Proving the Incompatibility of Efficiency and Strategyproofness via SMT Solving

by Brandl, Brandt, Eberl, and Geist (2018)

Philemon Huising

Advanced Topics in Computational Social Choice
Institute for Logic, Language and Computation
University of Amsterdam

11 November, 2021

The Model

- A, a finite set of m alternatives.
- $N = \{1, \ldots, n\}$, a finite set of agents.
- The preference relation reported by agent i is a complete and transitive relation on A, and is denoted $\succsim_i$.
- The set of all possible preference relations is denoted $\mathcal{R}(A)$.
- A preference profile is a tuple, $R = (\succsim_1, \ldots, \succsim_n)$, that specifies a preference relation for each agent $i \in N$.
- The set of all preference profiles is then $\mathcal{R}(A)^n$

Social Decision Schemes

A social decision scheme (SDS) maps preference profiles to lotteries.

Why? Fairness, e.g., in light of the GS-theorem.

Social Decision Schemes

A social decision scheme (SDS) maps preference profiles to lotteries.

Why? Fairness, e.g., in light of the GS-theorem.

The model continued:

- A lottery over A is simply a probability distribution on A, i.e., $p : A \to [0, 1]$, where $\sum_{a \in A} p(a) = 1$.
- The collection of all lotteries over A is denoted $\Delta(A) = \{p \in \mathbb{R}^A_{\geq 0} \mid \sum_{a \in A} p(a) = 1\}$.
- An SDS is defined as a function

$$F : \mathcal{R}(A)^n \to \Delta(A)$$

Axioms: Anonymity and Neutrality

The same as before (kind of):

- F is anonymous if $F(\succsim_1, \ldots, \succsim_n) = F(\succsim_{\sigma(1)}, \ldots, \succsim_{\sigma(n)})$ for any profile $(\succsim_1, \ldots, \succsim_n)$ and permutation $\sigma : N \to N$.

Axioms: Anonymity and Neutrality

The same as before (kind of):

- F is anonymous if $F(\succsim_1, \ldots, \succsim_n) = F(\succsim_{\sigma(1)}, \ldots, \succsim_{\sigma(n)})$ for any profile $(\succsim_1, \ldots, \succsim_n)$ and permutation $\sigma : N \to N$.
- F is neutral if $F(R)(a) = F(\pi(R))(\pi(a))$ for any profile R, alternative $a \in A$ and permutation $\pi : A \to A$.

Axioms: Anonymity and Neutrality

The same as before (kind of):

- F is anonymous if $F(\succsim_1, \ldots, \succsim_n) = F(\succsim_{\sigma(1)}, \ldots, \succsim_{\sigma(n)})$ for any profile $(\succsim_1, \ldots, \succsim_n)$ and permutation $\sigma : N \to N$.
- F is neutral if $F(R)(a) = F(\pi(R))(\pi(a))$ for any profile R, alternative $a \in A$ and permutation $\pi : A \to A$.

But what about efficiency and strategyproofness?

Utility Representations

We need a way to reason about the preferences that agents have over lotteries.

Utility Representations

We need a way to reason about the preferences that agents have over lotteries.

- For each preference profile R and agent i, we have a utility function $u_i^R : A \to \mathbb{R}$.

Utility Representations

We need a way to reason about the preferences that agents have over lotteries.

- For each preference profile R and agent i, we have a utility function $u_i^R : A \to \mathbb{R}$.
    - A utility function for an agent and profile must be consistent with the ordinal preferences of that agent: for $a, b \in A$,

$$u_i^R(a) \geq u_i^R(b) \text{ iff } a \succsim_i b$$

.

Utility Representations

We need a way to reason about the preferences that agents have over lotteries.

- For each preference profile R and agent i, we have a utility function $u_i^R : A \to \mathbb{R}$.
  - A utility function for an agent and profile must be consistent with the ordinal preferences of that agent: for $a, b \in A$,

$$u_i^R(a) \geq u_i^R(b) \text{ iff } a \succsim_i b$$

.

