

A Note on Testing Truthfulness

Ahuva Mu'alem*

Abstract

This work initiates the study of algorithms for the testing of monotonicity of mechanisms. Such testing algorithms are useful for searching dominant strategy mechanisms. An ϵ -tester for monotonicity is given a query access to a mechanism, accepts if monotonicity is satisfied, and rejects with high probability if more than ϵ -fraction of the mechanism values must be modified to obtain the property.

The notion of the distance from monotonicity essentially suggests a notion of distance from truthfulness. A direct mechanism is $(1 - \epsilon)$ -truthful if reporting the true valuation is a dominant strategy for every player i with probability $1 - \epsilon$ (assuming v_{-i} are uniformly distributed). This raises the question of how a local measure of violation, representing the point of view of the individual player, relates to the global measure of violation, representing the point of view of the mechanism designer.

1 Introduction

In this work we study testing algorithms for mechanisms. A motivating scenario is as follows. Suppose the designer of some mechanism argues that a certain behavior is optimal for the agents. One such recommendation might be “truthfulness”: in this case the designer would like to promote the true reporting of the individual preference as a dominant strategy. The agents then would like to verify this best response recommendation. In computational scenarios, it might be preferable to exhibit a quick test that examines a small “sampled” portion of the mechanism, allowing reasonable errors.

Many common collective decision making procedures can be modeled as mechanisms: various elections and resource allocation procedures, several markets, auctions and routing protocols in the Internet. A typical mechanism aggregates the individual preferences (“valuations”), then chooses a socially desired outcome (“alternative”), and possibly assigns payments for the agents. More formally, a (direct) mechanism is a tuple $m = (f, p)$, where the social choice function $f : V \rightarrow A$ maps an n -tuple of valuations $v = (v_1, v_2, \dots, v_n) \in V$ ($\subseteq R^{n|A|}$) to an outcome $a \in A$, and the function $p : V \rightarrow R^n$ assigns payments. A mechanism is *truthful* if true reporting of the valuation is a dominant strategy for every player. Formally, if $v_i(f(v_i, v_{-i})) - p_i(v_i, v_{-i}) \geq v_i(f(v'_i, v_{-i})) - p_i(v'_i, v_{-i})$, for all v_i, v'_i and v_{-i} .

It turns out that testing truthfulness is closely related to testing Boolean monotonicity. Our work is inspired by the work of Goldreich, Goldwasser et al. [4], who considered the problem of testing whether a given function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ is monotone. They presented a tester whose query complexity and running time are linear in n and $1/\epsilon$ (note that the size of the input is 2^n).

*Email: ahumu@yahoo.com. CS Department, Bar-Ilan University, Ramat-Gan, Israel. This work was carried out in School of Engineering and Computer Science, The Hebrew University of Jerusalem. Supported by grants from the Israeli Academy of Sciences, Israeli Ministry of Sciences, and the USA-Israel Binational Science Foundation.

Their tester performs a simple local test: It verifies whether (Boolean) monotonicity is maintained for randomly chosen pairs of strings that differ exactly on a single bit. Their analysis relates the measure of local violation of monotonicity to the global measure of the minimum distance of the given function to any monotone function. They also considered extensions of the form $f : \Sigma^n \rightarrow \Sigma'$, where Σ is a poset and Σ' is a total order.

It is easy to verify that monotonicity of Boolean functions $f : \{0,1\}^n \rightarrow \{0,1\}$ is a special case of monotonicity for social choice functions $f : V \rightarrow A$. Note that in our setting the general case is different from the binary case. First, each agent evaluates every possible outcome $a \in A$ (as opposed to choosing between “0” or “1”). Moreover, the set A of the mechanism’s possible outcomes is not necessarily ordered.

It is already known that the truthfulness of $m = (f, p)$ implies the monotonicity of f [12, 8, 7, 3]¹. However, monotonicity is not strong enough to imply truthfulness (see e.g., [7]). The *extended* monotonicity condition was (constructively) shown to fully characterize truthfulness [14, 9].

Our contribution. Monotonicity and extended monotonicity testing reflect the point of view of the mechanism designer. Intuitively, if all the payments are known, then the agents would like to test the “truthfulness” directly. That is, player i would like to estimate the probability of a local violation ($\exists v'_i$ such that $v_i(f(v_i, v_{-i})) - p_i(v_i, v_{-i}) < v_i(f(v'_i, v_{-i})) - p_i(v'_i, v_{-i})$). The following theorem relates both points of view.

