Hard-to-Manipulate Combinatorial Auctions

Hard-to-Manipulate VCG-Based Auctions

Saurabh Sanghvi * David C. Parkes †

Jan 21, 2004

Abstract

Mechanism design provides a framework to solve distributed optimization problems in systems of self-interested agents. The combinatorial auction is one such problem, in which there is a set of discrete items to allocate to agents. Unfortunately, recent results suggest that it is impossible to implement reasonable approximations without losing robustness to manipulation. Furthermore, the Vickrey-Clarke-Groves (VCG) mechanism is known to be vulnerable to manipulation when agents can bid under multiple false names. In this paper we relax incentive constraints and require only that useful manipulation be **NP**-hard. We prove that any tractable approximation algorithm can be made to produce a hard-to-manipulate (VCG-based) mechanism, providing a useful counterpoint to these negative results. We also show that false-name bid manipulation in the VCG is **NP**-hard.

*Harvard College, ssanghvi@post.harvard.edu

†Division of Engineering and Applied Sciences, Harvard University, parkes@eecs.harvard.edu

1 Introduction

The problem of mechanism design is that of designing an effective system-wide solution to a decentralized optimization problem where the inputs are the private information of self-interested agents. The goal is a mechanism that, through a particular means for choosing an outcome and payments, can induce agents to communicate their information truthfully and can implement a desirable social choice function.

A widely studied problem in mechanism design is the combinatorial auction problem. This is the problem of efficiently auctioning a basket of different goods to agents who may have nonlinear preferences over various subsets. One of the most attractive solutions to this problem is the Generalized Vickrey Auction (GVA), which is a special case of the general class of Vickrey-Clarke-Groves (or VCG) mechanisms. The GVA has the advantages of being strategyproof (that is, it is the dominant strategy for each player to report its true preferences) and allocatively efficient (it maximizes the value of the allocation over all agents) (Jackson 2000).

Several problems arise, however. First, since maximizing allocative efficiency is **NP**-hard in this setting, the GVA is necessarily intractable. Approximation algorithms must be used. Unfortunately, recent results show that introducing a tractable approximation algorithm into the GVA must either leave the mechanism manipulable or "unreasonable" in its approximation properties.[1] A second drawback is the susceptibility of the GVA to sophisticated types of manipulation, like that of false-name-bids, whereby agents bid under multiple identities. In fact, it is provably impossible to provide an auction that is both false-name-proof and allocatively efficient (Yokoo, Sakurai, and Matsubara 2002).

In this paper, we consider how these negative results change when the agents *themselves* are limited in their computational power. By relaxing strategyproofness to a type of "hardness-of-manipulation," one might hope to move the difficulty of the problem from the mechanism to the manipulators, thereby achieving reasonable and tractable mechanisms. We adopt a notion of hard-to-manipulate that states that the problem of computing a useful non-truthful bid given knowledge of the bids from other agents is **NP**-hard.

Specifically, this paper presents two new results. *First*, we provide sufficient conditions on the approximation algorithm of a VCG-based combinatorial auction that make manipulation **NP**-hard. The result holds even if all bidders are single-minded—that is, they only value a single bundle of goods at a certain value. We demonstrate that any approximation algorithm can be modified to satisfy these properties with no loss in efficiency and only a linear-time slowdown. *Second*, we prove that false-name bid manipulation in the GVA is **NP**-hard. This implies that we can sidestep the negative result ruling out efficiency and false-name-proofness, within our model of hard-to-manipulate mechanisms.

The guarantee that our notion of hard-to-manipulate provides to a mechanism designer

[1] Recent results due to Lavi et al. (Lavi, Mu'alem, and Nisan 2003) suggest a general negative result, even looking beyond the VCG at alternative payment schemes.

is substantially weaker than that of strategyproofness, given its worst-case nature. However, the demands of practicality, in light of substantial negative results make it at least a necessary and somewhat comforting recourse. We postpone further discussion of our results to the end of the paper.

2 Related Work

The literature of computational mechanism design generally focuses on how bounded-rationality is a problem: for agents, it leads to a push for mechanisms with simple-to-compute dominant strategy equilibria (Varian 1995) and a consideration of steps to mitigate unnecessary effort in valuation and bidding (Parkes 2003); for the mechanism infrastructure, it leads to a push for mechanisms that are tractable to implement (Nisan and Ronen 2001).