- A utility representation associates with each profile R a tuple $(u_1^R, \dots, u_n^R)$
- The expected utility for agent i with utility function u_i of a lottery p is then $u_i(p) = \sum_{a \in A} p(a) u_i(a)$, and

Utility Representations

We need a way to reason about the preferences that agents have over lotteries.

- For each preference profile R and agent i, we have a utility function $u_i^R : A \to \mathbb{R}$.
  - A utility function for an agent and profile must be consistent with the ordinal preferences of that agent: for $a, b \in A$,

$$u_i^R(a) \geq u_i^R(b) \text{ iff } a \succsim_i b$$

.

- A utility representation associates with each profile R a tuple $(u_1^R, \ldots, u_n^R)$
- The expected utility for agent i with utility function u_i of a lottery p is then $u_i(p) = \sum_{a \in A} p(a) u_i(a)$, and
- agent i prefers p to q if $u_i(p) \geq u_i(q)$.

Axioms: Efficiency and Strategyproofness

Efficiency

- Given a utility representation u and a profile R, a lottery p u-dominates a lottery q if
  (i) $u_i^R(p) \geq u_i^R(p)$ for all $i \in N$, and
  (ii) $u_i^R(p) > u_i^R(p)$ for some $i \in N$.

Axioms: Efficiency and Strategyproofness

Efficiency

- Given a utility representation u and a profile R, a lottery p u-dominates a lottery q if
  - (i) $u_i^R(p) \geq u_i^R(p)$ for all $i \in N$, and
  - (ii) $u_i^R(p) > u_i^R(p)$ for some $i \in N$.
- Attempt 1: an SDS F is u-efficient if it never returns u-dominated lotteries.

Axioms: Efficiency and Strategyproofness

Efficiency

- Given a utility representation u and a profile R, a lottery p u-dominates a lottery q if
  (i) $u_i^R(p) \geq u_i^R(p)$ for all $i \in N$, and
  (ii) $u_i^R(p) > u_i^R(p)$ for some $i \in N$.
- Attempt 1: an SDS F is u-efficient if it never returns u-dominated lotteries.

Strategyproofness

- Given a utility representation u, an SDS F can be u-manipulated at R by agent i reporting $\succsim_i'$ if $u_i^R(F(\succsim_i', R_{-i})) > u_i^R(F(R))$.

Axioms: Efficiency and Strategyproofness

Efficiency

- Given a utility representation u and a profile R, a lottery p u-dominates a lottery q if
  (i) $u_i^R(p) \geq u_i^R(p)$ for all $i \in N$, and
  (ii) $u_i^R(p) > u_i^R(p)$ for some $i \in N$.
- Attempt 1: an SDS F is u-efficient if it never returns u-dominated lotteries.

Strategyproofness

- Given a utility representation u, an SDS F can be u-manipulated at R by agent i reporting $\succsim_i'$ if $u_i^R(F(\succsim_i', R_{-i})) > u_i^R(F(R))$.
- Attempt 1: an SDS F is strategyproof if there is no profile R, agent i and preference relation $\succsim_i'$, such that it can be u-manipulated at R by agent i reporting $\succsim_i'$.

Axioms: Efficiency and Strategyproofness continued

Problem! How do we decide on a specific utility function for each agent? We can't!

Solution: quantify over all consistent utility function $\Longrightarrow$ weaker notions.

Axioms: Efficiency and Strategyproofness continued

Problem! How do we decide on a specific utility function for each agent? We can't!

Solution: quantify over all consistent utility function $\Longrightarrow$ weaker notions.

Definition (Efficiency)

An SDS is efficient if it never returns a lottery that is u-dominated for all utility representations u.

Axioms: Efficiency and Strategyproofness continued

Problem! How do we decide on a specific utility function for each agent? We can't!

Solution: quantify over all consistent utility function $\Longrightarrow$ weaker notions.

Definition (Efficiency)

An SDS is efficient if it never returns a lottery that is u-dominated for all utility representations u.

