximizes its throughput (given the current transmission rates of the other flows).

Theorem C.1 *The prescribed best-reply strategies are incentive compatible and converge to a pure Nash equilibrium for which the max-min fairness value is optimized. This holds even in asynchronous settings.*

Theorem C.1 follows from the following lemma:

Lemma 2 *Congestion-control games are universally-max-solvable.*

Proof: A congestion-control game is a game with private information in which the actions of the players (flows) are their transmission rates, and their utilities are their throughput values (that cannot exceed their d_is). Observe, that not allowing flows to transmit is a penalty. We consider a full-information game induced by the actual utilities of the players, and show that it is universally-max-solvable. We do so by exhibiting an elimination sequence of universally-max-dominated action sets. We later discuss how to overcome the fact that the action sets are infinite.

Let e be an edge for which $\frac{c(e)}{k_e}$ is minimized, where k_e is the number of flows whose routes go through e. Consider the flow i with the lowest d_i whose route goes through e. If $d_i < \frac{c(e)}{k_e}$ then the set of transmission rates lower than d_i is universally-max-dominated by d_i (because of the way that capacity is shared by the routers). We can therefore remove all these actions. Observe, that the resulting subgame is equivalent to a congestion-control game in which i is removed and the capacity of all edges on r_i is decreased by d_i (for which we apply the arguments in this proof).

Otherwise, consider the set of transmission-rates of i lower than $\frac{c(e)}{k_e}$. Observe, that this set is universally-max-dominated by all actions greater or equal to $\frac{c(e)}{k_e}$ (again, because of the way that capacity is shared by the routers). This is true for all flows whose routes go through e. We can therefore remove all actions of these flows lower than $\frac{c(e)}{k_e}$.

Let e be defined as before. Observe, that in the subgame in which every node i whose route goes through e cannot choose an action lower than $\frac{c(e)}{k_e}$, all actions of such an i that are greater than $\frac{c(e)}{k_e}$ are universally-max-dominated by the action $\frac{c(e)}{k_e}$. Hence, we can iteratively remove universally-max-dominated action-sets until each flow whose route goes through e is only left with the action $\frac{c(e)}{k_e}$.

Observe, that in the resulting subgame every flow i whose route goes through e has to transmit at a rate of $\frac{c(e)}{k_e}$. This subgame is equivalent to the congestion-control game in which e is removed, all flows that go through e are removed, and for each flow whose route goes through e, we decrease the capacity of each edge on its route by $\frac{c(e)}{k_e}$. We apply the arguments in this proof to this new game.

It is easy to verify the the tie-breaking rules implied by the prescribed best-reply strategies are consistent with the described elimination sequence.

In order to see why the convergence rate is polynomial let us revisit the elimination sequence. Observe, that in every $2n$ rounds (going over all flows twice) we reach a subgame equivalent to a congestion-game in which either a flow or an edge is removed. It is easy to show, via similar techniques, that the resulting action-profile (that survives the elimination sequence) is optimal with respect to max-min fairness. □

C.1.2 Routing Games

The Game. A *routing game* [28], is played over a graph G with vertex set V, and an edge set E. The players are the vertices, and each vertex has a private full order of *strict* preferences over all the simple paths in the graph between itself and a common fixed target node d. Each vertex must choose an outgoing edge representing a choice of neighbouring node to forward traffic to. Vertices are selfish and wish to be assigned as highly ranked routes as possible (given their preferences).

We consider the well-known *"No Dispute Wheel"* [21] condition, which is known to generalize the Gao-Rexford setting [18] (see [28] and references therein).

The Mechanism. We go over the vertices in round-robin order. In each step, a single vertex chooses an outgoing edge. This continues until all vertices do not switch their outgoing edges (i.e., repeat their previous edge-choices). Then, the outputted route assignment consists of the final outgoing edges. If after $|V|^3 - 2|V|^2 + |V|$ rounds this does not occur, then no vertex is assigned an outgoing edge (no traffic is sent at all).

The Prescribed Best-Reply Strategies: In each step, a vertex a should choose a single outgoing edge such that the route from a to d induced by this choice (given current choices of other vertices) is most preferred by a (if no such route is induced then a does not choose an outgoing edge at all).

Theorem C.2 *If No Dispute Wheel holds, then the prescribed best-reply strategies are incentive-compatible and converge to a unique pure Nash equilibrium. This holds even in asynchronous settings.*

This sheds light on the recent results presented in [28]. Theorem C.2 follows from the following lemma:

Lemma 3 *Routing games are universally-max-solvable.*

Proof: A routing game is a game with private information in which the actions of the players (vertices) are their choices of outgoing edges. A player's utility function u_i assigns a value to every action profile in a way that is consistent with the preferences. Observe, that not assigning outgoing edges to the vertices is a penalty. By saying that the No Dispute Wheel condition holds for a routing game we mean that it holds for any full-information game induced by the actual utilities of the players. (For a definition of Dispute Wheels see [21].) We consider such an underlying game and show that it is universally-max-solvable. We do so by exhibiting an elimination sequence of universally-max-dominated action sets. We shall show how the first universally-max-dominated action-set T in the elimination sequence can be found. Finding the following universally-max-dominated action-sets is done via the recursive use of the same arguments.