Our approach is to view bounded-rationality as a potential *boon* to the mechanism. The literature concerned with using bounded rationality in mechanism design in this way is fairly limited. The most directly similar approach is taken by Conitzer and Sandholm (2002, 2003), who prove hardness of manipulation results for common voting protocols, building on earlier work (Bartholdi 1989). There are also characterization results in the literature; for instance Archer et. al. (2003) give a precise characterization of the group manipulation opportunities available when using the marginal cost mechanism for multicast cost sharing. It is a straightforward corollary of their characterization that such group manipulation is *easy* under the model that we have chosen.

Other authors have adopted satisficing notions, such as those of ϵ-strategyproofness (Kothari, Parkes, and Suri 2004) and truth-revelation as a best-response with high probability whatever the bids of other agents (Archer, Papadimitriou, Talwar, and Tardos 2003). Indeed, there is a tradition of ϵ-equilibrium models in the economic literature (Schummer 2002). Another approach is to provide an explicit model of the bounded-knowledge or bounded-reasoning process available to an agent and design specific mechanisms to try to mitigate this constraints and still achieve useful outcomes (Nisan and Ronen 2000, Parkes 2003).

Finally, it is worth mentioning that our approach closely parallels that of another in computer science, cryptography. Whereas cryptography allows an agent to hide information by making it too hard for other agents to force a revelation, we hope to force an agent to reveal information by making it too hard for that same agent to *avoid* revelation. Just as in cryptography, we desire a certain "one-way" function wherein the mechanism's computation is easy but any attempt to backsolve the computation (to manipulate the result) will meet with difficulty.

3 Preliminaries

Our approach for mechanism design with bounded-rational agents is to exhibit a mechanism for which the agent manipulation problem is an **NP**-hard problem. We will consider the agent manipulation problem to be that of, given an agent i and true type vector θ for all agents, determining a new strategy θ_i' such that agent i's overall utility from the allocation and transfers improves by ϵ, for some fixed, small ϵ.

This definition is in a few ways stronger than it could be. By requiring the agents to manipulate only to improve their utility by ϵ, we are showing the stronger result that *any* manipulation is hard, not, for instance, that manipulating *optimally* is hard. Moreover, we give the agent access to the true type declarations of the other agents. Conceivably one could prove hardness of manipulation results by depending on the uncertainty an agent faces in the potential types of the other agents (Conitzer and Sandholm 2002). This will lead to the problem, however, of modeling the partial knowledge that an agent might have (since assuming the agent knows nothing will yield a very tenuous guarantee). We sidestep this problem by assuming the agent knows exactly what the other agents will report—showing that even then the manipulation is hard gives an only stronger result.

3.1 VCG-based mechanisms

Assume we have a set of bidders, $N = \{1, \ldots, n\}$, and a universe of goods $G = \{1, \ldots, m\}$. Each agent i privately observes a type θ_i, specifying the bidder's valuation for different subsets of the universe. Specifically, given a type θ_i and some $S \subseteq G$, we can define a valuation function $v(S, \theta_i) \to \mathbb{R}$. We make two assumptions about the bidders' valuation functions:

- Free Disposal: If $S \subseteq T$, then $v(S, \theta_i) \leq v(T, \theta_i)$
- Normalization: $v(\emptyset, \theta_i) = 0$

We let θ denote the type vector $(\theta_1, \ldots, \theta_n)$ of all agents.

Definition 1 *We define a combinatorial auction as a pair of functions $(k(\cdot), t(\cdot))$, known as the allocation and transfer functions, respectively, with the following constraints:*

- $k(\theta) = (k_1(\theta), \ldots, k_n(\theta))$*, where each* $k_i(\theta) \subseteq G$*, and for any* i, j*,* $i \neq j$*,* $k_i(\theta) \cap k_j(\theta) = \emptyset$*.*
- $t(\theta) = (t_1(\theta), \ldots, t_n(\theta))$ *where each* $t_i(\theta) \in \mathbb{R}$

The agents declare types $\hat{\theta}_i$ (not necessarily equal to true types θ_i) and from this are computed the allocations $k_i(\hat{\theta})$ and monetary payments (to the mechanism) $t_i(\hat{\theta})$. The utility of each player u_i is then defined to be $v(k_i(\hat{\theta}), \theta_i) - t_i(\hat{\theta})$ (quasilinearity). Let $\theta_{-i} = (\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \theta_n)$ denote the types without agent i.