Definition (Strategyproofness)

An SDS is manipulable if there is a profile R, agent i and a preference relation $\succsim_i'$ such that it is u-manipulable at R by agent i reporting $\succsim_i'$ for all utility representations u.

Axioms: Efficiency and Strategyproofness continued

Problem! How do we decide on a specific utility function for each agent? We can't!

Solution: quantify over all consistent utility function $\Longrightarrow$ weaker notions.

Definition (Efficiency)

An SDS is efficient if it never returns a lottery that is u-dominated for all utility representations u.

Definition (Strategyproofness)

An SDS is manipulable if there is a profile R, agent i and a preference relation $\succsim_i'$ such that it is u-manipulable at R by agent i reporting $\succsim_i'$ for all utility representations u.

An SDS is strategyproof if it is not manipulable.

Why are these notions weaker?

The Result

Theorem (3.1)

If $m \geq 4$ and $n \geq 4$, then there is no anonymous and neutral SDS that satisfies efficiency and strategyproofness.

The Result

Theorem (3.1)

If $m \geq 4$ and $n \geq 4$, then there is no anonymous and neutral SDS that satisfies efficiency and strategyproofness.

A new result!

Generalises other outcomes that concern:

- Restricted class of SDSs.
- Stronger notions of efficiency and strategyproofness (i.e., weaker statement).

Some related results for assignments are implied.

Proving It

Lemma ("Base Case")

If $m = 4$ and $n = 4$, then there is no anonymous and neutral SDS that satisfies efficiency and strategyproofness.

Computer aided proof using an SMT solver.

Lemma (Reduction/Preservation)

If there is an anonymous and neutral SDS F satisfying efficiency and neutrality for m alternatives and n agents, then for all $m' \leq m$ and $n' \leq n$, there is an SDS F' defined for m' alternatives and n' agents that satisfies these four properties.

Satisfaction Modulo Theories

Satisfaction modulo theories is the problem of determining whether a mathematical formula is satisfiable given a theory in which it is interpreted.

Satisfaction Modulo Theories

Satisfaction modulo theories is the problem of determining whether a mathematical formula is satisfiable given a theory in which it is interpreted.

The language is (usually quantifier-free) first order logic, augmented with a number of predicates ($=$, $\geq$) and functions ($+$, $-$), where variables need not be binary. So SMT generalizes SAT.

Satisfaction Modulo Theories

Satisfaction modulo theories is the problem of determining whether a mathematical formula is satisfiable given a theory in which it is interpreted.

The language is (usually quantifier-free) first order logic, augmented with a number of predicates ($=$, $\geq$) and functions ($+$, $-$), where variables need not be binary. So SMT generalizes SAT.

As the outcomes of SDSs are lotteries, we are concerned with the theory of (quantifier-free) linear real arithmetic.

Encoding the problem in SMT

Four kinds of SMT constraints:

- lottery definitions,
- the orbit condition (deals with a part of neutrality)
- strategyproofness
- efficiency

Other constraints, e.g., anonymity, are encoded in the representation of preference profiles.

Variables and the Lottery Constraints

Given a number of agents n and a set of alternatives A, we encode an SDS $F : \mathcal{R}(A)^n \rightarrow \Delta(A)$ with real-valued variables $p_{R,a}$, where $p_{R,a}$ represents the probability with which a is selected in profile R ($F(R)(a) = p_{R,a}$).

Variables and the Lottery Constraints

Given a number of agents n and a set of alternatives A, we encode an SDS $F : \mathcal{R}(A)^n \rightarrow \Delta(A)$ with real-valued variables $p_{R,a}$, where $p_{R,a}$ represents the probability with which a is selected in profile R ($F(R)(a) = p_{R,a}$).

Lottery constraints

$$\sum_{a \in A} p_{R,a} = 1 \text{ for all } R \in \mathcal{R}(A)^n$$

$$p_{R,a} \geq 0 \text{ for all } R \in \mathcal{R}(A)^n \text{ and } a \in A$$

Neutrality and Anonymity: Canonical Representations

We consider only the canonical representation $R_c \in \mathcal{R}(A)^n$ for every $R \in \mathcal{R}(A)^n$.