Fix a vertex a_0 with at least 2 actions. Let R_0 be a_0's most preferred existing route to d. Let a_1 be the vertex closest to d on R_0 such that u_1 prefers some other route R_1 to the suffix of R_0 that leads from a_1 to d.

Let us first handle the case that there is no such vertex a_1. This means that (u, d) is the most preferred route of the vertex u closest to d on R_0. Assume that u has at least two actions (otherwise, u always chooses (u, d) and hence we can apply the same arguments to the vertex that comes before u on R_0, and so on). This implies that (u, d) universally-max-dominates all other actions of u, which in turn implies that we can choose all other actions of u to be T.

If such an a_1 exists, then we choose a_2 to be the vertex closest to d on R_1 such that a_2's most preferred route R_2 is preferred over the suffix of R_1 that leads from a_2 to d. Once again if there is no such a_2 we are done (for the same reasons as before). Similarly, we choose $a_3, a_4, \ldots$. Since there is a finite number of vertices, at some point some vertex will appear twice in this sequence $(a_0, a_1, \ldots)$. This would result in the formation of a Dispute Wheel (in which the a_is are the pivot nodes and the R_is are the routes) – a contradiction. $\square$

C.2 Cost-Sharing Games

The Game. A cost-sharing problem is specified by a cost function C defined over a universe of users U. For every subset of the users S, $C(S)$ represents the total cost that players in S are charged if serviced. We assume that C is nonnegative ($C \geq 0$) and nondecreasing ($\forall S \subseteq T \; C(S) \leq C(T)$. There is a *cost-sharing method* ρ which specifies, for every subset of the users S, how the cost $C(S)$ will be shared between the users in S (each user $i \in S$ is assigned a nonnegative cost-share $\rho(i, S)$ such that $\Sigma_{i \in S} \rho(i, S) = C(S)$. Every user $i \in U$ has a private value v_i for being serviced, and submits a single bid b_i from a pre-defined discrete set of bids, say, the integers $\{0, 1, \ldots, K\}$.

For every combination of private values, the outcome of the game is defined in the following way: If for all $i \in U$ $b_i \geq \rho(i, U)$ then all users are serviced and each user i is charged his bid b_i. If not, then the first user j (given some order $1, ..., n$) for whom $b_i < \rho(i, U)$ is removed and the same procedure is applied to the smaller universe $U \setminus \{j\}$. This recursion continues until a set of users whose bids exceed their cost-shares is reached, or no users are left (in which case no one is serviced).

The Mechanism: Users repeatedly, one by one and in round-robin order, submit bids. This goes on until all users pass (i.e., repeat their previous bids), at which point the final bids are used. If after nK rounds this does not occur then no user is serviced.

The Prescribed Best-Reply Strategies: In each step, user i bids the lowest bid that is a best-reply to the current bids of the other users.

Definition 11 *[33, 34] A cost-sharing method ρ is said to be cross-monotonic if $\forall i \in S \subseteq T$ $\rho(i, S) \geq \rho(i, T)$.*

Theorem C.3 *For every cross-monotonic cost-sharing method, the prescribed best-reply strategies are incentive compatible.*

Our result shows that the famous Moulin mechanisms [33, 34] can be implemented in an incentive-compatible distributed fashion via best-reply dynamics. These results can be extended to the more general class of acyclic mechanisms [29]. Theorem C.3 follows from the following lemma:

Lemma 4 *Cost-sharing games are universally-max-solvable.*

Proof: We demonstrate an elimination sequence in the game that leaves every player with only a single action profile. The first steps in our elimination sequence are to eliminate for each player the bids that are above his valuation v_i. These are clearly universally-max-dominated actions, as bidding above one's valuation can yield only one of two outcomes: Either the player is awarded the service and is then requested to pay the amount he bid (in which case he has a negative utility), or he is not awarded the service and does not pay (a utility of 0). For any other (lower) bid, his utility can be 0 or higher, and so the set of all other bids universally-max-dominates the set of bids above the valuation.

We now proceed to iteratively remove the bids of players. Let us assume that we are in step t of the elimination sequence. In any step, our current game G_t (after some actions have been eliminated) contains for each player a possible set of bids $B_{i,t}$ (initially this is the set $\{0, \ldots, v_i\}$ since we've already eliminated higher bids). Let us define the set of players S_t as the set of players that have some chance of being awarded the service. I.e., let S_t be the set of players that have some remaining bid $b_i \in B_{i,t}$ that allows them to win in some eventuality (e.g., when any other player j bids the maximum: $\max(B_{j,t})$).

