Computer Science and Artificial Intelligence Laboratory
Technical Report

MIT-CSAIL-TR-2008-041 July 2, 2008

Knowledge Benchmarks in Adversarial Mechanism Design and Implementation in Surviving Strategies (Part I)

Jing Chen and Silvio Micali

massachusetts institute of technology, cambridge, ma 02139 usa — www.csail.mit.edu

Knowledge Benchmarks in Adversarial Mechanism Design and Implementation in Surviving Strategies (Part I)

Jing Chen and Silvio Micali

July 2, 2008

Abstract

We put forward new benchmarks and solution concepts for *Adversarial Mechanism Design*, as defined by [MV07.a], and we exemplify them in the case of truly combinatorial auctions.

We benchmark the *combined performance* (the sum of the auction's efficiency and revenue) of a truly combinatorial auction against a very relevant but private knowledge of the players: essentially, *the maximum revenue that the best informed player could guarantee if he were the seller.* (I.e., by offering each other player a subset of the goods for a take-it-or-leave-it price.)

We achieve this natural benchmark within a factor of 2, by means of a new and probabilistic auction mechanism, in *surviving strategies.* That is, the above performance of our mechanism is guaranteed in any rational play, independent of any possible beliefs of the players. Indeed, our performance guarantee holds for *any possible choice of strategies,* so long as each player chooses a strategy among those surviving iterated elimination of dominated strategies.

Our mechanism is extremely robust. Namely, its performance guarantees hold even if all but one of the players collude (together or in separate groups) in any possible but reasonable way. Essentially, the only restriction for the collective utility function of a collusive subset S of the players is the following: the collective utility increases when one member of S is allocated a subset of the goods "individually better" for him and/or his "individual price" is smaller, while the allocations and prices of all other members of S stay the same.

Our results improve on the yet unpublished ones of [MV07.b]. The second part of this paper, dealing with a more aggressive benchmark (essentially, the maximum welfare privately known to the players) is forthcoming.

1 Combinatorial Auctions

Auction Contexts. A combinatorial auction *context* C consists of a triple (N, G, TV), where N is the (finite) set of *players,* G is the finite set of *goods,* and TV is a *profile* (i.e., a vector indexed by the players) of valuations[1] of G, the *true valuations.* We say that C is a $n \times m$ auction context if $|N| = n$ and $|G| = m$. If C is such a context, we assume $N = \{1, \ldots, n\}$ and $G = \{g_1, \ldots, g_m\}$. As customary, let $\mathbb{C}$ be a subset of N, we denote the set $N \backslash \mathbb{C}$ by $-\mathbb{C}$ when N is clear from context. (For simplicity, we denote the set $N \backslash \{i\}$ by $-i$, instead of $-\{i\}$.) Let p be a profile, we use $p_{\mathbb{C}}$ to denote a sub-profile with respect to $\mathbb{C}$, that is, a vector indexed by the players in $\mathbb{C}$.

An *outcome* Ω for an $n \times m$ auction context (N, G, TV) consists of: (1) an *allocation* A, that is, a partition of G into $n+1$ subsets, $A = (A_0, A_1, \ldots, A_n)$, and (2) a *price profile* P, that is, a profile of real numbers. We refer to such an Ω as a $n \times m$ outcome, and to A_0 as the set of unallocated goods. For each player i, we refer to A_i as the set of goods allocated to i, and to P_i as the price of i. We refer to $\sum_i TV_i(A_i)$ as the *social welfare* of allocation A, and denote it by $SW(A, TV)$ —or more simply by SW when TV and A are clear from context. We refer to $\sum_i P_i$ as the *revenue* of Ω, and denote it by $REV(\Omega)$ —or more simply by REV when Ω is clear from context.

Player i's *utility function*, U_i, is defined as follows: for any context (N, G, TV) and outcome $\Omega = (A, P)$, $U_i(\Omega, TV) = TV_i(A_i) - P_i$. When Ω and TV are clear from context, i's *utility* refers to the value $u_i = U_i(\Omega, TV)$. (I.e., u_i is i's true valuation on the goods allocated to him minus his price.)

Public Extensive-Form Mechanisms and Auctions. Together with a $n \times m$ context C, a $n \times m$ mechanism M yields an auction (C, M). We shall use mechanisms M of *public extensive-form.* An auction with such a mechanism M is played in k stages, where k is an integer greater than 1. At each stage j, each player i publicly announces a string x_i^j simultaneously with the other players. Then M is evaluated on the profile x, where x_i is the sequence $x_i^1, \ldots, x_i^k$, so as to produce a final outcome. Here, a strategy of a player i is a deterministic function σ_i —although player i may want to use a mixture of different strategies. (A player i chooses σ_i based on his private inputs; that is, his true valuation TV_i and —as we shall soon see— his knowledge K_i.) For each stage j, σ_i selects the string x_i^j on inputs (1) j itself; (2) all strings announced by the players up to stage $j-1$ included; and (3) player i's private inputs.

We refer to a strategy profile σ as a *play* of (C, M), or of M when C is clear. Since a play σ

[1] A *valuation* of a set S is a mapping from S's subsets to the non-negative reals.

determines the profile x, it is convenient to view our M as mapping plays σ to outcomes $M(\sigma)$, and to refer to its allocation and price component as M_a and M_p: that is, $M(\sigma) = (M_a(\sigma), M_p(\sigma))$ for all plays σ.

Utilities, Social Welfare, and Revenue. In a play σ of an auction (C, M), i's *utility*, $u_i(\sigma)$, is defined as $U_i((M_a(\sigma), M_p(\sigma)), TV)$. To emphasize player i, we write $\sigma = (\sigma_i \sqcup \sigma_{-i})$.