Definition 2 *The Generalized Vickrey Auction (GVA) is defined as follows:*

- $k(\theta) = k^*(\theta)$ *is chosen to maximize* $\sum_{i \in N} v(k_i(\theta), \theta_i)$
- $t(\theta)$ *is defined with each* $t_i(\theta) = \sum_{j \neq i} v(k_j^*(\theta_{-i}), \theta_j) - \sum_{j \neq i} v(k_j^*(\theta), \theta_j)$

where $k^*(\theta_{-i})$ *maximizes* $\sum_{j \neq i} v(k_j(\theta_{-i}), \theta_j)$.

The GVA has several important properties, primary among them being strategyproofness and allocative efficiency:

- The GVA is **strategyproof** in the sense that it is the dominant strategy of each agent to report its true type θ_i. That is, $v(k_i^*(\hat{\theta}_i, \theta_{-i}), \theta_i) - t_i(\hat{\theta}_i, \theta_{-i})$ is maximized for $\hat{\theta}_i = \theta_i$, for all θ_i and all θ_{-i}.
- The GVA is **allocatively efficient** in the sense that the allocation maximizes the total value across agents.

Unfortunately, we also have the following:

Proposition 1 *(Rothkopf, Pekec, and Harstad 1998) Computing the optimal allocation in the combinatorial auction problem is* **NP***-hard.*

One approach to combating this problem is to use a suboptimal allocation function within a VCG payment scheme.

Definition 3 *The class of VCG-based auctions are those where* $t_i(\theta) = \sum_{j \neq i} v(k_j(\theta_{-i}), \theta_j) - \sum_{j \neq i} v(k_j(\theta), \theta_j)$*, where* k *can be any feasible allocation function.*

Clearly, a VCG-based auction using the optimal allocation k^* is the GVA. The potential price of using suboptimal algorithms is strategyproofness. Nisan and Ronen (2000) show that this price is usually paid:

Theorem 1 *(Nisan and Ronen 2000) A VCG-based mechanism with a reasonable*[2] *and tractable allocation algorithm will not be strategyproof.*

3.2 False-name bids

False-name bids are a form of manipulation in which a bidder is able to assume multiple identities, submit different bids under each of those identities, and possibly benefit from the result (Yokoo, Sakurai, and Matsubara 2002).

Within the GVA, we allow false-name bid manipulation by using the following:

[2] An allocation algorithm is reasonable if, whenever only one agent values a particular item, that item is assigned to it.

Definition 4 *We allow each agent i to submit multiple bids $\{\hat{\theta}_i\} = \{\hat{\theta_{i_1}}, \ldots \hat{\theta_{i_k}}\}$ to the mechanism, which then operates on the bids $\bigcup_i \{\hat{\theta}_i\}$ (Notice that the mechanism cannot distinguish bids coming from the same agent). The utility to agent i is the utility it receives from allocations to any of its bids, minus the total net transfers.*

Definition 5 *A combinatorial auction is **false-name-proof** if it is the dominant strategy of agent i to submit only a single bid, as $\{\theta_i\}$.*

It is known that the GVA is not false-name-proof for general agent valuation functions. In fact, Yokoo et. al. (2002) show that this is true in any efficient mechanism:

Theorem 2 *(Yokoo, Sakurai, and Matsubara 2002) A combinatorial auction that is allocatively efficient will not be false-name-proof.*

4 Intractability of VCG-based Auction Manipulation

In this subsection we formally define the manipulation problem for VCG-based mechanisms, and prove it is **NP**-hard.