Throughout our elimination sequence the set S_t will be monotonically non-increasing. I.e., if at some point $i \notin S_t$ then $\forall t' > t \quad i \notin S_{t'}$. This property is maintained due to cross–monotonicity and the fact that we are only removing actions from the game. If actions are removed from some player's available set, he can bid no higher and is therefore not going to improve his ability to acquire the service. A player will also not gain from the removal of other players' strategies, since if they are suddenly not able to acquire the resource, the cost to the player itself may only increase or remain the same (due to the cross-monotonicity property). In any step t in our elimination one of the following cases must hold:

- **There exists a player with more than one action, and that player is not in S_t.** This player is never going to be awarded the service in G_t, and so will never pay his bid. We can therefore eliminate all bids except the highest one for such a player (since all actions yield a utility of 0, they are now universally-max-dominated in the weak sense by the highest bid that yields the exact same payoff). From now on, this player is effectively out of the game in one elimination step.

- **All players with more than one action are in S_t, but there exists a player in S_t for which $\min(B_{i,t}) < \rho(i, S_t)$.** Note that since players that are not in S_t have no chance of ever getting the service, this lowers the price that player $i \in S_t$ can ever hope to pay and be awarded the service to $\rho(i, S_t)$. Because of cross-monotonicity, a smaller group than S_t will increase the cost for player i, and no larger group is possible. Therefore, we can eliminate all bids below $\rho(i, S_t)$ for any such player. They never give a payoff of more than 0 in the game G_t and are therefore universally-max-dominated by higher bids that give a payoff of at least 0.

- **All players with more than one action are in S_t, and have $\min(B_{i,t}) \geq \rho(i, S_t)$.** In this case we can determine that all possible outcomes of the game G_t end with the set of player S_t being awarded the resource. This is because each of these players pays at least $\rho(i, S_t)$ – no other bids are available to them. Since this outcome is assured, each player is always better off paying as little as possible, and so we can deduce that all the higher bids of any players are universally-max-dominated by his lowest remaining bid. These eliminations bring us to a game in which all players have only a single strategy remaining and we are done.

Since we have shown that there are bids to be eliminated in any case (unless every player has only one strategy remaining, we have shown how to construct the full sequence of eliminations. □

C.3 Stable-Roommates Games

The Game. [17] There are n students $1, \ldots, n$ who must be paired to dorm rooms. Each student has a private full order of *strict* preferences over the other students. Each student must choose another student to be his roommate. Students are selfish and wish to be paired with students that are ranked as high as possible (given their preferences).

A stable matching is a matching of students such that no pair of students that are not matched to each other prefer each other to their matched roommates.

The Mechanism. We go over the students in round-robin order. At each step a single student announces another student. We call this an offer made by the student to the other student. This goes on until all students do not change their offers (i.e., repeat their previous roommate choices), at which point every two students that made an offer to each other (in their final announcements) are matched, and all other students remain unmatched. If after $n^3 - 2n^2 + n$ rounds this does not occur then all students remain unmatched.

The Prescribed Best-Reply Strategies: If student i announces a student j, and j prefers i over all other offers made to it (in the most recent student announcements) we say that i makes j a better offer. In each step, a student must choose the student he prefers most out of all students to whom he can make a better offer.

We show the the prescribed best-reply strategies are incentive-compatible a converge to a unique stable matching in the following natural special cases:

C.3.1 Intern-Assignment Games

The n students consist of interns and hospitals. Each intern has a private full order of *strict* preferences over the hospitals. All hospitals have identical preferences over interns (not necessarily known to the interns).

Theorem C.4 *For intern-assignment games the prescribed best-reply strategies are incentive compatible and converge to a unique stable matching.*

This follows from the following lemma:

Lemma 5 *Intern-assignment games are universally-max-solvable.*

Proof: We consider a simple condition on the preferences of the students in stable-roommates games, which we call "No Preference Cycle".

Definition 12 *A Preference Cycle in the stable-roommates game, is a cyclic order defined over $k > 2$ students $p_1, \ldots, p_k$, such that each player p_i prefers p_{i+1} over p_{i-1} (All indices are considered modulo $k+1$).*

No Preference Cycle holds if no Preference Cycles can be induced by the stable-roommates game. It is easy to show that in intern-assignment games there can be no Preference Cycles.