When M is probabilistic, the expected social welfare and the expected revenue of a strategy profile σ, $\mathbb{E}[SW_\sigma]$ and $\mathbb{E}[REV_\sigma]$, are respectively defined as $\mathbb{E}[SW(M_a(\sigma), TV)]$ and $\mathbb{E}[REV(M_p(\sigma))]$. Player i's expected utility with respect to σ, $\mathbb{E}[u_i(\sigma)]$, is defined as $\mathbb{E}[U_i((M_a(\sigma), M_p(\sigma)), TV)]$. More simply, we use $\mathbb{E}[SW]$, $\mathbb{E}[REV]$ and $\mathbb{E}[u_i]$ respectively, when σ is clear from context.

2 A General Model of Collusion

In our analysis, *we* distinguish between independent and collusive players. By contrast, *our mechanisms* have no idea of which players are collusive or independent. Indeed, we assume that collusion is "illegal". Accordingly, no special bids are envisaged for collusive players, and collusive players have no desire to disclose themselves. Therefore, our mechanisms produce outcomes specifying allocations and prices treating each player as independent. But since legal deterrents are often insufficient, we want to design our mechanisms so as to be robust in the presence of collusion.

We now allow for collusive players. We assume that the set of players comprises a collection of mutually disjoint sets, the *collusive sets.* The members of a collusive set S do not try to maximize their *individual* utilities. Rather, they try to maximize the *collective* utility of S, as determined by a *collusive function* u_S: an unrestricted function of TV_S, the true-valuation sub-profile of the members of S, of P_S —that is the profile of prices of the members of S, and of A_S, that is of the subset of goods allocated to the members of S.

Note that, in order to maximize S's collective utility, it is not necessary that the set S itself, the collusive utility function u_S, the sub-profile TV_S, or any other knowledge S's members may have, be common knowledge within S. For instance, it suffices that just one member of S has all this information, and that he is indeed in a position to dictate the action for all members of S.

Note too that a player i belonging to a collusive set S does not imply that i's individual utility function u_i ceases from "existing". It only means that i will now act so as to maximize S's collective utility. But the function u_i, that is $TV_i(A_i) - P_i$, could still be relevant. For instance, i may have joined a collusive set S under an agreement that would compensate him —e.g., by a monetary side payment— for any losses he individually incur when playing a strategy that maximizes S' collective utility. And to compute such compensation, u_i continues to be relevant. Also, to be reasonable, the collective utility of S must ultimately be related to the individual utilities of its members. For example, the members of S may have agreed to maximizing the sum of their individual utilities, but then split the proceeds in some, not necessarily fair, way. As for another example, the collective utility of S my coincide with the individual utility of just one of his members, who may convince the other to go along, in a variety of ways, including blackmail, promise of future cooperation in totally different settings, a lump-sum payment, etc.

Since the enormous diversity among collusive utility functions, it is important to point out that our results hold for essentially all reasonable choices of them. That is, they hold for all u_S satisfying a "minimal property. The property states that, all other things being equal, the utility of a collusive set S solely depends on the value that an individual member of the set receives. Namely, the collective utility stays the same, if the only change consists of swapping the goods allocated to a member i with another set that i values the same (for which he pays the same price); and that the collective utility increases if the only change is allocating to i another subset of goods which he values more (keeping paying the same price for them). Let us now be more precise.

Definition 1. (Minimal Reasonableness Axiom) *We say that the utility function u_S of a collusive set S satisfies the minimal reasonableness axiom (MRA) if, for (1) all players $i \in S$ and (2) all outcomes (A, P) and (A', P') such that $(A_j, P_j) = (A'_j, P'_j)$ whenever $j \in S \setminus \{i\}$, we have:*

- *If $TV_i(A_i) - P_i = TV_i(A'_i) - P'_i$, then $u_S((A, P), TV) = u_S((A', P'), TV)$;*
- *If $TV_i(A_i) - P_i > TV_i(A'_i) - P'_i$, then $u_S((A, P), TV) > u_S((A', P'), TV)$.*

(Note that the first implication in the above definition does not imply that $(A_i, P_i) = (A'_i, P'_i)$.)

For uniformity of treatment and ease of notation, we find it convenient to assume that every player i belongs to a collusive set. That is,

Definition 2. *Let $\mathbb{C}$ be a profile of subsets of players, and u a vector indexed by the subsets $\mathbb{C}_i$. We say that $(\mathbb{C}, I, u)$ is a* minimally reasonable collusion system *if the following five properties are satisfied:*

1. *For each player i, $i \in \mathbb{C}_i$*
2. *For all players i and j, either $\mathbb{C}_i = \mathbb{C}_j$ or $\mathbb{C}_i \cap \mathbb{C}_j = \emptyset$.*

3. $I = \{i : \mathbb{C}_i$ *has cardinality* $1\}$

4. For any player i, $u_{\mathbb{C}_i}$ *is a (collective) utility function for* $\mathbb{C}_i$ *satisfying the MRA*

5. $u_{\mathbb{C}_i}((A,P),TV) = TV_i(A_i) - P_i$ *whenever* $i \in I$.

We refer to a player in I *as* independent*, and to one not in* I *as* collusive*.*

Whether or not i is independent, we define i's *individual utility function* u_i as follows:

$$u_i((A,P),TV) = TV_i(A_i) - P_i.$$

Accordingly, u_i coincides with $u_{\mathbb{C}_i}$ (and $u_{\{i\}}$) if and only if i is independent.

Since outcomes are ultimately determined by the players' strategies, when the true-valuation profile TV is understood, we may consider each $u_{\mathbb{C}_i}$ (respectively, u_i) to be the (possibly probabilistic) function that, for any strategy profile σ, returns the utility of the collusive set $\mathbb{C}_i$ (respectively, the individual utility of player i) under σ.