Thm1 Let $V = \{1, 2, \dots, q\}^{n \cdot |A|}$. If $f : V \rightarrow A$ is ϵ -close to extended monotonicity then there exists an associate generic payment function $p : V \rightarrow \mathbb{R}^n$ such that $m = (f, p)$ is $(1 - 2\epsilon)$ -truthful.

From a computational point of view this theorem suggests the following straight-forward implication: In the search for (almost) dominant strategy mechanisms the designer should only look for social choice functions that are (close to) extended monotonicity. The agents then will test the extended monotonicity. In particular, this test queries the social choice function alone and not the payment function: this means that the mechanism designer does not need to compute any part of the entire “table” of payments in advance for every possible input v_{-i} . Note that intuitively, computing the above generic payment $p_i(v)$ takes $|V_i|^{\mathcal{O}(c)}$ (essentially a shortest path computation), so that querying the payment function might be costly. A trusted authority can compute the specific payments after the agents report their bids.

The following shows that beyond extended monotone functions, the class of almost truthful mechanisms contains several approximations to the social welfare.

Thm2 Let $V = \{1, 2, \dots, q\}^{n \cdot |A|}$. If $f : V \rightarrow A$ is $(1 + \epsilon)$ -approximation to the social welfare then it is $(\epsilon 2qn)$ -close to monotone and $(\epsilon \cdot |A| \cdot qn)$ -close to extended monotonicity.

We then observe that Thm1 and the proof of Thm2 suggest a generic way to derive (almost) truthful mechanisms, using a shifting technique.²

We also show a monotonicity tester for a special case. It is worth noting that the tester accepts not only monotone functions with high probability but also almost monotone functions.

Open. As the above discussion suggests, describing a monotonicity testing algorithm for the general

¹The monotonicity property is called W-MON in [7].

²Such a technique was used in [2] to reduce the influence of Boolean variables. In our case the influence of an individual player may not be decreased (by a single shift operation), but rather the total social welfare is increased. Thus, intuitively, every monotone Boolean function corresponds to a local minimum of the total influence, and every monotone social choice function corresponds to a local maximum of the total social welfare.

case $f : \{1, 2, \dots, q\}^{n \cdot |A|} \rightarrow A$ is an interesting open problem. Besides the general case (where A is not ordered “at all”), the “weakly” ordered case (assuming some “universally bad” outcome $z_0 : v_i(a) \geq v_i(z_0)$ for all a and i) seems interesting as well (e.g. for the testing of Combinatorial Auctions).

1.1 Related Work

Several Algorithmic Mechanism Design [10] papers consider various notions of almost truthfulness (see [1, 6] and references therein). The paper [1] shows an almost truthful mechanism for Combinatorial Auctions with single parameter agents. [6] shows that every approximation algorithm w.r.t. the social welfare, can be coupled with a payment function that ensures “almost” truthfulness (using a different notion of almost truthfulness from ours).

Testing algorithms for the convexity and submodularity of valuations are studied in [11]. Such testers are applicable for mechanisms which guarantee certain properties (such as truthfulness or an approximation ratio) only if the input follows some specific structure.

2 Preliminaries

- A social choice function $f : V \rightarrow A$ is ϵ -close to a property P if there exists a function $g : V \rightarrow A$ satisfying P such that $\frac{|\{v | f(v) \neq g(v)\}|}{|V|} \leq \epsilon$, where V is finite. The function is ϵ -far, otherwise.
- A social choice function is monotone if for every i, v_{-i}, v_i, v'_i :

$$v_i(f(v_i, v_{-i})) - v_i(f(v'_i, v_{-i})) + v'_i(f(v'_i, v_{-i})) - v'_i(f(v_i, v_{-i})) \geq 0.$$

- A social choice function is extended monotone if for every $i, v_{-i}, k, v_i^1, v_i^2, \dots, v_i^k$:

$$v_i^1(f(v_i^1, v_{-i})) - v_i^1(f(v_i^2, v_{-i})) + \dots + v_i^k(f(v_i^k, v_{-i})) - v_i^k(f(v_i^1, v_{-i})) \geq 0.$$

This is also called the non-negative cycle property. As mentioned, extended monotonicity implies monotonicity but not necessarily vice versa. Clearly, if f is ϵ -close to extended monotonicity, then it is ϵ -close to monotonicity.