We focus first on a special case of the general combinatorial auction problem, in which bidders are single-minded:

Definition 6 *A type vector θ is **single-minded** if for every agent i, there exists a bundle of goods S_i and number r such that the valuation function v_i defined by θ satisfies $v_i(S_i) = r$ if $T \subseteq S_i$, and $v_i(S_i) = 0$ otherwise. We then denote the agent's bid as a pair, (S_i, r).*

In other words, the agent only values a single bundle for a single amount, and gets no additional or partial value from any other goods.

What we will do is consider the necessary properties of VCG-based auctions that will allow us to say that their manipulation is hard.

Definition 7 *An allocation algorithm k satisfies **greedy optimality** ("G-OPT") iff, whenever all bidders are single-minded[3], it never allocates to a set of bids that is a proper subset of the set allocated to by another valid allocation, and it never allocates goods to agents who don't request them.*

Intuitively, G-OPT just ensures that the algorithm at the very least does not leave goods unassigned (or assigned to agents who don't want them) when other bids request them.

Claim 1 *Any allocation algorithm can be modified to satisfy greedy optimality with at worst a linear-time slowdown and no loss in efficiency.*

[3]We state our assumptions with respect to the single-minded case, because that is all we require to prove the results. Both this definition and the following one could be naturally extended to general types, but this is not necessary for this paper.

Proof: First, check to make sure all bidders are single-minded (if not, return the regular allocation). Then, take the allocation and remove any goods assigned to an agent when those goods were not in the agent's desired bundle. Scan through the remaining bids, allocating to any that are consistent with the current allocation. At the end, either all bids are allocated to, or any remaining bids could not possibly be allocated to given the current allocation. ■

Definition 8 *An allocation algorithm satisfies* ***strong consumer sovereignty*** *("SCS") iff whenever all bidders are single-minded, no single agent would receive more value from receiving its bundle than the total value the algorithm achieves.*

Generally, consumer sovereignty requires that there exists an amount an agent can bid such that the algorithm will assign the desired bundle to that agent. Strong consumer sovereignty simply makes that bid amount more specific—the value of the allocation achieved without allocating to that bundle.

Claim 2 *Any allocation algorithm can be modified to satisfy strong consumer sovereignty with at worst a linear-time slowdown and no loss in efficiency.*

Proof: First, verify that all bidders are single-minded. Take the allocation, and scan through the remaining bids, and find the one with the most value (break ties arbitrarily). If this value is greater than the total value achieved by the original allocation, then alter the allocation to allocate to that single bid (removing all conflicts). ■

Definition 9 *The* ***single-minded manipulation problem*** *for a combinatorial auction is, given (θ, i) (all single-minded), returning a bid $\hat{\theta}_i = (\hat{S}_i, \hat{v}_i)$ that gives agent i at least ϵ additional units than what it would have received from being truthful in that auction, or outputting "impossible" if no such manipulation exists.*[4]

We can also define:

Definition 10 *The* ***general manipulation problem*** *for a combinatorial auction is, given (θ, i), returning a new type $\hat{\theta}_i$ that gives agent i at least ϵ additional units than what it would have received from being truthful in that auction, or outputting "impossible" if no such manipulation exists.*

Theorem 3 *Suppose the allocation algorithm of a combinatorial auction satisfies greedy optimality and strong consumer sovereignty and runs in polynomial time. Then the single-minded manipulation problem is* **NP***-hard.*

Corollary 1 *Suppose the allocation algorithm of a combinatorial auction satisfies greedy optimality and strong consumer sovereignty and runs in polynomial time. Then the general manipulation problem is* **NP***-hard.*

[4] ϵ is a parameter chosen to be sufficiently smaller than the precision otherwise provided by the bidding language. Specifically, choose ϵ such that for all bid amounts b, b', we have that $b - b' > \epsilon$.

Proof of Corollary: Since this problem generalizes the single-minded case, it is only harder. ■

To prove the theorem, we require the following fact:

Fact 1 *(Lehmann, O'Callaghan, and Shoham 2002) The following problem (*SINGLE-MINDED-WINNER-DETERMINATION*) is* **NP***-hard: given a single-minded type vector θ and a number K, determining whether an allocation can achieve efficiency of K units, and outputting that allocation if it exists.*

Proof of Theorem: Suppose the single-minded manipulation problem could be solved in polynomial time. We provide a solution to SINGLE-MINDED-WINNER-DETERMINATION.