Central idea: R_c and R'_c are equal iff one can be obtained from the other by renaming the agents and alternatives. I.e., iff $F(R_c)$ and $F(R'_c)$ are equal (modulo renaming alternatives) for any neutral and anonymous SDS F.

Neutrality and Anonymity: Canonical Representations

We consider only the canonical representation $R_c \in \mathcal{R}(A)^n$ for every $R \in \mathcal{R}(A)^n$.

Central idea: R_c and R'_c are equal iff one can be obtained from the other by renaming the agents and alternatives. I.e., iff $F(R_c)$ and $F(R'_c)$ are equal (modulo renaming alternatives) for any neutral and anonymous SDS F.

Advantages: simple encoding (no permutations) and computationally lean! But how?

Neutrality and Anonymity: Canonical Representations

We consider only the canonical representation $R_c \in \mathcal{R}(A)^n$ for every $R \in \mathcal{R}(A)^n$.

Central idea: R_c and R'_c are equal iff one can be obtained from the other by renaming the agents and alternatives. I.e., iff $F(R_c)$ and $F(R'_c)$ are equal (modulo renaming alternatives) for any neutral and anonymous SDS F.

Advantages: simple encoding (no permutations) and computationally lean! But how?

Anonymity: identify each R with a function $r : \mathcal{R}(A) \to \mathbb{N}$ that tells us how often each preference relation is submitted in R.

$$r(\succsim) = |\{i \in N \mid \succsim_i = \succsim\}|$$

Canonical Representations continued

Neutrality:

(1) Given r, compute all (!) 'anonymous' preference profiles $\pi(r)$ that can be achieved via a permutation $\pi : A \to A$.

Canonical Representations continued

Neutrality:

(1) Given r, compute all (!) ‘anonymous’ preference profiles $\pi(r)$ that can be achieved via a permutation $\pi : A \to A$.

(2) Choose the lexicographically minimal profile $\pi_{\text{lexmin}}(r)$ (using some ordering on $\mathcal{R}(A)$.

Canonical Representations continued

Neutrality:

(1) Given r, compute all (!) 'anonymous' preference profiles $\pi(r)$ that can be achieved via a permutation $\pi : A \to A$.

(2) Choose the lexicographically minimal profile $\pi_{\text{lexmin}}(r)$ (using some ordering on $\mathcal{R}(A)$.

(3) Choose the smallest profile R' (in the ordering on $\mathcal{R}(A)$) that agrees with $\pi_{\text{lexmin}}(r)$.

Canonical Representations continued

Neutrality:

(1) Given r, compute all (!) 'anonymous' preference profiles $\pi(r)$ that can be achieved via a permutation $\pi : A \to A$.

(2) Choose the lexicographically minimal profile $\pi_{\text{lexmin}}(r)$ (using some ordering on $\mathcal{R}(A)$.

(3) Choose the smallest profile R' (in the ordering on $\mathcal{R}(A)$) that agrees with $\pi_{\text{lexmin}}(r)$.

This is sufficient for the result, but does not fully capture neutrality. We need the orbit condition.

The Orbit Condition

Two alternatives $a, b \in A$ are said to be equivalent if $\pi(a) = b$ for some permutation $\pi : A \to A$ that maps the anonymous preference relation associated with R to itself.

The orbit of profile R is then class of all equivalent alternatives.

The orbit condition requires that any anonymous and neutral SDS has to assign equal probabilities to all equivalent alternatives:

Orbit constraint

For each canonical profile R_c, orbit O of R_c, and two alternatives $a, b \in O$:

$$p_{R,a} = p_{R,b}.$$

Stochastic Dominance

Informally, lottery p stochastically dominates lottery q for agent i (denoted $p \succsim_i^{SD} q$) if for any alternative $a \in A$, p is at least as likely as q to yield an alternative at least as good as a.

Stochastic Dominance

Informally, lottery p stochastically dominates lottery q for agent i (denoted $p \succsim_i^{SD} q$) if for any alternative $a \in A$, p is at least as likely as q to yield an alternative at least as good as a.