3 A Safe Solution Concept

Typically, mechanism design aims at guaranteeing a given property $\mathbb{P}$ "at equilibrium." Equilibria, however, are a fragile notion, because they do not solely depend on the players' rationality, but also on their *beliefs.* In case of a combinatorial auction, saying that a profile of strategies σ is an equilibrium only means that, for every player i, deviating from σ_i is an irrational thing to do (i.e., yields a lower utility for i) only if he believes that any other player j will stick to his strategy σ_j. Accordingly, if some players believe that the equilibrium about to be played is σ while others believe it is τ, the auction may not end up in any equilibrium at all, since "mixing and matching" the strategies of σ and τ needs not to result in an equilibrium!

Accordingly, our mechanism works for a much more robust set of plays: namely,

For any profile of strategies that survives iterated elimination of dominated strategies.

That is, our notion of implementation is "equilibrium-less," depends solely on the players' rationality (rather than their beliefs), and is immune to any "equilibrium selection" problem.

We formalize this strong notion below, directly for the collusive setting.

Definition 3. *Relative to a minimally reasonable collusion system* $(\mathbb{C}, I, u)$ *for an auction context* C, *we say that a strategy* σ_i *for player* i *is* dominated *if there is another strategy* σ'_i *for* i *such that*

- *For all strategy sub-profiles* σ'_{-i}*:* $\mathbb{E}[u_{\mathbb{C}_i}(\sigma_i \sqcup \sigma'_{-i})] \le \mathbb{E}[u_{\mathbb{C}_i}(\sigma')]$*; and*
- *For some strategy sub-profile* σ'_{-i}*:* $\mathbb{E}[u_{\mathbb{C}_i}(\sigma_i \sqcup \sigma'_{-i})] < \mathbb{E}[u_{\mathbb{C}_i}(\sigma')]$.

We say that σ_i *is* undominated *if it is not dominated.*

The following definition is a variant of Definition 60.2 of [OR94].[2]

Definition 4. *Let* Σ *be a profile such that* Σ_i *is the set of all available strategies for player* i*, and* X *be a profile such that* $X_i \subseteq \Sigma_i$ *for any* i*. We say that* X survives iterated elimination of dominated strategies *if there is a collection* $(X^t)_{t=0}^T$ *of profiles such that*

1. $X^0 = \Sigma$ *and* $X^T = X$.
2. $X_j^{t+1} \subseteq X_j^t$ *for any player* $j \in N$ *and any* $t = 0, \ldots, T-1$.
3. *For any* $t = 0, \ldots, T-1$ *and any player* $j \in N$*, every strategy in* X_j^{t+1} *is undominated over* X^t*, i.e., undominated assuming that for every player* i*, the set of all available strategies is* X_i^t.
4. *For any player* j*, no action in* X_j^T *is dominated over* X^T.

Definition 5. (Implementation in Surviving Strategies) *We say that a mechanism* M *implements a property* $\mathbb{P}$ *in surviving strategies if, for any strategy profile* σ *that survives iterated elimination of dominated strategies,* $\mathbb{P}$ *is expected to hold over outcomes distributed according to* $M(\sigma)$.

4 Our Knowledge Benchmark

4.1 Our Knowledge Model

At a minimum, each player i knows his own true valuation TV_i, but he might also have other knowledge, K_i, about the true valuations of the other players. The kind of knowledge that we deem relevant is the revenue that i could guarantee to himself if he were the seller of the goods to the other players. This is made more precise as follows.

Definition 6. *Let* (A, P) *be an outcome for a combinatorial context* (N, G, TV)*. We say that* (A, P) *is* canonical *if* $\forall i \in N$*: (1)* P_i *is a non-negative integer, and (2)* $P_i = 0$ *whenever* $A_i = \emptyset$*. We say that* (A, P) *is* feasible *if it is canonical and* $P_j \le TV_j(A_j)$ *whenever* $A_j \neq \emptyset$.

[2]The difference between the two definitions is in item 3, where they allow X_j^{t+1} to still contain some strategies dominated over X^t.

If (A, P) is a canonical (respectively, feasible) outcome such that $A_j = \emptyset$ whenever j lies outside a given subset S of the players, we further say that (A, P) is canonical (respectively, feasible) for S.

Notice that a feasible outcome (A, P) essentially consists of a way of selling the goods that generates revenue equal to $\sum_j P_j$: namely, offer to each player j to either (1) buy bundle A_j for price P_j, or (2) receive no good and pay nothing. (Receiving positive utility, each player j prefers to accept such an offer to buy.)

Definition 7. *Let $C = (N, G, TV)$ be a combinatorial auction context, and K a profile.*

- *We say that C has* external knowledge K *if (1) $K_i \subset \mathbb{V}_{-i}$, where $\mathbb{V}_{-i}$ is the set of all possible valuation sub-profiles for the players in $-i$, and (2) $TV_{-i} \in K_i$.*

  We refer to K_i as the external knowledge of i. In essence, it represents what player i knows about the other players true valuations: namely, he knows that TV_{-i} must be one of the sub-profiles in K_i, but he does not know which one. The external knowledge of i is i's private information. At least, no information about K_i is known to the auction designer. (The results of this paper would continue to hold under a more general definition of K. In particular, players may know also information about each other's knowledge. For instance, certain true-valuation information can be common knowledge.)