- A social choice function $f : V \rightarrow A$ is a $(1 + \epsilon)$ -approximation with respect to the social welfare if: $\frac{\sum_i v_i(f^*(v))}{1 + \epsilon} \leq \sum_i v_i(f(v))$, where the function $f^* : V \rightarrow A$ always outputs an alternative with the highest social welfare, that is $f^*(v) = \operatorname{argmax}_{a \in A} \sum_i v_i(a)$.
- A tester algorithm for property P is a randomized algorithm that for any given input function f and a distance parameter ϵ :
 - The tester accepts with probability $\geq \frac{2}{3}$ if f has the property.
 - The tester rejects with probability $\leq \frac{1}{3}$ if f is ϵ -far from the property.

3 Almost Truthfulness

3.1 Proof of Theorem 1

The following theorem shows that every almost extended monotone function can be coupled with an “associate generic” payment function that ensures almost truthfulness.

Theorem 1 *Let $V = \{1, \dots, q\}^{n \cdot |A|}$. If $f : V \rightarrow A$ is ϵ -close to extended monotonicity then there is $p : V \rightarrow R^n$ such that the mechanism $m = (f, p)$ is $(1 - 2\epsilon)$ -truthful.*

proof: Using a bound from [4] we shall see that the payment in [13, 14] fits our setting. We start with some definitions. For every distinct $b, c \in A$ and v_{-i} define:

$$\delta_{bc}^i(v_{-i}) = \inf \{v'_i(b) - v'_i(c) \mid v'_i \in V_i \text{ and } f(v'_i, v_{-i}) = b\}, \text{ and}$$

$$\tau_{bc}^i(v_{-i}) = \inf_{a^i \in A, k \in N} \delta_{ba^i}^i(v_{-i}) + \delta_{a^i a^2}^i(v_{-i}) + \dots + \delta_{a^{k-1} a^k}^i(v_{-i}) + \delta_{a^k c}^i(v_{-i}).$$

Clearly $\tau_{bc}^i(v_{-i}) \leq \delta_{bc}^i(v_{-i})$. Note that in presence of a negative cycle (with respect to v_{-i}) $\tau_{bc}^i(v_{-i}) = -\infty$. Now, fix some arbitrary alternative $a \in A$, and define the *associate generic payment* function:

$$p_d^i(v_{-i}) = \begin{cases} \tau_{da}^i(v_{-i}) & d \neq a \text{ and } \tau_{da}^i(v_{-i}) \neq -\infty \\ 0 & \text{otherwise.} \end{cases}$$

Consider the graph G with the vertices $V(G) = V$, and the edges:

$$E(G) = \{(v, u) \mid \text{there exist } v_i, u_i \in V_i \text{ such that } v = (v_i, v_{-i}) \text{ and } u = (u_i, v_{-i})\}.$$

Clearly, $|E(G)| = \frac{1}{2} \cdot |V| \cdot n \cdot (q^{|A|} - 1)$. Let $\Delta \subseteq E(G)$ be a minimal set of violating edges that must be omitted from G in order to avoid negative cycles. Let $\delta = \frac{|\Delta|}{|E(G)|}$. The proof follows from the following two lemmas.

Lemma 1 $\epsilon \geq \frac{\delta}{2}$.

proof: The argument in [4] is applicable for this generalized case: To obtain the property, $\epsilon|V|$ changes must be made. Each change can fix at most $n \cdot (q^{|A|} - 1)$ violating edges, and so:

$$\epsilon|V| \geq \frac{|\Delta|}{n \cdot (q^{|A|} - 1)} = \frac{\delta \cdot |E(G)|}{n \cdot (q^{|A|} - 1)}. \blacksquare$$

To simplify the notation, we next assume a single player.