Claim C.5 *No Preference-Cycle holds for any intern-assignment game.*

Proof: By contradiction. Assume that a Preference Cycle $p_1, ..., p_k$ exists. Consider a hospital p_i in the Preference Cycle (observe that a Preference Cycle always has the form ...-intern-hospital-intern-hospital-...). p_i must prefer p_{i+1} over p_{i-1}. Now, consider the hospital p_{i+2}. It prefers p_{i+3} over p_{i+1}. Hence, p_i also prefers p_{i+3} over p_{i+1}, which implies that it prefers p_{i+3} over p_{i-1}. If we continue this line of argument we can show that p_i's desire for an intern only increases as we move forward along the Preference Cycle. However, because of the cyclical structure of the Preference Cycle this would lead to a contradiction (because the intern that comes before p_{i-1} would be both preferred over p_{i-1} and less preferred than p_{i-1}. □

We shall prove the theorem for the more general case in which No Preference Cycle holds. A stable roommates game is a game with private information in which the actions of the players (students) are their choices of roommates. We define the utility utility u_i of player i as follows: Let v_i be a valuation function of i that assigns a value to every other student in a way that is consistent with i's preferences. For every set of actions (roommate-choices) $a = (a_1, ..., a_n)$, u_i has a value of $v_i(a_i)$ if i is preferred by a_i (given v_{a_i}) over all other students j for which $a_j = a_i$). Otherwise,

$u_i(a) = 0$. That is, i get his value for the student he chose if he is making that student a better offer. Observe, that leaving all players unmatched is a penalty. By saying that the No Preference Cycle condition holds for a stable-roommates game we mean that it holds for any full-information game induced by the actual utilities of the players. We consider such an induced game and show that it is universally-max-solvable. We do so by exhibiting an elimination sequence of universally-max-dominated action sets. We shall show how the first universally-max-dominated action-set T in the elimination sequence can be found. Finding the following universally-max-dominated action-sets is done via the recursive use of the same arguments.

Pick a student i in the game and look at his most preferred roommate j. Then, look at j's most preferred roommate, and so on. Eventually one of the players must appear twice, thus forming a cycle. The cycle must be of length 2 because otherwise we get a Preference-Cycle. This implies two students who prefer each other over any other roommate. For each of these two students, the set of actions of selecting anyone other than the other student is universally-max-dominated by the action of selecting this ideal partner. We therefore eliminate all these actions (in two elimination steps), and remain with a subgame that is equivalent to a stable roommates game without these two students. Using the same line of reasoning we can find another pair of students that can be ideally matched in this subgame, and repeat the same procedure. □

C.3.2 Correlated Two-Sided Markets

The students consist of sellers and buyers. The game is represented by a weighted bipartite graph (for simplicity, assume that no two edges have the same weight) in which vertices on one side represent buyers, and vertices on the other side represent sellers (markets). Each buyer has a preference order over sellers, such that a seller A is ranked higher than another seller B by that buyer if the edge connecting that buyer to A has a bigger weight than the edge connecting the buyer to B. Similarly, sellers prefer buyers that are connected to them by heavier edges.

Theorem C.6 *For correlated two-sided markets the prescribed best-reply strategies are incentive compatible and converge to a unique stable matching.*

This follows from the following lemma:

Lemma 6 *Correlated two-sided markets are universally-max-solvable.*

Proof: We shall demonstrate an elimination sequence for the Correlated Two-Sided Market game. Let e be the edge in the graph that has the biggest weight. Both the buyer and the seller at each end of this edge prefer each other over all other players in the game. They are therefore able to make each other a better offer than any other player, and thus an action of choosing each other is guaranteed to give them their maximal payoff. It therefore universally-max-dominates all other actions they may chose (that yield a lower profit in any eventuality). So in two eliminations, the two players on e are left with only a single action and are effectively out of the game. The elimination sequence now proceeds as it would with the matching problem on the reduced graph (where the two players and all edges connected to them have been removed). □

C.4 Iterative 1st-Price Auctions

In Section 4 we handle the case of auctions with unit-demand bidders. While the general case is not universally-max-solvable we show that the toy special case in which only one item is sold is.

The Game: A single item is sold in a first price auction. There are n bidders (players) and each bidder i has a private value v_i for the item and submits a single bid b_i from a pre-defined discrete set of bids (actions), say, the integers $\{0, 1, \ldots, K\}$. The winner is the bidder with the highest bid b_i (with ties broken, say, lexicoraphically) who pays his bid b_i for the item (and gets utility $v_i - b_i$.)

The Mechanism: Bidders repeatedly announce, one by one and in round-robin order, bids from the set of bids until all bidders "pass" (i.e., repeat their previous bid), at which point the final bids are used. If after $n^2K + n$ rounds this does not occur then the item is left unsold.

The Prescribed Best-Reply Strategies: In each step, bidder i bids the lowest bid that beats all other current bids, provided that this bid is feasible for him, i.e., less than v_i. If no feasible bid can beat the other current bids, he bids the largest feasible bid (i.e., at most v_i).

Theorem C.7 *The prescribed best-reply strategies are incentive compatible and converge to the pure Nash equilibrium in which the highest bidder wins the item and is charged the (discretized) second highest bid.*

When the initial bids of the bidders are all 0, we get the familiar English auction behavior in which players repeatedly increase their bids by the minimal increment, and drop from the auction (i.e., pass) once their values are exceeded. However, note that our mechanism does not force ascending behavior, does not forbid jumps or stalling (passing and then increasing the bid), does not require 0 as the initial bid, and yet is guaranteed to end up with the same equilibrium. So, we show that in this fundamental setting best-reply dynamics (with given tie-breaking rules) mimic, in certain cases, the behavior of the well-known ascending English auction. Notably, in our framework the ascending behavior is completely endogenous and not dictated in any way by the auction rules.