Denote θ as a vector $((S_1, v_1), \ldots (S_n, v_n))$ where n is the number of agents. First, run the allocation algorithm on θ. Let the efficiency of the result be x. If $x \geq K$, then we are done: output "yes" and return the allocation. Suppose otherwise, so $K > x$.

Denote the set of all (relevant) goods $S = \bigcup_i S_i$. Define $\theta' = ((S_1 \cup \{g_1\}, v_1), \ldots, (S_n \cup \{g_n\}, v_n), (S, K - \epsilon))$, where each g_i is a new unique good, and we have introduced a new agent $n + 1$ that values the entire set of goods of the original auction at the value $K - \epsilon$. Run the polynomial-time algorithm for the manipulation problem on $(\theta', n + 1)$. Return the same answer that this algorithm returns. If the answer is "yes," then let the type vector of the manipulation for agent $n+1$ be $\hat{\theta'}_{n+1}$. Obtain the allocation achieving K by running the allocation algorithm on $(\theta_{-(n+1)}, \hat{\theta'}_{n+1})$ and returning the allocation to players 1 through n.

We now prove the main reduction: θ has an allocation achieving K units of efficiency if and only if agent $n + 1$ can achieve at least ϵ units of manipulation in θ'.

First, however, a lemma:

Lemma 1 *When run on θ', the allocation algorithm will allocate S to agent $n + 1$ (the new agent), giving it utility $K - x - \epsilon$.*

Proof of Lemma: We assumed that the allocation algorithm satisfies strong consumer sovereignty, and that the algorithm achieves x units of welfare when agent $n + 1$'s bid is not in the system. But we know that $K > x$, and thus $K - \epsilon > x$ (since ϵ is of arbitrarily small precision), and thus it follows by SCS that agent $n + 1$ is allocated S.

Since no one else can be allocated to, agent $n + 1$'s payment is x units, and its utility is $K - x - \epsilon$. ■

Suppose that an allocation of θ can give K units of efficiency (call this allocation $g^*(\theta)$). Recall x is the efficiency of the allocation algorithm $k(\theta)$, and $K > x$. By the lemma we know agent $n+1$ gets $K-x-\epsilon$ units of utility from being truthful. But then agent $n+1$ can achieve $K - x$ units of utility by forsaking its desired bundle and just bidding for the unique goods of the agents that are *not* in allocation g^* (and bidding high enough that the bid will be taken—such a number exists by SCS). The payment to agent $n + 1$ is just K (efficiency

for everyone else with it in the system) minus x (efficiency for everyone else without it). The algorithm will achieve K units for everyone else with agent $n+1$ because the remaining bids form the feasible allocation g^*, and so by G-OPT, the allocation algorithm will allocate to them. Thus, agent $n+1$ can manipulate to achieve an extra ϵ unit of utility.

Suppose, on the other hand, that we know agent $n+1$ can manipulate to achieve ϵ more utility, so that it receives $K-x$ units of utility. First, it is clear that this utility cannot consist in it actually making its bid for S. If it did, then (1) it would get $K-\epsilon$ units of utility and (2) all other agents would receive zero utility (because no other bid could possibly be made). Since the other agents would get x without agent $n+1$, the most it could receive in this way is the same $K-x-\epsilon$ achieved before. But then all of its utility must come from payment (i.e., it must be paid $K-x$ units). Since x is the utility for everyone else without agent $n+1$, the allocation its manipulation induces must achieve K units of efficiency for everyone else in this auction. Moreover, since the same allocation is possible in the original (unique goods create no new conflicts), the same K units must have been achievable in the original auction. ■

Nisan (2003) achieves an algorithm with a polynomially-bounded approximation ratio for general combinatorial auctions.[5] Combined with our theorem and the claims above, this leads to the following corollary, a counterpoint to the negative result in Nisan & Ronen (2000):

Corollary 2 *There exists a tractable and hard-to-manipulate mechanism with a polynomially-bounded approximation ratio.*

5 Intractability of GVA False-Name Bid Manipulation

In this section we define the false-name manipulation problem and demonstrate that this manipulation problem in the GVA is **NP**-hard.

