CS386 Assignment 6

Due: Thursday 11/5 10 PM ET

Throughout, a *profile* is a list of strict rankings of the candidates, one per voter. We write $a \succ_i b$ if voter i ranks a above b, and write rankings top-to-bottom as strings, so *bac* means $b \succ a \succ c$. A *social choice function* outputs one winner; a *social welfare function* outputs a ranking. Unless said otherwise, ties are broken alphabetically.

1. **Warm-ups.** Do 4 of the following.

   (a) *Plurality is not IIA.* View plurality as a social welfare function: rank the candidates by their number of first-place votes. Consider the profile P with 3 voters *abc*, 2 voters *bca*, and 2 voters *cba*.

      i. What ranking does plurality output on P?

      ii. Build a new profile P' from P in which *no voter changes how they rank a versus b*, but plurality outputs a ranking with b above a. Explain why this shows that plurality is not IIA. (Which voters can change their ranking without changing their a vs. b preference, and what does that do to the first-place counts?)

   (b) *IRV is not monotone.* Consider the profile with 2 voters *acb*, 3 voters *bac*, and 3 voters *cba*.

      i. Run IRV. Who wins?

      ii. One of the *bac* voters changes their ranking to *acb*. Run IRV again. Who wins now?

      iii. Explain why this violates monotonicity (see problem 2 if you need definition).

   (c) *Borda is not Condorcet consistent.* Consider the profile with 3 voters *bca* and 2 voters *cab*. Find the Condorcet winner, and compute all the Borda scores. Who wins under Borda?

   (d) *Pareto.* A social choice function satisfies **Pareto** if whenever every voter ranks a above b, then b does not win. Show that Borda and IRV both satisfy Pareto. (For IRV: while a is still in the race, how many first-place votes can b have?)

   (e) *The Condorcet paradox, in general.* For every $m \geq 3$ candidates $c_1, \ldots, c_m$, give a profile with no Condorcet winner. (Use m voters, where voter k ranks the candidates $c_k, c_{k+1}, \ldots, c_m, c_1, \ldots, c_{k-1}$. Show every candidate loses a head-to-head majority vote to someone.)

   (f) *Intransitive dice.* A pair of dice is rolled independently, and "$X > Y$" means that die X shows a larger number than die Y. Each die has 6 sides, and the faces can be any integers (repeats allowed). Choose numbers for three dice A, B, C so that $\Pr[A > B] = \Pr[B > C] = \Pr[C > A] > \frac{1}{2}$. (Recall the condorcet paradox from class)

(g) *A branching algorithm for* 3SAT. This one is just for fun and is *not* material for the second midterm. Let φ be a formula in conjunctive normal form in which every clause has at most 3 literals, over n variables. Consider the following algorithm SOLVE(φ).

- If φ has no clauses, accept. If φ contains an empty clause, reject.
- Otherwise pick any clause C, and let $x_1, \ldots, x_k$ ($k \leq 3$) be the variables appearing in it. For each of the $2^k - 1$ assignments to $x_1, \ldots, x_k$ that make C true, substitute the assignment into φ (delete the clauses it satisfies and delete the literals it falsifies), and call SOLVE on the result, which has $n - k$ variables left. Accept if and only if some call accepts.

i. Explain why the algorithm is correct. (Why do we lose nothing by only trying assignments that satisfy C?)

ii. Show that the recursion tree has at most $7^{n/3}$ leaves. (Induct on n, and check each of the cases $k = 1, 2, 3$. Note $3 \leq 7^{2/3}$.) Conclude that the algorithm runs in time $\left(7^{1/3}\right)^n \cdot \text{poly}(|\varphi|)$.

iii. Write the running time as $2^{\delta n}$ (ignoring the polynomial factor) and compute δ. Check that $\left(7^{1/3}\right)^n < 2^n$, so this beats brute force over all 2^n assignments. Recall that the Exponential Time Hypothesis (ETH) says that there is some constant $\delta^* > 0$ such that no algorithm solves 3SAT in time $O(2^{\delta^* n})$. Why does the algorithm above not disprove the ETH? What would an algorithm need to achieve in order to do so?

(h) *Arrow's theorem, part 1: Pareto.* Suppose F is a social welfare function on $|A| \geq 3$ candidates that satisfies unanimity and IIA. Show that if every voter ranks b above a in $\prec$, then $b \succ_F a$, where $x \succ_F y$ means that x is above y in $F(\prec)$.

(i) *Arrow's theorem, part 2: a pivotal voter exists.* Let F be as in (h), and fix two candidates $a \neq b$. For $i = 0, 1, \ldots, n$, let $\prec^{(i)}$ be any profile in which voters $1, \ldots, i$ rank a above b and voters $i+1, \ldots, n$ rank b above a (the rest of each ranking is arbitrary). Write $x \succ_F^{(i)} y$ if x is above y in $F(\prec^{(i)})$.

(1) Using Pareto, what do $F(\prec^{(0)})$ and $F(\prec^{(n)})$ say about a and b?