- *By $F(K_i)$ we denote the set of outcomes that are feasible for every valuation sub-profile in K_i.*

- *By $MEW_i(K)$ we define the* maximum external welfare *known to player i, that is,*

  $$MEW_i(K) = \max_{\Omega \in F(K_i)} REV(\Omega).$$

  Thus, $MEW_i(K)$ represents the maximum revenue that i could guarantee if he were in charge to sell the goods of G to the other players by making each one of them a take-it-or-leave-it offer for some subset of the goods.

4.2 The MEW Benchmark

As argued, a benchmark maps a profile of knowledge to a non-negative number, and we might as well restrict our benchmarks to just the knowledge of the independent players. Accordingly, we put forward the following definition.

Definition 8. *In a truly combinatorial auction with external-knowledge profile K, we define the* maximum (known) external welfare *benchmark, MEW, as follows: for all collusion systems* $(\mathbb{C}, I, u)$*,*

$$MEW(K) = \max_{i \in I} MEW_i(K).$$

(It is easy to see that MEW is "player-monotone" —see [MV07.a].)

5 Our Mechanism

Our mechanism is of public extensive-form, and actually consists of three stages: two player stages followed by a final mechanism stage, where the mechanism produces the final outcome (A, P).

In the first stage, each player i publicly (and simultaneously with the others) announces (1) a canonical outcome Ω_i for the players in $-i$; and (2) a subset of goods S_i. (Allegedly, Ω_i is actually feasible, and indeed represents the "best way known to i to sell the goods to the other players." Allegedly too, S_i is i's favorite subset of goods, that is the one i values the most.)

After the first stage, everyone can compute (a) the revenue R_i of Ω_i for each player i, (b) the highest and second highest of such revenues, respectively denoted by $R_\star$ and R', and (c) the player whose announced outcome has the highest revenue —the lexicographically first player in case of "ties". Such player is called the "star player" and is denoted by "$\star$". (Thus, $\star \in N$.)

In the second stage, each player i, envisioned to receive a non-empty set of goods (for a positive price) in $\Omega_\star$, publicly (and simultaneously with the other such players) answers yes or no to the following implicit question: "are you willing to pay your envisioned price for your envisioned goods?" (The players not receiving any goods according to $\Omega_\star$ announce the empty string.)

After the second stage, for each asked player i who answers no, the star player is punished with a fine equal to the price he envisioned for i.

In the third and final stage, the mechanism flips a fair coin. If Heads, $S_\star$ is given to the star player at no *additional* charge (and thus player $\star$ pays nothing altogether if no player says no in the second stage). If Tails, (1) the goods are sold according to $\Omega_\star$ to the players who answered yes in the second stage, (2) all the revenue generated by this sale is given to the star player, and (3) the star player additionally pays R' to the seller/auctioneer. (Thus, the star player pays only R' if he has not been fined.) A more precise description of our mechanism is given below. In it, for convenience, we also include three "variable-update stages" and mark them by the symbol "$\bullet$". In such stages the contents of some public variables are updated based on the strings announced so far.

Mechanism $\mathcal{M}$

- Set $A_i = \emptyset$ and $P_i = 0$ for each player i.

1. Each player i simultaneously and publicly announces (1) a canonical outcome for $-i$, $\Omega_i = (\alpha^i, \pi^i)$, and (2) a subset S_i of the goods.

- Set: $R_i = REV(\Omega_i)$ for each player i, $\star = \arg\max_i R_i$, and $R' = \max_{i \neq \star} R_i$.

  (We shall refer to player $\star$ as the "star player", and to R' as the "second highest revenue".)

2. Each player i such that $\alpha_i^\star \neq \emptyset$ simultaneously and publicly announces YES or NO.

- For each player i who announces NO, $P_\star = P_\star + \pi_i^\star$.

3. Publicly flip a fair coin.

   – If Heads, reset $A_\star = S_\star$.

   – If Tails: (1) reset $P_\star = P_\star + R'$; and (2) for each player i who announced YES in Stage 2, reset: $A_i = \alpha_i^\star$, $P_i = \pi_i^\star$, and $P_\star = P_\star - P_i$.

Comment. The outcome (A, P) may not be canonical, as the price of the star player may be non-zero even though he may receive nothing.

6 Analysis of Our Mechanism

Theorem 1. $\forall$ *integers* n *and* m*;* $\forall$ *true-valuation profiles* TV*, external knowledge* K*, and minimally reasonable collusion system* $(\mathbb{C}, I, u)$ *for* n *players and* m *goods; and* $\forall$ *surviving play* σ *of* $\mathcal{M}$*:*

the sum of the expected social welfare and the expected revenue of $\mathcal{M}(\sigma)$ *is at least*

$$\frac{MEW_I(K)}{2}.$$

Proof. We base our proof on that of three simpler claims about the actions of the players in any execution of $\mathcal{M}(\sigma)$ when σ —as well as n, m, TV, K, $(\mathbb{C}, I, u)$— are as above.

Without loss of generality, we assume that $\forall i \notin \mathbb{C}_\star$ i answers YES in Stage 2 if $TV_i(\alpha_i^\star) = \pi_i^\star$.[3]

[3] Else, we can easily modify $\mathcal{M}$ such that when Stage 3 gets Tails, reset $P_i = \pi_i^\star - \epsilon$ where ϵ is an arbitrarily small positive number. Doing this only changes our benchmark in Theorem 1 from $\frac{MEW_I(K)}{2}$ to $\frac{MEW_I(K)}{2} - \tau$, where τ is another arbitrarily small number. Therefore we ignore this point in the analysis later.