Definition 11 *The **false-name-manipulation problem** for a combinatorial auction is, given a type vector for all agents θ and agent i, determining a set of bids $\{\theta_{i_1}, \ldots \theta_{i_k}\}$ such that when the auction operates on $(\theta_{-i}, \{\theta_{i_1}, \ldots \theta_{i_k}\})$, the resulting utility to agent i increases by ϵ units.*

Theorem 4 *The false-name-manipulation problem for the GVA is* **NP***-hard (even when there are only two agents).*

Fact 2 *(Garey and Johnson 1979) The following problem (*EXACT-COVER-BY-THREE-SETS*) is* **NP***-hard: Given a universe X, $|X| = 3q$, and a collection C of 3-element subsets of X, does there exist a subcollection of size q that partitions X?*

[5] The algorithm effectively works by finding sets S_i for each bidder maximizing $v(S_i, \theta_i)/|S_i|$ and then allocating those sets to each in decreasing order of $v(S_i, \theta_i)/|S_i|$.

Proof: We reduce from EXACT-COVER-BY-THREE-SETS. Create two agents, 1 and 2. The set of goods will be precisely the elements of X. Agent 1 has the following valuation function θ_1:

$$v_1(S) = \begin{cases} q+\epsilon & : \quad S = X \\ 0 & : \quad S \neq X \end{cases}$$

(where ϵ is some arbitrary, small constant). Agent 2, alternatively, has the following valuation function:

$$v_2(S) = \begin{cases} \overline{v} & : \quad |S| \geq 4 \\ 1 & : \quad |S| \leq 2 \\ 1 & : \quad |S| = 3, S \in C \\ 2 & : \quad |S| = 3, S \notin C \end{cases}$$

(where $\overline{v}$ is some arbitrary number much larger than q). Clearly, the appropriate allocation under truthful reporting is to give at least four elements to agent 2 and give nothing to agent 1. By strategyproofness there is no way for agent 1 to lie to achieve more than this zero utility. However, with false-name bids this is sometimes possible. Specifically, agent 1 can manipulate by false-name bids to achieve at least ϵ units of utility if and only if there is a partition of X through elements of C.

Suppose there is a partition of X, call it $C' \subseteq C$. Then consider the following strategy for agent 1: agent 1 takes on q identities, each of whom bids for only an exact subset $C_i \in C'$, with any bid amount sufficiently high to win (say, $\overline{v}+1$). Clearly, agent 1 will win all of the goods. Moreover, each bidder will be charged exactly 1 unit, because if any individual bidder would be removed from the system, agent 2 would get those three elements, which it values only at 1 unit apiece. Thus, the total utility to agent 1 is ϵ (bundle gives $q+\epsilon$, but it pays q).

Suppose, alternatively, that we have a false-name bid manipulation achieving ϵ utility for agent 1. Then we must demonstrate there is a partition of X using subsets from C. First, some straightforward but necessary observations. Since all the transfers are nonnegative, and since agent 1 only gets utility from the entire bundle, it can only improve its utility by obtaining the entire bundle of goods. Also, agent 1 will gain nothing by submitting bids that would not win—such bids would either have no effect or would just increase its payments. Moreover, if a bidder wins a certain subset, specifying that bidder's valuations on any other subset will have no, or only an adverse, effect on agent 1's total utility. Thus, we can reduce the false-name bids of agent 1 to specifying a partition of X, with each identity bidding for a single disjoint subset, and with the bid amount just some large number guaranteeing the bid is made (notice that changing the size of the bid amount, given that the bid will be made, will not affect the agent's utility). To reason about potential partitions, notice that no identity can afford to take a subset of size larger than 3, since it would have to pay $\overline{v}$ (agent 2 would value that subset at $\overline{v}$), which is already far more than agent 1's valuation

for the entire bundle. Moreover, every identity must pay at least 1 unit, since no matter what the identity receives, agent 2 would value that bundle for at least that much. Thus, there can be at most q identities that are allocated to, since otherwise agent 1 will pay more than its value for the entire bundle. Given that $|X| = 3q$ and that no identity can get more than 3 goods, it follows that each identity must receive 3 distinct goods. But the only way such a manipulation will succeed in not costing agent 1 more than its value is if each such bid is charged 1 unit, and this happens if and only if each subset is a member of C. In such a case, we have the partition as desired. ■