Lemma 2 *If $f(v) = b$, then $v(b) - p_b \geq v(c) - p_c$ with probability at least $1 - \delta$, for every $c \in A$.*

proof: There are 3 cases to consider:

1. If $c = a$, then $p_b - p_a$ equals τ_{ba} with probability $1 - \delta$, and $\tau_{ba} \leq \delta_{ba} \leq v(b) - v(a)$, by definition.
2. If $b = a$, then with probability $1 - \delta$, $\tau_{ca} + \delta_{ac} \geq 0$, and so w.h.p $p_a - p_c \leq -\tau_{ca} \leq \delta_{ac} \leq v(a) - v(c)$.
3. If $a \neq b, c$, again w.h.p $\tau_{ba} \leq \tau_{ca} + \delta_{bc}$, and so w.h.p $p_b - p_c = \tau_{ba} - \tau_{ca} \leq \delta_{bc} \leq v(b) - v(c)$. \blacksquare

3.2 Proof of Theorem 2

In this subsection we show that the class of almost extended monotone social choice functions contains the class of approximations to the social welfare (for small enough ϵ).

Theorem 2 *Let $V = \{1, \dots, q\}^{n \cdot |A|}$. If $f : V \rightarrow A$ is $(1 + \epsilon)$ -approximation to the social welfare then it is $(\epsilon 2qn)$ -close to monotone and $(\epsilon \cdot |A| \cdot qn)$ -close to extended monotonicity.*

We start with the definition of the extended shift operator. Choose arbitrary collection of inputs that form a negative cycle: $C = \{(v_i^{(1)}, v_{-i}), (v_i^{(2)}, v_{-i}), \dots, (v_i^{(k)}, v_{-i})\}$. More formally, there exist $a^{(1)}, \dots, a^{(k)}$ such that: $f(v_i^{(j)}, v_{-i}) = a_j$ and $v_i^{(1)}(a_1) - v_i^{(1)}(a_2) + v_i^{(2)}(a_2) - v_i^{(2)}(a_3) + \dots + v_i^{(k)}(a_k) - v_i^{(k)}(a_1) < 0$.

The “shifted” f would be $\tilde{f}(v_i^{(j)}, v_{-i}) = a_{j+1}$, where $k+1$ equals 1, and $\tilde{f}(v') = f(v')$, otherwise. That is, \tilde{f} repairs the given negative cycle, and is similar to f elsewhere.

proof: Let $w_f(v) = \sum_i v_i(f(v))$, and $w_{f^*}(v) = \sum_i v_i(f^*(v))$, be the resulted social welfare of f (and the optimal social welfare, respectively) for the input v . Let $W_f = \sum_{v \in V} w_f(v)$, and $W_* = \sum_{v \in V} w_{f^*}(v)$ be the resulted total welfare over all possible inputs. The $(1 + \epsilon)$ -approximability of f implies that $\frac{W_*}{1 + \epsilon} \leq W_f$, so that:

$$W_* - W_f \leq W_* - \frac{W_*}{1 + \epsilon} \leq \epsilon W_* \leq \epsilon \cdot nq \cdot |V|. \quad (*)$$

It is easy to verify that $W_{\tilde{f}} \leq W_*$, and that $W_f + 1 \leq W_{\tilde{f}}$. To see this, $W_f = \sum_{v \in V \setminus C} w_f(v) + \sum_{v \in C} (v_i(f(v)) + \sum_{j \neq i} v_j(f(v))) < \sum_{v \in V \setminus C} w_f(v) + \sum_{v \in C} (v_i(\tilde{f}(v)) + \sum_{j \neq i} v_j(f(v))) = W_{\tilde{f}}$.

By (*) we get that the given approximation function might have at most $\epsilon nq|V|$ negative cycles, as each shift iteration increases the total welfare by at least 1. In each shift iteration at most $|A|$ entries are being updated. This shows $(\epsilon \cdot |A| \cdot qn)$ -closeness to extended-monotonicity. Similarly, to show the closeness to monotonicity, the operator can be restricted to shift only negative cycles of length 2. ■

3.3 Generic Design

For finite (not necessarily single-parameter) domains the above shifting technique suggests a generic way to construct (almost) truthful mechanisms. Note that, as demonstrated in [5], for bounded domains (as opposed to unrestricted domains [14]) there are truthful mechanisms other than the celebrated VCG and weighted VCG mechanisms, so the resulted mechanisms after the shifting may be different.

Generic Design Technique: Start with an arbitrary social choice function f . Use the extended shift operator to make it (close to) extended monotone. Construct a (almost) truthful mechanism by combining the modified social choice function with the associate generic payment function.