We prove the following lemma (that implies the theorem):

Lemma 7 *1st-price auction games are universally-max-solvable.*

Proof: The 1st price auction game is a game with private information in which the actions of the players are their bids, and their utilities are a function of their private values (if a player wins $u_i = v_i - b_i$ otherwise $u_i = 0$). Observe, that leaving the item unsold is indeed a penalty. We prove that any full-information game induced by the actual utilities is universally-max-solvable by exhibiting an elimination sequence of universally-max-dominated action sets. We shall show how the first universally-max-dominated action-set T in the elimination sequence can be found. Finding the following universally-max-dominated action-sets is done via the recursive use of the same arguments.

Consider the case that there is a bidder i that has at least 2 actions, and the highest possible bid of this player is at least v_i. Then, we can choose T to be that player's highest bid (i's utility from bidding a bid higher or equal to v_i is at most 0). By repeating this argument, we are left with the case that all the possible actions of bidders (with more than one possible bid) are strictly less than their values for the item.

Let b_i be the lowest possible bid of player i. Let j be a bidder with at least 2 actions for which b_j is minimized (If more than one such bidder choose one that would lose if all of them bid their lowest bid. If there is no such bidder then there is a single action-profile and we are done). There are now two possibilities: Either there exists another bidder with at least 2 actions, or no such bidder exists. If there is such a bidder, then j never wins when bidding b_j, and therefore b_j is universally-max-dominated. Otherwise, each other bidder has only one possible bid. If j beats these bids when bidding b_j then b_j universally-max-dominates all other actions of j (and we can choose T to be the other actions of j). If this is not the case, then b_j is universally-max-dominated by j's other actions, and we choose T to be b_j.

It is easy to see that the tie-breaking rules implied by the the prescribed best-reply strategies are consistent with the elimination sequence described above. □

D Beyond Universally-Max-Solvable Games

D.1 Auctions with Unit-Demand Bidders

Theorem D.1 *The prescribed best-reply strategies are incentive compatible and converge to the pure Nash equilibrium in which items are allocated as in the maximal-weighted-matching and prices are the corresponding VCG prices.*

Proof:

Let us recall some classic results in auction theory regarding competitive equilibria:

Definition 13 *Let $S = (j_1, ..., j_n)$ be an allocation of the items to the n bidders (every bidder is either assigned a unique item or the empty set). Let p be a price vector. We say that (S, p) is a* competitive equilibrium *if for every $i \in [n]$ $j_i \in D_i(p)$ (every bidder gets an item in his most demanded set), and for every $i \in [n]$ such that $j_i = \emptyset$ it holds that bidder i pays 0.*

Theorem D.2 *[23] If (S, p) is a competitive equilibrium then it must hold that S is the social-welfare-maximizing item-assignment.*

Theorem D.3 *[40, 27] There is a unique minimum competitive equilibrium price vector, and it corresponds to the VCG prices.*

Informally speaking, we prove that the prescribed best-reply strategies converge to the VCG allocation and prices by interpreting the prescribed best-reply strategies as a form of Kuhn's Hungarian Method for computing optimal weighted matchings in a bipartite graph [25]. The following claim guarantees that incentive compatibility follows once we show that best-reply dynamics converge to the VCG outcome.

Claim D.4 *If best-reply strategies converge to the VCG allocation and prices, then they are incentive compatible.*

Proof: Assume for the point of contradiction that there is some bidder i that reaches an allocation A and prices p_A that are superior for him than the VCG outcome by deviating from the VCG outcome. Consider the case in which we alter bidder i's valuation function to be such that his value for every item j is exactly the price paid for j by the highest bidder in p_A. Observe that because the valuations of the other bidders remain unaltered, and they all "passed" (otherwise the mechanism would not have allocated any items at all) this means that the items they receive must be in their most demanded sets. Also observe, that given bidder i's new valuation function the item he gets must also be in his most demanded set. So, A and p_A are a competitive equilibrium for this new valuation function (and the previous valuations for the other bidders). Since the VCG prices are always the minimal competitive equilibrium this means that i could reach this outcome, or a better one (for him), simply by pretending to have this new valuation function. However, this contradicts the truthfulness of VCG prices. □

The main part of the proof is proving convergence of best-reply dynamics. The following definition will be helpful:

Definition 14 *An n-tuple of vectors of bids and its associated allocation of items S is a* maximal competitive allocation *if no bidder can* increase *any of his bids so as to induce an allocation in which his utility is improved.*

Thus, at a maximal competitive allocation, it must be the case that given the vector of maximum bids $b^* = (b_1^*, \ldots, b_m^*)$ as prices, each bidder receiving an item in the allocation is getting (one of) his most-demanded item(s) or one that is at most ϵ worse than his most-demanded item; also, for each bidder i not receiving an item in S it is true that $b^*j \geq v_{i,j} - \epsilon$.