(2) Conclude that there is an index $i^* \in \{1, \ldots, n\}$ such that $b \succ_F^{(i^*-1)} a$ but $a \succ_F^{(i^*)} b$.

(3) Use IIA to explain why i^* does not depend on how you filled in the rest of the rankings in each $\prec^{(i)}$.

(j) *Arrow's theorem, part 3: the pivotal voter is a dictator.* **This one counts as 2 problems.** Let F, a, b, i^* be as in (i). Show that for every profile $\prec$, if $c \neq d$ and $c \succ_{i^*} d$, then $c \succ_F d$. Conclude that $F(\prec) = \prec_{i^*}$ for every profile $\prec$, which finishes the proof of Arrow's theorem. You may use the pairwise neutrality lemma from Problem 4.

(1) Pick $e \notin \{c, d\}$ (why does one exist?). Let $\prec'$ be the profile obtained from $\prec$ as follows. For voters $j < i^*$, move e to the bottom of their ranking. For voters $j > i^*$, move e to the top. For voter i^*, move e so that $c \succ'_{i^*} e \succ'_{i^*} d$. Show that $\prec$ and $\prec'$ rank c versus d the same way for every voter.

(2) Show that for every voter j,

$$c \succ'_j e \iff a \succ_j^{(i^*)} b \qquad \text{and} \qquad e \succ'_j d \iff b \succ_j^{(i^*-1)} a.$$

(3) Apply pairwise neutrality twice, to the pairs (a, b) and (c, e), and to the pairs (b, a) and (e, d), to get $c \succ'_F e$ and $e \succ'_F d$. Then use IIA to conclude $c \succ_F d$.

(4) Why does this show that voter i^* is a dictator?

2. **Incentive compatibility and monotonicity.** Let f be a social choice function on profiles of strict rankings. Recall that f is **incentive compatible** (IC) if no voter can get an outcome they prefer by misreporting their ranking: for every profile and every voter i with true ranking $\prec_i$ and any alternative report $\prec_i'$, if $a = f(\prec_i, \prec_{-i})$ and $a' = f(\prec_i', \prec_{-i})$, then either $a = a'$ or $a' \prec_i a$ (here $x \prec_i y$ means i prefers y to x). Recall also that f is **monotone** if $f(\prec_i, \prec_{-i}) = a \neq a' = f(\prec_i', \prec_{-i})$ implies that $a' \prec_i a$ and $a \prec_i' a'$.

   (a) Show that if f is IC then f is monotone. (Apply the definition of IC twice: once when $\prec_i$ is the true ranking, and once when $\prec_i'$ is.)

   (b) Show that if f is monotone then f is IC.

   (c) In the profile from 1(b), one voter can profitably lie under IRV. Say which voter, what they report, and why this is a manipulation. Which of (a) or (b) tells you this must be possible?

3. **The hypotheses in Arrow's theorem matter** Arrow's theorem says that every social welfare function on $|A| \geq 3$ candidates that satisfies unanimity and IIA is a dictatorship. Here a *dictatorship* is a function for which some fixed voter d has $f(\prec_1, \ldots, \prec_n) = \prec_d$ for every profile, and *unanimity* says that if $\prec_i = \prec$ for all i, then $f(\prec_1, \ldots, \prec_n) = \prec$. For each hypothesis below, give a social welfare function that satisfies all the *other* hypotheses and prove it does. Assume $n \geq 3$ voters.

   (a) Drop unanimity: find a non-dictatorial, IIA social welfare function.

   (b) Drop IIA: find a non-dictatorial social welfare function satisfying unanimity.

   (c) Drop $|A| \geq 3$: for $|A| = 2$ and n odd, find a non-dictatorial social welfare function satisfying unanimity and IIA.

4. **Pairwise neutrality** In class we proved Arrow's theorem using the following lemma, which we stated without proof. Throughout, F is a social welfare function on $|A| \geq 3$ candidates that satisfies unanimity and IIA. For a profile $\prec$, write $x \succ_F y$ if x is above y in $F(\prec)$, and write $x \succ_F' y$ if x is above y in $F(\prec')$. You may use that F satisfies **Pareto**: if every voter ranks b above a, then $b \succ_F a$ (see warm-up 1(h)).

   Lemma (pairwise neutrality). Let $a \neq b$ and $c \neq d$ be candidates, and let $\prec$ and $\prec'$ be profiles such that for every voter i,

   $$a \succ_i b \iff c \succ_i' d.$$

   Then $a \succ_F b \iff c \succ_F' d$.

   (a) First assume that a, b, c, d are all distinct. It suffices to show the forward direction $a \succ_F b \Rightarrow c \succ_F' d$. (For the reverse direction, apply the forward direction to the pairs (b, a) and (d, c).) Build a new profile $\prec''$ from $\prec$ by changing each voter's ranking as follows: move c to just above a, and move d to just below b, leaving everything else where it was.

      i. Check that every voter ranks a versus b in $\prec''$ the same way as in $\prec$.

      ii. Check that every voter ranks c versus d in $\prec''$ the same way as in $\prec'$. (Look at the two cases $a \succ_i b$ and $b \succ_i a$ separately, and use the assumption on $\prec$ and $\prec'$.)

      iii. Check that every voter ranks c above a and b above d in $\prec''$.

iv. Put these together. Using IIA and Pareto, show that $a \succ_F b$ implies $c \succ'_F d$. (First get $a \succ''_F b$, and from there get a chain $c \succ''_F a \succ''_F b \succ''_F d$. Then come back to $\prec'$.)