6 Discussion

Hardness of VCG-based Auction Manipulation: It is important at this stage to ask what is driving our result that although strict truthfulness and reasonableness may be impossible, reasonableness and hardness of manipulation are certainly possible with tractable mechanisms. In fact, the hardness is driven by the difficulty of the winner-determination problem and the so-called "self-correcting" property of VCG-based mechanisms. Specifically, it is an often noted property of VCG-based mechanisms that any potential manipulation always improves the efficiency of the allocation—i.e., if by manipulating an agent forces the allocation from A to A', then it must be the case that A' has overall better welfare for *all* agents than A did.

Thus, the reason that manipulation in these cases is hard is that the agents are ultimately trying to solve the same problem as the allocation algorithm—trying to find the allocation best maximizing welfare for all agents. The nature of the reduction and the required assumptions are all just there to ensure that (1) any (as opposed to say, optimal) manipulation by the agent will require solving a hard problem and (2) the allocation algorithm is sufficiently sophisticated (or "controllable") to take advantage of the "hints" a manipulative agent would provide.

Hardness of GVA False-name Manipulation: Unlike the hard-to-manipulate result for an agent with a single identity, our false-name manipulation result does not depend on the difficulty of the winner-determination problem. On the contrary, the result requires that the winner-determination problem is solved optimally, and shows that in such a case, manipulation is **NP**-hard. This is a drawback— with false-name bids, we (currently) must settle for a mechanism that is *intractable*, allocatively efficient, and hard to manipulate.

We leave as an open question whether it is possible to design a tractable, reasonable, and hard-to-false-name-manipulate mechanism. Certainly, an agent is able to derive a substantial amount of power when given false-name bids and a suboptimal allocation algorithm. For instance, it is possible for an agent to create (through several false names) a situation where the algorithm will be inefficient, then create a new agent who fixes this inefficiency (through unique goods, like in the intractability result above), and get paid for the result.

6.1 Critique and Future Work

Mechanism design for bounded-rational agents is an important but difficult endeavor. Many negative results point to the complexity of game-theoretic solution concepts and suggest the importance of finding a way to relax the assumption that agents will always play an exact equilibrium. A new theory is required that can provide a normative approach for design with respect to self-interested but bounded-rational agents. This is a difficult problem, however, because models of bounded-rationality are hard to formalize and do not allow ready analysis, especially in equilibrium.

Our results are worst-case in nature: all we guarantee is that manipulation is hard in some cases. Manipulation may, in general, remain possible for an agent. This issue is certainly not unique to this work—the results of Conitzer & Sandholm (2002) have the same nature. Yet, in a mechanism design community spoiled by the ironclad, "every-case" guarantees of dominant strategy equilibrium, the difference is striking.

What we would really like is some statement like: "the dominant strategy equilibrium for polynomial-time bounded agents is truth revelation". That is, whatever the other (polynomial-time) agents are capable of reporting, given knowledge of the mechanism rules any polynomial-time bounded agent will always do best by being truthful. O'Connell (2000) has suggested some definitions of this form. Our current model invokes no equilibrium concept, but instead adopts truth-revelation as a default strategy and shows that, whatever other agents are doing (we need not assume they are polynomially-bounded or truthful), a polynomially-bounded agent will not be able to manipulate every time such a manipulation exists. Notice that agents are not reasoning about the bounded-rationality of other agents, but merely supposed to fall-back on truth-revelation because manipulation is hard.

We believe that real progress will require a formal reintroduction of equilibrium concepts, albeit across bounded-rational agents. In this vein, Rubinstein (1998) has considered the equilibrium across agents represented as finite-state machines with limited state. Other work has modeled agents with cost-based models of computation (assumed common-knowledge), but assumed that agents can perform optimal meta-deliberation which is itself often intractable (Larson and Sandholm 2001). In our setting, it is not clear how a bounded-rational equilibrium can be implemented when the strategy space for an agent is exponential yet an agent has polynomial-time bounds. What section of the search space should that polynomial-time bound be spent searching, and how might agents "coordinate" their reasoning with others?