Formally:

$$p \succsim_i^{SD} q \iff \sum_{b \succsim_i a} p(b) \geq \sum_{b \succsim_i a} q(b) \text{ for all } a \in A.$$

Stochastic Dominance

Informally, lottery p stochastically dominates lottery q for agent i (denoted $p \succsim_i^{SD} q$) if for any alternative $a \in A$, p is at least as likely as q to yield an alternative at least as good as a.

Formally:

$$p \succsim_i^{SD} q \iff \sum_{b \succsim_i a} p(b) \geq \sum_{b \succsim_i a} q(b) \text{ for all } a \in A.$$

Lemma (4.3)

Let $\succsim_i \in \mathcal{R}(A)$. A lottery p SD-dominates another lottery q for agent i iff $u_i(p) \geq u_i(q)$ for every utility function u_i compatible with $\succsim_i$.

Stochastic Dominance

Informally, lottery p stochastically dominates lottery q for agent i (denoted $p \succsim_i^{SD} q$) if for any alternative $a \in A$, p is at least as likely as q to yield an alternative at least as good as a.

Formally:

$$p \succsim_i^{SD} q \iff \sum_{b \succsim_i a} p(b) \geq \sum_{b \succsim_i a} q(b) \text{ for all } a \in A.$$

Lemma (4.3)

Let $\succsim_i \in \mathcal{R}(A)$. A lottery p SD-dominates another lottery q for agent i iff $u_i(p) \geq u_i(q)$ for every utility function u_i compatible with $\succsim_i$.

Stochastic dominance allows us to avoid quantifying over utility functions!

Stochastic Dominance, Efficiency, and Strategyproofness

Corollary (4.3.1 - Efficiency)

An SDS F is efficient iff, for all $R \in \mathcal{R}(A)^n$, there is no lottery p such that:

(i) $p \succsim_i^{SD} F(R)$ *for all* $i \in N$, *and*

(ii) $p \succ_i^{SD} F(R)$ *for some* $i \in N$.

Stochastic Dominance, Efficiency, and Strategyproofness

Corollary (4.3.1 - Efficiency)

An SDS F is efficient iff, for all $R \in \mathcal{R}(A)^n$, there is no lottery p such that:

(i) $p \succsim_i^{SD} F(R)$ *for all* $i \in N$, *and*

(ii) $p \succ_i^{SD} F(R)$ *for some* $i \in N$.

Corollary (4.3.2 - Strategyproofness)

An SDS F is manipulable iff there exist a profile R, agent i, and a preference relation $\succsim_i'$ such that $F(\succsim_i', R_{-i}) \succ_i^{SD} F(R)$.

Encoding Strategyproofness

For each (canonical) profile R, agent i and preference relation $\succsim_i'$, we encode that the manipulated outcome $F(\succsim_i', R_{-i})$ is not SD-preferred by the the truthful outcome $F(R)$:

Encoding Strategyproofness

For each (canonical) profile R, agent i and preference relation $\succsim_i'$, we encode that the manipulated outcome $F(\succsim_i', R_{-i})$ is not SD-preferred by the the truthful outcome $F(R)$:

$$
\begin{aligned}
&\neg \left(f(R^{i\mapsto\succsim}) \succ_i^{SD} f(R) \right) \\
&\quad \equiv f(R^{i\mapsto\succsim}) \not\succsim_i^{SD} f(R) \vee f(R) \succsim_i^{SD} f(R^{i\mapsto\succsim}) \\
&\quad \equiv \left((\exists x \in A) \sum_{y \succsim_i x} f(R^{i\mapsto\succsim})(y) < \sum_{y \succsim_i x} f(R)(y) \right) \vee \left((\forall x \in A) \sum_{y \succsim_i x} f(R^{i\mapsto\succsim})(y) \overset{(*)}{\leq} \sum_{y \succsim_i x} f(R)(y) \right) \\
&\quad \equiv \left(\bigvee_{x \in A} \sum_{y \succsim_i x} p_{(R^{i\mapsto\succsim})_c, \pi_c^{R^{i\mapsto\succsim}}(y)} < \sum_{y \succsim_i x} p_{R,y} \right) \vee \left(\bigwedge_{x \in A} \sum_{y \succsim_i x} p_{(R^{i\mapsto\succsim})_c, \pi_c^{R^{i\mapsto\succsim}}(y)} \overset{(**)}{=} \sum_{y \succsim_i x} p_{R,y} \right),
\end{aligned}
$$