(b) Now allow the pairs $\{a, b\}$ and $\{c, d\}$ to overlap.

i. If $a = c$ or $b = d$, explain why the same construction still works. (Which of the moves is unnecessary?)

ii. If $b = c$ or $a = d$, the construction in part (a) breaks. Say why, and then fix it by chaining the lemma through intermediate pairs. Each step of the chain should be a case you've already handled, and you are free to choose the intermediate profiles. (For the pairs (a, b) and (b, d) with $d \neq a$, go through (a, d). For (a, b) and (b, a), you need a candidate $e \notin \{a, b\}$; why does one exist?)

5. **The Gibbard–Satterthwaite theorem** A social choice function f is a *dictatorship* if there is a voter i^* such that $f(\prec)$ is the top candidate of $\prec_{i^*}$ for every profile $\prec$. We will prove: *if f is IC and onto A with $|A| \geq 3$, then f is a dictatorship.* You may use Problem 2 (so f being IC means f is monotone) and Arrow's theorem.

For a profile $\prec$ and a set of candidates $S \subseteq A$, let $\prec^S$ be the profile obtained from $\prec$ by moving the candidates in S to the top of every voter's ranking *without changing the order within S or the order within $A \setminus S$.*

(a) *Lifting a set to the top.* Show that if f is IC and onto A, then $f(\prec^S) \in S$ for every profile $\prec$ and every nonempty $S \subseteq A$. Since f is onto, there is a profile $\tilde{\prec}$ with $f(\tilde{\prec}) \in S$. Let $\tilde{\prec}^{(k)} = (\prec_1^S, \ldots, \prec_k^S, \tilde{\prec}_{k+1}, \ldots, \tilde{\prec}_n)$, so that the first k voters have already been lifted. Show by induction on k that $f(\tilde{\prec}^{(k)}) \in S$. (Suppose $f(\tilde{\prec}^{(k-1)}) = c \in S$ but $f(\tilde{\prec}^{(k)}) = c' \notin S$. What does monotonicity say about how voter k ranks c and c' in each of the two profiles? Why is that impossible when S is at the top of $\prec_k^S$?)

(b) *From f to a ranking.* Define a relation $\prec_{\underline{f}}$ on A by

$$a \prec_{\underline{f}} b \iff f\big(\prec^{\{a,b\}}\big) = b.$$

Show that this is a strict total order, so that $\underline{f}$ is a social welfare function.

- *Total and antisymmetric:* use (a) with $S = \{a, b\}$.
- *Transitive:* suppose, for contradiction, that $a \prec_{\underline{f}} b \prec_{\underline{f}} c \prec_{\underline{f}} a$ on some profile $\prec$. Let $S = \{a, b, c\}$; by (a), $f(\prec^S) \in S$. Say $f(\prec^S) = a$ (the other cases are the same with the roles rotated). Change the voters one at a time from $\prec_i^S$ to $\prec_i^{\{a,b\}}$. Show that the outcome stays a at every step. (By monotonicity, a change from a to some a' would need a' to be above a in the new ranking. Who could that be, when $\{a, b\}$ is at the top and a, b are ordered the same way as before?) Conclude that $f(\prec^{\{a,b\}}) = a$, contradicting $a \prec_{\underline{f}} b$.

(c) *Unanimity.* Show that $\underline{f}$ satisfies unanimity. (Suppose every voter ranks b above a. In $\prec^{\{a,b\}}$, b is already at the top of every ranking. Apply (a) to the profile $\prec^{\{a,b\}}$ with $S = \{b\}$.)

(d) *IIA.* Show that $\underline{f}$ satisfies IIA. (Suppose $\prec$ and $\prec'$ order a and b the same way for every voter. Compare $f(\prec^{\{a,b\}})$ and $f(\prec'^{\{a,b\}})$ by changing voters one at a time from $\prec_i^{\{a,b\}}$ to $\prec_i'^{\{a,b\}}$. In every intermediate profile, what does (a) say about the outcome? Then use monotonicity.)

(e) *Conclusion.* By Arrow's theorem, $\underline{f}$ is a dictatorship: there is a voter i^* with $\prec_{\underline{f}} = \prec_{i^*}$ for every profile. Show that $f(\prec)$ is the top candidate of $\prec_{i^*}$ for every profile $\prec$, so f is a dictatorship. (Let t be i^*'s top candidate and suppose $f(\prec) = x \neq t$. Change the voters one at a time to $\prec^{\{t,x\}}$. Use monotonicity to show that the outcome stays x throughout, so $f(\prec^{\{t,x\}}) = x$. What does that say about t and x in $\prec_{\underline{f}}$?)