We show that once best-reply dynamics reach a maximal competitive allocation, the bids will monotonically fall until the VCG prices and allocation are reached, at which point the best-reply dynamics will have converged. This is true because of our careful choice of tie-breaking rules (between different possible best-replies).

Lemma 8 *Starting from any maximal competitive allocation, and associated bids, the prescribed best-reply dynamics monotonically decrease the prices and converge to the VCG prices and allocation.*

Proof: First note that as long as we remain at a maximal competitive allocation, no bidder will ever raise his bid for any item. Furthermore, after a single bidder's turn, the new set of bids and induced allocation must still be a maximal competitive allocation. Thus, it suffices to show that in every complete round of bidding, at least one bidder decreases his price vector, unless the VCG allocation and payments have been reached.

W.l.o.g, assume that k items are allocated to bidders $1, \ldots, k$ in the current maximal competitive allocation. We begin by arguing that if $k < n$, then there is some at least one bidder who will decrease some of his bids. Consider a bidder i who has the highest bid for an item j that is currently unallocated. Since the mechanism is not allocating item j, it must be the case that bidder i is receiving at least as much utility from the item j' that he is allocated than if he were to receive item j. In the case that item j was his target item, then he is receiving ϵ more utility from the current allocation and bids than if he were to receive j. Thus, best-reply dynamics (with our tie-breaking rules) prescribe that either he maintains his current bid on item j' while decreasing his bid on item j by 2ϵ and all other bids by ϵ, or he has a best-reply move that increases his utility, in which case his bids will also decrease. In the case that item j was not his target, by making item j his target in his next round of bidding, he would decrease all bids by at least ϵ.

Hence, bids will decrease until all items are being allocated. Given that all items are being allocated, to see that prices will decrease until VCG payments are reached, assume for the point of contradiction that all bidders have passed in a round of bidding, but we are not at VCG prices. Since each bidder has passed, his current bid for the item he receives must be identical to the second highest bid for that item. Since we are not at the VCG prices, there exists a set of $k \leq n$ items $j_1, \ldots, j_k$ such that the prices may be decreased by some amounts inducing a smaller competitive equilibrium (namely the VCG allocation and prices). Assume (w.l.o.g.) that bidders $1, \ldots, k$ receive these items in the original allocation. It must be the case that each of these bidders also receives items from the set $\{j_1, \ldots, j_k\}$ in the VCG allocation, otherwise they would have less utility than in the previous allocation, and thus would want to switch items. Therefore, we conclude that in the original allocation, the second-highest bids for each item are held exclusively by bidders in the set $\{1, \ldots, k\}$.

Consider creating the directed graph, with the k bidders as nodes, and an edge from i to i' if bidder i has the second-highest bid for the item i' receives. Every node must have an in-degree of at least 1, and thus there exists at least one directed cycle. Observe that augmenting the allocation according to the cycle will increase the utility of each bidder in the cycle by exactly ϵ, contradicting the efficiency of the equilibrium. We conclude that prices fall until the VCG prices are reached, implying the VCG allocation. $\square$

To complete our proof of Theorem 4.1, the following lemma considers the prescribed best-reply dynamics given a configuration of the game that is not a competitive efficient equilibrium.

Lemma 9 *From any initial configuration, the prescribed best-reply dynamics will result in a maximal competitive allocation.*

Proof: Given any configuration of bids that has been reached after at least one full round of bidding, assume that the allocation allocates k items. First observe that at any subsequent stage in the prescribed best-reply dynamics the allocation will always allocate at least k items. This can easily be seen by noting that in a single best-reply, if a bidder lowers his bid, he must still be allocated an item, and lowering his bids can only increase the size of the allocation. If a bidder increases his bid, he must not have been assigned an item in the previous allocation, and he will raise his bid in such a way that he will have the unique highest bid for at most one item. Therefore, his bidding will preserve the size of the allocation. We now show that the size of the allocation will increase until we have reached a maximal competitive equilibrium.

Given some configuration of bids that has been reached after at least one full round of bidding, let $X \subset \{1, \ldots, m\}$ be the union of all minimal overdemanded sets of items at the current set of bids b^*, and let Y be the associated set of bidders,

$$Y \subset \{1, \ldots, n\} := \{i : argmax_j(v_{i,j} - b^*_j) \subset X\}.$$

Clearly $|X| < |Y|$, and every bidder not in X either receives some item, or has sufficiently low valuations that he desires no item at the current bids. If the set X has remained the union of all minimally overdemanded sets throughout a complete round of bidding, and the corresponding set Y has also remained unchanged, it must be the case that the bids for each item in X has increased by at least ϵ in the previous round of bidding, since by definition each bidder will bid so as to receive an item in the allocation, and thus must ϵ-jump the previous bid for a most-desired item. There are two possible ways that the set X might be augmented: the prices might get sufficiently high that a bidder drops out and prefers no items, or the prices get high enough so that some bidder in Y prefers some new item that is not in X. In the first case the number of bidders that receive an item in the allocation or do not desire any items increases. In the second case, either the size of the matching increases by 1, or $|X|$ increases by at least 1.