Claim 1. *In Stage 2, the following two implications hold for all $i \notin \mathbb{C}_\star$:*

1. *i answers YES if $TV_i(\alpha_i^\star) > \pi_i^\star$, and*
2. *i answers NO if $TV_i(\alpha_i^\star) < \pi_i^\star$.*

Proof of Claim 1. We restrict ourselves to just prove the first implication. (The proof of the second implication is indeed totally symmetric.) We proceed by contradiction. Assume σ is such that in Stage 2 some player $i \notin \mathbb{C}_\star$ announces NO when $TV_i(\alpha_i^\star) > \pi_i^\star$, and consider the following alternative strategy for player i.

Strategy σ_i':

- *Stage 1.* Run σ_i (with stage input "1" and private inputs TV_i and K_i) and announce Ω_i and S_i as σ_i does.
- *Stage 2.* If $i \in \mathbb{C}_\star$ or $TV_i(\alpha_i^\star) = \pi_i^\star$, run σ_i and answer whatever σ_i does.[4]
  Else, answer YES if $TV_i(\alpha_i^\star) > \pi_i^\star$, and NO if $TV_i(\alpha_i^\star) < \pi_i^\star$.

We derive a contradiction by proving that σ_i is dominated over Σ^0 by the strategy σ_i'. Therefore $\sigma_i \notin \Sigma_i^1$, and thus $\sigma_i \notin \Sigma_i$, which contradicts the fact that σ_i is a surviving strategy.

Towards proving that σ_i is dominated by σ_i', observe that, since σ_i' coincides with σ_i in Stage 1, the outcome profile Ω is the same in the plays σ and $\sigma_i' \sqcup \sigma_{-i}$. Accordingly, the star player too is the same in both plays, and thus so is $\mathbb{C}_\star$. Finally, since (by hypothesis) $i \notin \mathbb{C}_\star$ in play σ, $i \notin \mathbb{C}_\star$ also in play $\sigma_i' \sqcup \sigma_{-i}$.

Now consider any other player $j \in \mathbb{C}_i$ —if any— and notice that the subset of the goods allocated to him, A_j, is the same, under strategy profiles σ and $\sigma_i \sqcup \sigma_{-i}$, whenever $\mathcal{M}$'s coin tosses are the same. And the same holds for his final price, P_j. That is, the following proposition holds

(*) $\quad \forall j \in \mathbb{C}_i \setminus \{i\}$: $\quad \mathcal{M}_a(\sigma)_j = \mathcal{M}_a(\sigma_i' \sqcup \sigma_{-i})_j \quad$ and $\quad \mathcal{M}_p(\sigma)_j = \mathcal{M}_p(\sigma_i' \sqcup \sigma_{-i})_j$.

We now distinguish two cases, each occurring with probability 1/2.

(1) $\mathcal{M}$'s coin toss comes up Heads.

In this case, $u_i(\sigma) = u_i(\sigma_i' \sqcup \sigma_{-i}) = 0$, that is the individual utility of player i is 0, since only the star player receives goods. Therefore, because of this and Proposition (*), the MRA implies that

$$u_{\mathbb{C}_i}(\sigma) = u_{\mathbb{C}_i}(\sigma_i' \sqcup \sigma_{-i}).$$

[4]The statement of Claim 1 specifies that $i \notin \mathbb{C}_\star$ and $TV_i(\alpha_i^\star) \neq \pi_i^\star$. However, a strategy must be specified in all cases, and thus σ_i' must be specified also when $i \in \mathbb{C}_\star$ or $TV_i(\alpha_i^\star) = \pi_i^\star$.

That is, the collective utility of the collusive set $\mathbb{C}_i$ is the same in both play σ and play $\sigma'_i \sqcup \sigma_{-i}$.

(2) $\mathcal{M}$'s coin toss comes up Tails.

In this case, by hypothesis $TV_i(\alpha_i^\star) > \pi_i^\star$, while player i answers NO in play σ and answers YES in play $\sigma'_i \sqcup \sigma_{-i}$. Thus the individual utility of i is different in the two plays: specifically $u_i(\sigma) = 0$ and $u_i(\sigma'_i \sqcup \sigma_{-i}) = TV_i(\alpha_i^\star) - \pi_i^\star > 0$. Therefore, by this and Proposition (*), the MRA implies that

$$u_{\mathbb{C}_i}(\sigma) < u_{\mathbb{C}_i}(\sigma'_i \sqcup \sigma_{-i}).$$

Combining these two cases yields

$$\mathbb{E}[u_{\mathbb{C}_i}(\sigma)] < \mathbb{E}[u_{\mathbb{C}_i}(\sigma'_i \sqcup \sigma_{-i})].$$

That is, σ_i is dominated by σ'_i over Σ^0. ■

Claim 2. *Every independent player i chooses S_i to be his favorite subset of goods, that is*

$$S_i = \arg\max_{S \subseteq G} TV_i(S).$$

Proof of Claim 2. We again proceed by contradiction. Assume σ is such that, for some subset T of the goods, in Stage 1 some independent player i announces S_i such that $TV_i(S_i) < TV_i(T)$. Now consider the following alternative strategy for player i.

Strategy σ'_i:

- *Stage 1.* Run σ_i so as to compute the outcome Ω_i and the "desired" subset of goods S_i;
  Store S_i for future use; and
  Announce Ω_i and S'_i such that $S'_i = \arg\max_{S \subseteq G} TV_i(S)$.
- *Stage 2.* Run σ_i with the following four inputs:
  (1) stage input 2;
  (2) private inputs TV_i and K_i;
  (3) the announced outcome profile Ω; and
  (4) The profile $S_i \sqcup S_{-i}$ of subsets of the goods.
  Answer YES or NO as σ_i does.