Encoding Efficiency

Problem: we also have to quantify over the set of all lotteries $\Delta(A)$.

Solution: two lemmas from Aziz et al. (2015).

Encoding Efficiency

Problem: we also have to quantify over the set of all lotteries $\Delta(A)$.

Solution: two lemmas from Aziz et al. (2015).

Lemma (4.4)

Let $R \in \mathcal{R}(A)^n$. A lottery $p \in \Delta(A)$ is efficient iff every lottery $p' \in \Delta(A)$ with $supp(p') \subseteq supp(p)$ is efficient.

Encoding Efficiency

Problem: we also have to quantify over the set of all lotteries $\Delta(A)$.

Solution: two lemmas from Aziz et al. (2015).

Lemma (4.4)

Let $R \in \mathcal{R}(A)^n$*. A lottery* $p \in \Delta(A)$ *is efficient iff every lottery* $p' \in \Delta(A)$ *with* $supp(p') \subseteq supp(p)$ *is efficient.*

Lemma (4.5)

Whether a lottery $p \in \Delta(A)$ *is efficient for a given profile* R *can be computed in polynomial time by solving a linear program.*

Encoding Efficiency continued

Lemma 4.4 tells us that the efficiency of a lottery depends only on its support, thus we can speak of efficient and inefficient support.

Via lemma 4.3, an SDS is efficient iff it never returns a lottery with insufficient support.

Consequently, an SDS is efficient iff for any (canonical) profile R and any inefficient support $I_R \subseteq A$ for R, the lottery assigned to R must assign a probability of 0 to at least one alternative in the inefficient support.

Encoding Efficiency continued

Lemma 4.4 tells us that the efficiency of a lottery depends only on its support, thus we can speak of efficient and inefficient support.

Via lemma 4.3, an SDS is efficient iff it never returns a lottery with insufficient support.

Consequently, an SDS is efficient iff for any (canonical) profile R and any inefficient support $I_R \subseteq A$ for R, the lottery assigned to R must assign a probability of 0 to at least one alternative in the inefficient support.

Efficiency Constraint

For each (canonical) profile $R \in \mathcal{R}(A)^n$ and each inefficient support $I_R \subseteq A$:

$$\bigvee_{a \in I_R} p_{R,a} = 0.$$

Verification of Correctness

Drawbacks of the SMT-based proof:

(i) one must trust the SMT solver,

(ii) one must trust the correctness of the program that performs the encoding, and

(iii) the proof is virtually impossible to be checked by humans.

Verification of Correctness

Drawbacks of the SMT-based proof:

(i) one must trust the SMT solver,
(ii) one must trust the correctness of the program that performs the encoding, and
(iii) the proof is virtually impossible to be checked by humans.

Solutions:

(i) Generate a MUS and use other solvers to verify that it is indeed unsatisfiable.
(ii) Run solvers on different variants of the encoding to reproduce known results.
(iii) Translate MUS into an independent proof in HOL using a generic interactive theorem prover (not automated!).

Concluding Remarks and...

Concluding Remarks and...

Questions?

References

[1] Brandl, F., Brandt, F., Eberl, M., & Geist, C. (2018). Proving the incompatibility of efficiency and strategyproofness via SMT solving. *Journal of the ACM (JACM), 65*(2), 1-28.

[2] Aziz, H., Brandl, F., & Brandt, F. (2015). Universal Pareto dominance and welfare for plausible utility functions. *Journal of Mathematical Economics, 60*, 123-133.