To conclude, note that $b^*_j \leq \max_i v_{i,j}$ and thus the prices cannot increase indefinitely; additionally, $|X| \leq m$, and thus it cannot increase indefinitely, so after every $\frac{m \cdot \max_{i,j} v_{i,j}}{\epsilon}$ rounds of bidding, the size of the set of bidders either receiving an item or not wanting any item must increase by at least 1. Hence, after at most

$$\frac{nm \max_{i,j} v_{i,j}}{\epsilon}$$

rounds of bidding, all bidders that have a most-desired item are receiving an item (which must then be a most-desired item), and thus we must be in a maximal competitive equilibrium. $\square$

$\square$

D.2 Adword Auctions

The Game. There are k slots $1, ..., k$. Each slot j has a clique-through-rate (CTR) α_j such that $\alpha_1 > \alpha_2 > \ldots > \alpha_k$ (we shall call the low-index slots "higher slots", as higher slots are preferred to lower ones). There are n bidders. Each bidder i has a private value v_i per click. If bidder i is assigned slot j his value for that slot is defined to be $\alpha_j \cdot v_i$. Each bidder submits a single bid b_i from a pre-defined discrete set of bids (multiples of some $\epsilon > 0$). The allocation rule assigns the k slots to the k highest bidders (the highest bidder gets slot 1, etc.) and charges each winning bidder a cost per click that equals the bid of the bidder that was assigned the slot below his. We denote by $\pi(j)$ the player that was awarded the j'th slot. Formally, the utility of a player that is assigned to slot j is

$$u_{\pi(j)} = \alpha_j \cdot v_{\pi(j)} - \alpha_j \cdot b_{\pi(j+1)}$$

where $b_{\pi(j+1)}$ is the bid submitted by the bidder who is assigned the $(j+1)$'th slot.

The (Randomized) Mechanism. In each step choose one bidder, uniformly at random, and allow that bidder to change his bid. This goes on until all bidders do not change their bids (i.e., repeat their previous bids), at which point the final bids are used to allocate the slots. If after some large number of steps T (to be defined later) this does not occur then no slot is allocated.

The Prescribed Best-Reply Strategies. In each step, the bidder i chosen to play in that step chooses the bid b_i that maximizes his utility (given the bids of the other bidders and the allocation rule). Ties between different such bids will be broken as follows: A bidder i that bids so as to get slot j, should bid a value b_i such that

$$\alpha_j \cdot (v_i - b_{\pi(j+1)}) = \alpha_{j-1} \cdot (v_i - b_i)$$

That is, i chooses a bid b_i such that i is indifferent between getting the j'th slot and paying the next highest bid, or getting the $(j-1)$'th slot and paying b_j. If any best-reply bid of i leads to i not getting a slot at all, then i will bid his highest feasible bid, i.e, the maximal bid b_i that is smaller or equal to v_i.

Theorem D.5 *The prescribed best-reply strategies are incentive-compatible and converge to a pure Nash equilibrium in which every bidder pays his VCG price.*

[8] have shown that (randomized) best-reply dynamics indeed converge to this pure Nash equilibrium. We strengthen this result by showing that this is achieved in an incentive-compatible manner.[3]

Proof: Let us denote by $\pi(j)$ the player that ends up in the j'th slot if the Nash equilibrium is achieved (i.e., if all players follow the best-reply strategy and converge to a stable solution), and by b_i the bid submitted by player i in this scenario.

We shall assume, for the point of contradiction, the existence of a bad state s in which all players but one are stable, and such that this player prefers s to the unique Nahs equilibrium. Let $\sigma(j)$ be the player that is allocated the j'th slot, and let b'_i be the bid of player i in this scenario.

If the unstable player does not win any slot, then he has a utility of 0, and is certainly not better off than in the Nash equilibrium outcome. We therefore assume that the unstable player occupies some slot in the state s. We begin by proving the following claim on the bids submitted by players:

[3]We note that as the adwords mechanism is randomized, best-reply dynamics are guaranteed to converge to a specific action profile with high probability (arbitrarily close to 1). Therefore, as long as the adwords mechanism consists of sufficiently many steps, the possibility that the prescribed best-reply strategies do not converge has but a inconsequential affect on players' (expected) utilities. For ease of exposition we shall therefore ignore the effect of this event on players' (expected) utilities, rather than state the result in terms of ϵ-truthfulness.

Claim D.6 *Let j be the slot that is occupied by the single non-stable player in the state s.*

$$\forall j' > j \quad b_{\pi(j')} \leq b_{\sigma(j')} \quad \textit{and} \quad v_{\pi(j')} \leq v_{\sigma(j')}$$

if j is the lowest slot, the maximal bid of all losing players is also no lower than the maximal bid of losing players in the Nash equilibrium.