We derive a contradiction by proving that σ_i is dominated by σ'_i. Towards proving this, observe that outcome profile Ω is the same in the plays σ and $\sigma'_i \sqcup \sigma_{-i}$. (In fact, Ω is defined at the end of Stage 1, and in that stage σ_i and σ'_i announce the same outcome Ω_i.) Accordingly, the star player too is the same in both plays, and thus so are $\mathbb{C}_\star$ and the second-highest revenue R'.

Also observe that $S_i \sqcup S_{-i}$ is exactly the profile S (that would have been) announced in play σ. Therefore, in Stage 2, σ'_i announces the same answer as σ_i (would have done). Accordingly, since σ_i is undominated over Σ^0, so is σ'_i. Therefore $\sigma'_i \in \Sigma^1_i$. By hypothesis, we also have $\sigma_j \in \Sigma^1_j$ for any player $j \neq i$; and thus:

$$\sigma'_i \sqcup \sigma_{-i} \in \Sigma^1.$$

Let us now prove that σ_i is dominated by σ'_i over Σ^1, contradicting the fact that σ_i is a surviving strategy. To this end, as i is independent, we only need to compare his expected *individual* utilities; that is $\mathbb{E}[u_i(\sigma)]$ and $\mathbb{E}[u_i(\sigma'_i \sqcup \sigma_{-i})]$. We distinguish two cases.

(1) $i \neq \star$ in both plays.

There are three sub-cases.

(a) $\alpha^\star_i = \emptyset$. In this case we have $\mathbb{E}[u_i(\sigma)] = \mathbb{E}[u_i(\sigma'_i \sqcup \sigma_{-i})] = 0$.

(b) $\alpha^\star_i \neq \emptyset$ *and i answers NO.* Also in this case we have $\mathbb{E}[u_i(\sigma)] = \mathbb{E}[u_i(\sigma'_i \sqcup \sigma_{-i})] = 0$.

(c) $\alpha^\star_i \neq \emptyset$ *and i answers YES.* In this case, with probability $\frac{1}{2}$, $u_i(\sigma) = u_i(\sigma'_i \sqcup \sigma_{-i}) = 0$, and with probability $\frac{1}{2}$, $u_i(\sigma) = u_i(\sigma'_i \sqcup \sigma_{-i}) = TV_i(\alpha^\star_i) - \pi^\star_i$.

Overall therefore, also in case (c) we have $\mathbb{E}[u_i(\sigma)] = \mathbb{E}[u_i(\sigma'_i \sqcup \sigma_{-i})]$.

(2) $i = \star$ in both plays.

In this case we note that $j \notin \mathbb{C}_\star$ for all $j \neq i$ (because i is independent), and that σ_j is a surviving strategy (by hypothesis). According to Claim 1, every player $j \neq i$ answers YES in Stage 2 if and only if $TV_j(\alpha^i_j) \geq \pi^i_j$. This fact, the fact that by construction $\Omega_i = (\alpha^i, \pi^i)$ is the same in both plays, and the fact that there is only one true-valuation TV_j, imply that the answer of every player j in Stage 2 is the same in both plays. Let us now compare the expected individual utility of i in the two plays, using the notation "$\sum_{j:YES}$" (respectively,"$\sum_{j:NO}$") for the sum taken over every player j who answers YES (respectively, NO) in Stage 2.

When $\mathcal{M}$'s coin flip ends up Tails, i's individual utility is the same in both plays. In fact, in this case we have

$$u_i(\sigma) = u_i(\sigma'_i \sqcup \sigma_{-i}) = \sum_{j:YES} \pi^i_j - \sum_{j:NO} \pi^i_j - R'.$$

But when $\mathcal{M}$'s coin flip ends up Heads, i's individual utility is less in play σ than in play $\sigma'_i \sqcup \sigma_{-i}$.

Indeed:

$$(a)\ u_i(\sigma) = TV_i(S_i) - \sum_{j:NO} \pi_j^i, \quad (b)\ u_i(\sigma_i' \sqcup \sigma_{-i}) = TV_i(S_i') - \sum_{j:NO} \pi_j^i, \quad \text{and} \quad (c)\ TV_i(S_i) < TV_i(S_i').$$

Therefore in Case 2 we have $\mathbb{E}[u_i(\sigma)] < \mathbb{E}[u_i(\sigma_i' \sqcup \sigma_{-i})]$.

Combining Cases 1 and 2, we see that σ_i is dominated by σ_i' over Σ^1. ■

Claim 3. *Every independent player i does not "underbid," that is,he announces Ω_i such that*

$$REV(\Omega_i) \geq MEW_i(K).$$

Proof of Claim 3. We again proceed by contradiction. Assume that σ is such that, for some independent player i, we have $REV(\Omega_i) < MEW_i(K)$. Now consider the following alternative strategy for i.

Strategy σ_i':

- *Stage 1.* Run σ_i so as to compute the outcome Ω_i and the "desired" subset of goods S_i;
  Store Ω_i for further use; and
  Announce the outcome $\Omega_i' = \arg\max_{\omega \in F(K_i)} REV(\omega)$ and the subset of goods S_i.
- *Stage 2.* If $i = \star$ or $\alpha_i^\star = \emptyset$, announce the empty string.
  Else, run σ_i with the following four inputs:
  (1) stage input 2;
  (2) private inputs TV_i and K_i;
  (3) $\Omega_i \sqcup \Omega_{-i}$; and
  (4) the profile S of subsets of the goods.
  Answer YES or NO as σ_i does.[5]

We derive a contradiction by proving that σ_i is dominated by σ_i'. Towards proving this, observe that outcome sub-profile Ω_{-i} is the same in the plays σ and $\sigma_i' \sqcup \sigma_{-i}$, so is subset profile S. Also observe that $\Omega_i \sqcup \Omega_{-i}$ is exactly the profile Ω (that would have been) announced in play σ. Therefore in Stage 2, when $i \neq \star$ and $\alpha_i^\star \neq \emptyset$, σ_i' announces the same answer as σ_i (would have done). Accordingly, since

[5] By construction, $REV(\Omega_i) < REV(\Omega_i')$. As will become clear soon, $REV(\Omega_j)$ is the same for any $j \neq i$ in both plays. Thus if $i \neq \star$ in play $\sigma_i' \sqcup \sigma_{-i}$, $i \neq \star$ in play σ. Therefore the output of σ_i with such inputs is not an empty string.