A single shift operation repairs one given cycle of violation. It might be the case that some other new violated cycles will pop-up, as a result. However, the proof of theorem 2 shows that a “long enough” sequence of shifting operations leads eventually to extended monotone function.

4 Monotonicity Testers

In this section we present monotonicity tester for a special case. Let v_a denotes the following zero-one valuation: $v_a(a) = 1$, and $v_a(b) = 0$ for every $b \neq a$. The domain of these valuations is denoted $V_A = \{v_a \mid a \in A\}$. We show a tester for $f : V_A \rightarrow A$ and $f : V_A \times V_A \rightarrow A$. The following algorithm and its analysis is inspired by [4].

Algorithm 1 Repeat $O(\frac{|A|}{\epsilon})$ times:

- Uniformly select $v \in V_A \times \dots \times V_A$, $i \in 1, \dots, n$, and $v'_i \in V_A$. Query $f(v_i, v_{-i})$ and $f(v'_i, v_{-i})$.
(Assume w.l.o.g that $v_i = v_a$, $v'_i = v_b$)
- Reject if a violation is detected.
(That is, reject if either $f(v_a, v_{-i}) = b \neq f(v_b, v_{-i})$ or $f(v_b, v_{-i}) = a \neq f(v_a, v_{-i})$).

If all iterations are completed without a reject then accept.

Proposition 1 Algorithm 1 is a monotonicity tester for $f : V_A \rightarrow A$ and $f : V_A \times V_A \rightarrow A$.

proof: Clearly monotone functions are always accepted. We start with $f : V_A \rightarrow A$. Let $\Delta_f = \{(v_a, v_b) \mid f(v_a) = b \neq f(v_b) \text{ or } f(v_b) = a \neq f(v_a)\}$, be the set of all violating pairs. Let δ be the fraction of violating pairs, that is: $\delta = \frac{|\Delta_f|}{0.5 \cdot |A| \cdot |A-1|}$. It is easy to check that a similar shift operator as in [4] cannot decrease the number of non-violating pairs. Let $a, b \in A$ arbitrary elements, then the shift of f (w.r.t v_a and v_b) is as follows: if v_a, v_b form a violating pair then $\tilde{f}(v_a) = f(v_b)$, $\tilde{f}(v_b) = f(v_a)$ and $\tilde{f}(v_c) = f(v_c)$, $c \neq a, b$, otherwise $f = \tilde{f}$.

Claim 1 $|\Delta_{\tilde{f}}| \leq |\Delta_f|$.

proof: Define $\Delta_f(v_a) = \{(v_x, v_a) \mid (v_x, v_a) \text{ is a violating pair}\}$. If $f(v_a) = b$ and $f(v_b) = a$, then after the shift w.r.t v_a, v_b , $\Delta_{\tilde{f}}(v_a) = \Delta_{\tilde{f}}(v_b) = \emptyset$. Otherwise, assume w.l.o.g that $f(v_a) \neq a, b$ and $f(v_b) = a$. Again $\Delta_{\tilde{f}}(v_a) = \emptyset$. For every x such that v_b, v_x is a non-violating pair: If $f(v_x) = x$, it is easy to see that v_b, v_x is a non violating pair after the shift. If $f(v_x) \neq x$, it is easy to see that v_b, v_x is a violating pair only if $f(v_a) = x$. Clearly, this pair cancels out with v_a, v_b , and thus $|\Delta_{\tilde{f}}| \leq |\Delta_f|$. ■

By the above claim it is immediate that there is a “long enough” sequence of shifts which result in a monotone social choice function (say $g : V_A \rightarrow A$, as it might depend on the specific order of shifts). Note that this sequence is finite, as each shift operation increases the total social welfare W_f . Finally, $\epsilon|A| \leq \text{dist}(f, g) \leq 2 \cdot |\Delta_f| \leq \delta \cdot |A|^2$. This shows that if f is ϵ -far, then the test accepts with probability $(1 - \delta)^{O(\frac{|A|}{\epsilon})} \leq \text{const}$.