That is, in all lower slots, the valuation of the player that is awarded this slot is at least as high as the valuation of the player that got it in the Nash equilibrium. In addition, players in those slots bid at least as high as players in those slots do in the Nash equilibrium.

Proof: Notice that in the Nash equilibrium outcome, the losing players are the $n - k$ players with the lowest valuation, and that they all bid their valuation according to the best-reply strategy. In the state s, there are also $n - k$ players that are not awarded any slot, and they too bid their valuation (They must all be stable and thus follow the best-reply strategy). The highest bid they produce can therefore be no lower than the one given in the Nash equilibrium.

Induction base. We now proceed to prove the claim by induction on the slot j'. If j is the lowest slot, then we are done. Otherwise, let $j' = k$ be the lowest slot. The amount payed by the player in this slot is determined by the highest bid of the losing players. This amount, as we have shown can only be greater than that of the Nash equilibrium. The player position k has a higher valuation than the losing players. This is because all losing players are bidding their valuation, and the player at position k he must have bid higher than them. We know that this bid is lower than his valuation (otherwise he would get a negative payoff and would not be stable – bidding lower would be preferable). Since there are at least $n - k$ players with lower valuations than player $\sigma(k)$, we are assured that his valuation is at least as large as that of player $\pi(k)$ that had *exactly* $n - k$ other players with a lower valuation than his own.

Since the bid of the highest losing player is no lower, and $v_{\sigma(k)} \geq v_{\pi(k)}$ we can deduce that $b'_{\sigma(k)} \geq b_{\pi(k)}$. This is because the player at the k'th slot pays his price of indifference for the next slot which is defined as:

$$b_{\pi(k)} = \frac{\alpha_k}{\alpha_{k-1}} \cdot b_{\pi(k+1)} + \left(1 - \frac{\alpha_k}{\alpha_{k-1}}\right) \cdot v_{\pi(k)}$$

and we have:

$$b_{\pi(k)} = \frac{\alpha_k}{\alpha_{k-1}} \cdot b_{\pi(k+1)} + \left(1 - \frac{\alpha_k}{\alpha_{k-1}}\right) \cdot v_{\pi(k)} \leq \frac{\alpha_k}{\alpha_{k-1}} \cdot b'_{\sigma(k+1)} + \left(1 - \frac{\alpha_k}{\alpha_{k-1}}\right) \cdot v_{\sigma(k)} = b'_{\sigma(k)}$$

This concludes the proof for the base of the induction.

Induction step. We shall assume that the claim has been proven up to slot $j' + 1$ and will prove it for slot j'. The player in position j' must have a higher value than the player in position $j' + 1$. This is because the player in position $j' + 1$ is stable, and is bidding its price of indifference – the highest price it will be willing to pay to get slot j' rather than $j' + 1$. The player at position j' is indeed paying this price and is stable. That is, it is gaining more than it would gain for slot $j' + 1$. This implies that his price of indifference for slot $j' + 1$ would have been higher, and so it must be that $v_{\sigma(j')} > v_{\sigma(j'+1)}$. From the induction assumption we have that $v_{\sigma(j'+1)} \geq v_{\pi(j'+1)}$ this then implies that $v_{\sigma(j')} > v_{\pi(j'+1)}$, but $v_{\pi(j')}$ is the next valuation after $v_{\pi(j'+1)}$ and no player has valuation between them (we know this because in the Nash equilibrium outcome players are awarded slots in an order that is consistent with their valuations). Therefore, it must be that $v_{\sigma(j')} \geq v_{\pi(j')}$ as needed.

Now, since in both states the player in position j' is stable, it is bidding its price of indifference.

$$b_{\pi(j')} = \frac{\alpha_{j'}}{\alpha_{j'-1}} \cdot b_{\pi(j'+1)} + \left(1 - \frac{\alpha_{j'}}{\alpha_{j'-1}}\right) \cdot v_{\pi(j')} \leq \frac{\alpha_{j'}}{\alpha_{j'-1}} \cdot b'_{\sigma(j'+1)} + \left(1 - \frac{\alpha_{j'}}{\alpha_{j'-1}}\right) \cdot v_{\sigma(j')} = b'_{\sigma(j')}$$

□

Using our claim we can deduce that the unstable player $\sigma(j)$ is forced to pay for the j'th slot a price that is at least as high as the price it cost in the Nash equilibrium outcome (since it pays the bid of the player below it which could only increase). However, in the Nash outcome he could have bid for that position and gotten this payment but did not (Note that there is a case in which positions j and $j+1$ are awarded to two players with the same bid and the player $\sigma(j)$ cannot get position j because ties are not broken in the right way. In this case, if he goes for position $j-1$ he will pay the same amount and will only gain more clicks for the same price). This brings us to a contradiction. It cannot be that the unstable player is better off than in the Nash equilibrium. □