σ_i is undominated over Σ^1, so is σ'_i. Therefore $\sigma'_i \in \Sigma^2_i$. By hypothesis, we also have that $\sigma_j \in \Sigma^2_j$ for any player $j \neq i$; therefore:

$$\sigma'_i \sqcup \sigma_{-i} \in \Sigma^2.$$

Let's now prove that σ_i is dominated by σ'_i over Σ^2, contradicting the fact that σ_i is a surviving strategy. To this end, as i is independent, we only need to compare his expected *individual* utilities; that is $\mathbb{E}[u_i(\sigma)]$ and $\mathbb{E}[u_i(\sigma'_i \sqcup \sigma_{-i})]$. We distinguish three cases.

(1) *$i = \star$ in play σ.*

In this case, i's expected utility in play σ is the weighted sum of his utility when $\mathcal{M}$'s coin toss is Heads and his utility when $\mathcal{M}$'s coin toss is Tails.[6] Therefore we have

$$\begin{aligned}
\mathbb{E}[u_i(\sigma)] &= \frac{TV_i(S_i) - \sum_{j:NO} \pi^i_j}{2} + \frac{\sum_{j:YES} \pi^i_j - \sum_{j:NO} \pi^i_j - R'}{2} \\
&\leq \frac{TV_i(S_i)}{2} + \frac{\sum_j \pi^i_j - R'}{2} = \frac{TV_i(S_i)}{2} + \frac{REV(\Omega_i) - R'}{2} \\
&< \frac{TV_i(S_i) + MEW_i(K) - R'}{2}.
\end{aligned}$$

(The last, strict inequality follows from our hypothesis on the revenue of Ω_i.)

Let us now compare this expected utility with $\mathbb{E}[u_i(\sigma'_i \sqcup \sigma_{-i})]$. Towards computing the latter utility in Case 1, notice that the present case implies that $i = \star$ also in play $\sigma'_i \sqcup \sigma_{-i}$. In fact, we have already argued that the subprofile Ω_{-i} is the same in both plays and that $REV(\Omega'_i) > REV(\Omega_i)$. This fact also implies that the second-highest revenue R' is the same in both plays.

Because Ω'_i is a "feasible way of selling the goods to the players in $-i$," according to Claim 1, every player $j \neq i$ answers YES in Stage 2: in our notation $\sum_{j:YES} = \sum_j$. Accordingly, we have

$$\begin{aligned}
\mathbb{E}[u_i(\sigma'_i \sqcup \sigma_{-i})] &= \frac{TV_i(S_i)}{2} + \frac{\sum_j \pi^i_j - R'}{2} = \frac{TV_i(S_i) + REV(\Omega'_i) - R'}{2} \\
&= \frac{TV_i(S) + MEW_i(K) - R'}{2}.
\end{aligned}$$

Therefore we conclude that, in Case 1, $\mathbb{E}[u_i(\sigma)] < \mathbb{E}[u_i(\sigma'_i \sqcup \sigma_{-i})]$.

(2) *$i \neq \star$ in play σ and in play $\sigma'_i \sqcup \sigma_{-i}$.*

In this case, again because all outcomes in Ω_{-i} are the same in both plays, so are the star player, $\Omega_\star$, and $\mathbb{C}_\star$. Moreover, because i is independent, we also have $i \notin \mathbb{C}_\star$ in both plays. Finally, by construction, the answer announced in Stage 2 by σ'_i and σ_i are the same in both plays.

[6] Both individual utilities are expected, if the strategies of the other players are probabilistic.

Accordingly, to compare i's individual utility in play σ with his individual utility in play $\sigma'_i \sqcup \sigma_{-i}$ it suffices to consider three sub-cases.

(a) $\alpha_i^\star = \emptyset$. In this sub-case, $\mathbb{E}[u_i(\sigma)] = \mathbb{E}[u_i(\sigma'_i \sqcup \sigma_{-i})] = 0$.

(b) $\alpha_i^\star \neq \emptyset$ *and* i *answers NO*. In this sub-case, also $\mathbb{E}[u_i(\sigma)] = \mathbb{E}[u_i(\sigma'_i \sqcup \sigma_{-i})] = 0$.

(c) $\alpha_i^\star \neq \emptyset$ *and* i *answers YES*. In this sub-case, $\mathbb{E}[u_i(\sigma)] = \mathbb{E}[u_i(\sigma'_i \sqcup \sigma_{-i})] = \frac{TV_i(\alpha_i^\star)-\pi_i^\star}{2}$.

Therefore in Case 2 we have $\mathbb{E}[u_i(\sigma)] = \mathbb{E}[u_i(\sigma'_i \sqcup \sigma_{-i})]$.

(3) $i \neq \star$ *in play* σ *and* $i = \star$ *in play* $\sigma'_i \sqcup \sigma_{-i}$.

In this case, let us prove that $\mathbb{E}[u_i(\sigma)] \leq \frac{TV_i(S_i)}{2} \leq \mathbb{E}[u_i(\sigma'_i \sqcup \sigma_{-i})]$.

To upperbound $\mathbb{E}[u_i(\sigma)]$ we consider three sub-cases for play σ.