7 Conclusions

To conclude, we have demonstrated two primary complexity results—the **NP**-hardness of manipulating suboptimal and tractable VCG-based auctions and the **NP**-hardness of manipulating the GVA using false-name bids. In the process, we suggested a model for sidestepping

negative results in mechanism design—by relaxing strategyproofness to a type of "hardness-of-manipulation." Although the guarantee this latter notion provides is substantially weaker than strategyproofness, the demands of practicality, in light of substantial negative results, make it at least a necessary and somewhat comforting recourse.

References

Archer, Aaron, Joan Feigenbaum, Arvind Krishnamurthy, Rahul Sami, , and Scott Shenker, 2003, "Approximation and collusion in multicast cost sharing," in *Games and Economic Behavior*.

Archer, Aaron, Christos Papadimitriou, Kunal Talwar, and Eva Tardos, 2003, "An approximate truthful mechanism for combinatorial auctions with single parameter agents," in *Proc. 14th ACM Symposium on Discrete Algorithms*.

Bartholdi, J J, 1989, "The computatonal difficulty of manipulating an election," *Social Choice and Welfare*, 6, 227–241.

Conitzer, Vincent, and Tuomas Sandholm, 2002, "Complexity of Manipulating Elections with Few Candidates," in *Proceedings of the National Conference on Artificial Intelligence (AAAI)*.

Conitzer, Vincent, and Tuomas Sandholm, 2003, "Universal Voting Protocol Tweaks to Make Manipulation Hard," in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*.

Garey, Michael R, and David S Johnson, 1979, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W.H.Freeman and Company, New York.

Jackson, Matthew O., 2000, "Mechanism theory," in *The Encyclopedia of Life Support Systems*, EOLSS Publishers.

Kothari, Anshul, David C. Parkes, and Subhash Suri, 2004, "Approximately-Strategyproof and Tractable Multi-Unit Auctions," *Decision Support Systems*, Special issue dedicated to the Fourth ACM Conference on Electronic Commerce (EC'03), Forthcoming.

Larson, Kate, and Tuomas Sandholm, 2001, "Bargaining with limited computation: Deliberation equilibrium," *Artificial Intelligence*, 132, 183–217.

Lavi, Ron, Ahuva Mu'alem, and Noam Nisan, 2003, "Towards a Characterization of Truthful Combinatorial Auctions," in *44th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pp. 574–583.

Lehmann, Daniel, Liadan Ita O'Callaghan, and Yoav Shoham, 2002, "Truth Revelation in Approximately Efficient Combinatorial Auctions," *Journal of the ACM*, 49, 577–602.

Nisan, Noam, 2003, "Personal Communication," .

Nisan, Noam, and Amir Ronen, 2000, "Computationally feasible VCG mechanisms," in *ACM Conference on Electronic Commerce*, pp. 242–252.

Nisan, Noam, and Amir Ronen, 2001, "Algorithmic Mechanism Design," *Games and Economic Behavior*, 35, 166–196.

O'Connell, Thomas, 2000, "Bounded Rationality in Repeated Games and Mechanism Design in Computational Settings," Ph.D. thesis, University at Albany, State University of New York.

Parkes, David C., 2003, "Auction Design with Costly Preference Elicitation," *Annals of Mathematics and AI*, Special Issue on the Foundations of Electronic Commerce, Forthcoming.

Rothkopf, M. H., A. Pekec, and R. M. Harstad, 1998, "Computationally Manageable Combinatorial Auctions," *Management Science*, 44.

Rubinstein, Ariel, 1998, *Modeling Bounded Rationality*, MIT Press.

Schummer, James, 2002, "Almost Dominant Strategy Implementation," Discussion paper, MEDS Department, Kellogg Graduate School of Management.

Varian, Hal R, 1995, "Economic Mechanism Design for Computerized Agents," in *Proc. USENIX Workshop on Electronic Commerce*. Minor update, 2000.

Yokoo, Makato, Yuko Sakurai, and Shigeo Matsubara, 2002, "The Effect of False-name Bids in Combinatorial Auctions: New Fraud in Internet Auctions," in *Games and Economic Behavior*.