For $f : V_A \times V_A \rightarrow A$, first try the perform a similar shift operation (say over (v_a, v_2) , (v_b, v_2)). If the number of violating pairs increases as a result, by the above argument the number of violating pairs over the rows of $V = V_A \times V_A$ must be increased. There are several simple cases to consider (for $f(v_a, v_2) = b$, $f(v_b, v_2) = a : v_2 \in \{v_a, v_b, v_c\}$ and for $f((v_a, v_2) = x$, $f(v_b, v_2) = a : v_2 \in \{v_a, v_x, v_b, v_c\}$). In these cases we use some other natural shifts depending on the specific case.

- In the case $v_2 = v_a$ we use the following shift $f(v_a, v_2) = a$. In the case $v_2 = v_x$, $x \neq b$ we use the following shift $f(v_b, v_2) = x$. In the case $v_2 = v_b$ we use the following shift $f(v_b, v_2) = b$.

- In case $v_2 = v_c, c \neq a, b, x$. If there is a violating pair $(v_a, v_a), (v_b, v_a)$ repair first this shift as in the former case. Otherwise, $f(v_a, v_2) = a$.

Finally, $\epsilon|A|^2 \leq \text{dist}(f, g) \leq 2 \cdot |\Delta_f| \leq \delta \cdot |A|^2 \cdot 2 \cdot |A|$. ■

The following claim shows that the above algorithm accepts “almost monotonicity” w.h.p.

Claim 2 *Algorithm 1 accepts w.h.p social choice functions that are ϵ -close to monotone.*

proof: By Lemma 1, $2\epsilon \geq \delta$, and so the probability that the test rejects an ϵ -close function is $\delta \frac{|A|}{\epsilon}$. ■

Acknowledgements

We thank Michal Feldman, Daniel Lehmann, Noam Nisan, Amir Ronen, Alex Samorodnitsky and Michael Schapira for helpful discussions and comments.

References

- [1] Aaron Archer, Christos Papadimitriou, Kunal Talwar, and Eva Tardos. An approximate truthful mechanism for combinatorial auctions with single parameter agent. In *SODA*, 2003.
- [2] Michael Ben-Or and Nathan Linial. Collective coin flipping. In *Randomness and Computation*, pages 91–115. Academic Press, New York, 1990.
- [3] Sushil Bikhchandani, Shurojit Chatterji, and Arunava Sen. Incentive compatibility in multi-unit auctions, 2003. Working paper.
- [4] Oded Goldreich, Shafi Goldwasser, Eric Lehman, Dana Ron, and Alex Samorodnitsky. Testing monotonicity. *Combinatorica*, 20(3):301–337, 2000.
- [5] Philippe Jehiel, Moritz Meyer ter Vehn, and Benny Moldovanu. Potential and implementation, 2004. Working paper.
- [6] Anshul Kothari, David Parkes, and Subhash Suri. Approximately-strategyproof and tractable multi-unit auctions. In *Proc. of the ACM Conference on Electronic Commerce (EC'03)*, 2003.
- [7] Ron Lavi, Ahuva Mu'alem, and Noam Nisan. Towards a characterization of truthful combinatorial auctions, 2003. Working paper. Preliminary version presented in FOCS-03.
- [8] Daniel Lehmann, Liadan O'Callaghan, and Yoav Shoham. Truth revelation in approximately efficient combinatorial auctions. *Journal of the ACM*, 49(5):577–602, 2002.
- [9] Rudolf Muller and Rakesh Vohra. On dominant strategy mechanisms, 2003. Working paper.
- [10] Noam Nisan and Amir Ronen. Algorithmic mechanism design. *Games and Economic Behavior*, 35:166–196, 2001.
- [11] Michal Parnas, Dana Ron, and Ronitt Rubinfeld. On testing convexity and submodularity. *SIAM Journal on Computing*, 32(5):1158–1184, 2003.

- [12] Kevin Roberts. The characterization of implementable choice rules. In Jean-Jacques Laffont, editor, *Aggregation and Revelation of Preferences. Papers presented at the 1st European Summer Workshop of the Econometric Society*, pages 321–349. North-Holland, 1979.
- [13] J.C. Rochet. A necessary and sufficient condition for rationalizability in a quasi-linear context. *Journal of Mathematical Economics*, 16:191–200, 1987.
- [14] Irit Rozenshtrom. Dominant strategy implementation with quasi-linear preferences, 1999. Master's thesis, Dept. of Economics, The Hebrew University, Jerusalem, Israel.