(a) $\alpha_i^\star = \emptyset$. In this case, $\mathbb{E}[u_i(\sigma)] = 0$.

(b) $\alpha_i^\star \neq \emptyset$ *and* $TV_i(\alpha_i^\star) < \pi_i^\star$. In this case, according to Claim 1, player i answers NO in Stage 2, and thus $\mathbb{E}[u_i(\sigma)] = 0$.

(c) $\alpha_i^\star \neq \emptyset$ *and* $TV_i(\alpha_i^\star) \geq \pi_i^\star$. In this case, according to Claim 1, player i answers YES in Stage 2, and thus $\mathbb{E}[u_i(\sigma)] = \frac{TV_i(\alpha_i^\star)-\pi_i^\star}{2} \leq \frac{TV_i(\alpha_i^\star)}{2} \leq \frac{TV_i(S_i)}{2}$. In fact, by Claim 2, $S_i = \arg\max_{S \subseteq G} TV_i(S)$.

Therefore, in Case 3, $\mathbb{E}[u_i(\sigma)] \leq \frac{TV_i(S_i)}{2}$ as stated above.

Let us now lowerbound $\mathbb{E}[u_i(\sigma'_i \sqcup \sigma_{-i})]$ in Case 3. First of all, as in Case 1, we have that $\mathbb{E}[u_i(\sigma'_i \sqcup \sigma_{-i})] = \frac{TV_i(S_i)+REV(\Omega'_i)-R'}{2}$. But now, since $REV(\Omega'_i) \geq R'$, we also have $\frac{TV_i(S_i)}{2} \leq \frac{TV_i(S_i)+REV(\Omega'_i)-R'}{2}$, as stated above.

Combining these three cases, we see that σ_i is dominated by σ'_i over Σ^2. ■

We are finally ready to prove Theorem 1, that is: for our mechanism $\mathcal{M}$,

$$\mathbb{E}[REV] + \mathbb{E}[SW] \geq \frac{MEW_I(K)}{2}.$$

To this end, denote by $*$ the independent player "realizing" our benchmark: that is,

$$* = \arg\max_{i \in I} MEW_i(K).$$

(Notice that the players $*$ and $\star$ need not to coincide.)

According to Claim 3, in any surviving play, $*$ announces an outcome Ω_* such that $REV(\Omega_*) \geq MEW_*(K)$. Now, since by definition the star player is the one who announces an outcome with the largest revenue, we have $R_\star \geq REV(\Omega_*)$, and thus $R_\star \geq MEW_I(K) = MEW_*(K)$. To prove Theorem 1 we distinguish two cases.

(1) $\star = *$.

In this case, as player $*$ is independent, so is player $\star$, and thus $i \notin \mathbb{C}_\star$ for all players $i \neq \star$. Therefore Claim 1 guarantees that every $i \neq \star$ answers YES in Stage 2 if and only if $TV_i(\alpha_i^\star) \geq \pi_i^\star$. Accordingly, the following inequality holds for $\mathcal{M}$'s expected social welfare:

$$\mathbb{E}[SW] = \frac{TV_\star(S_\star)}{2} + \frac{\sum_{i:YES} TV_i(\alpha_i^\star)}{2} \geq \frac{\sum_{i:YES} TV_i(\alpha_i^\star)}{2} \geq \frac{\sum_{i:YES} \pi_i^\star}{2}.$$

At the same time,

$$\mathbb{E}[REV] = \frac{\sum_{i:NO} \pi_i^\star}{2} + \frac{R' + \sum_{i:NO} \pi_i^\star}{2} \geq \frac{\sum_{i:NO} \pi_i^\star}{2}.$$

Thus

$$\mathbb{E}[SW] + \mathbb{E}[REV] \geq \frac{\sum_{i:YES} \pi_i^\star + \sum_{i:NO} \pi_i^\star}{2} = \frac{R_\star}{2} \geq \frac{MEW_I(K)}{2}.$$

(2) $\star \neq *$.

In this case, $* \in -\star$, thus —since player $*$ is independent— $R' \geq REV(\Omega_*) \geq MEW_I(K)$. Therefore $\mathcal{M}$'s expected revenue is

$$\mathbb{E}[REV] = \frac{\sum_{i:NO} \pi_i^\star}{2} + \frac{R' + \sum_{i:NO} \pi_i^\star}{2} \geq \frac{R'}{2} \geq \frac{MEW_I(K)}{2}.$$

Because

$$\mathbb{E}[SW] = \frac{TV_\star(S_\star)}{2} + \frac{\sum_{i:YES} TV_i(\alpha_i^\star)}{2} \geq 0,$$

we have

$$\mathbb{E}[SW] + \mathbb{E}[REV] \geq \frac{MEW_I(K)}{2}.$$

Combining these two cases, Theorem 1 follows.

Q.E.D.

References

[MV07.a] S. Micali and P. Valiant, Revenue in Truly Combinatorial Auctions and Adversarial Mechanism Design, Technical Report, MIT-CSAIL- TR-2008-039, 2008.

(Previous versions include: MIT-CSAIL-TR-2007-052, 2007; submission to FOCS 2007, submission to SODA 2008; and the versions deposited with the Library of Congress in 2007)

[MV07.b] S. Micali and P. Valiant, Leveraging Collusion in Combinatorial Auctions, Unpublished Manuscript. 2007

[OR94] M. J. Osborne and A. Rubinstein, A Course in Game Theory, MIT Press, 1994.

[BLP06] M. Babaioff, R. Lavi and E. Pavlov, Single-Value Combinatorial Auctions and Implementation in Undominated Strategies, SODA'06, pages 1054-1063, 2006.