CS 405/6001: Game Theory and Algorithmic Mechanism Design **Jul–Nov 2026**

Lecture 15: The Social Choice Setup and Introduction to Voting

Lecturer: Swaprava Nath *Scribe(s): Claude Scribe*

Disclaimer: *These notes aggregate content from several texts and have not been subjected to the usual scrutiny deserved by formal publications. If you find errors, please bring to the notice of the Instructor.*

15.1 Recap

For the last several lectures we studied strategic-form games: fixed sets of players, each independently choosing an action, with equilibrium notions like PSNE and MSNE describing when nobody wants to unilaterally deviate. Today we turn to a different but related question: how does a *group* of agents, each with their own ranking over a common set of alternatives, arrive at a single collective decision? This is the domain of **social choice theory**, and it will occupy us for the rest of the week, building up to the Gibbard–Satterthwaite impossibility theorem.

15.2 The Social Choice Setup

Definition 15.1 (Alternatives, agents, preferences) *Let $A = \{a_1, \ldots, a_m\}$ be a finite set of **alternatives**, and $N = \{1, \ldots, n\}$ a finite set of **agents**. Each agent i has a linear (strict, complete, transitive) preference ordering P_i over A; write $\mathcal{P}$ for the set of all such orderings. A **preference profile** is $P = (P_1, \ldots, P_n) \in \mathcal{P}^n$.*

Definition 15.2 (Social welfare function and social choice function) *A **social welfare function** maps every profile P to a complete social ordering of A. A **social choice function (SCF)** is the simpler object*

$$f : \mathcal{P}^n \to A,$$

selecting a single winning alternative for every profile.

Example 15.3 (Choosing a generative AI model) *A company must deploy exactly one generative AI model $A = \{a, b, c, d, \ldots\}$ as its customer-support assistant. A review committee shortlists the candidates, and each member ranks them by their own criteria: one weighs benchmark accuracy most heavily, another latency and cost per token, another safety and refusal behaviour. An SCF f takes everyone's ranking and outputs the single model that ships. This is the running example we use throughout this lecture.*

SCFs are the right object when the goal is to pick a single winner (an election, a deployed model, an adopted policy), rather than to rank every alternative. This distinction matters because, as we see next, ranking *everything* turns out to be impossible under very mild conditions.

15.3 Arrow's Impossibility Theorem

Definition 15.4 (Three fairness conditions)

- ***Weak Pareto****: if every agent strictly prefers a to b ($a \succ_i b\ \forall i$), the social ordering places a above b. For an SCF, this means: if everyone's top choice is a, then $f(P) = a$.*
- ***Independence of Irrelevant Alternatives (IIA)****: the social comparison of a against b depends only on how each agent individually ranks a against b – not on where a third alternative c sits in anyone's ranking.*
- ***Non-dictatorship****: there is no single agent i whose top choice is always the social outcome, regardless of the rest of the profile.*

Theorem 15.5 (Arrow, 1951) *If $|A| \geqslant 3$, there is no social welfare function that simultaneously satisfies Weak Pareto, IIA, and Non-dictatorship.*

Each condition in Definition 15.4 looks like a bare minimum any reasonable aggregation procedure should satisfy. Theorem 15.5 says that once there are three or more alternatives, they cannot all hold together – not for any conceivable rule, however cleverly designed.

Remark 15.6 (Two escape routes) *Since a full social* ranking *is what Theorem 15.5 rules out, there are two natural ways to make progress:*

1. ***Change the object of choice****: ask only for a single winner, i.e. study SCFs $f : \mathcal{P}^n \to A$ rather than full social welfare functions. This is the route we take for the rest of the week.*
2. ***Restrict the domain of preferences****: assume agents' preferences have some structure (e.g. single-peaked preferences, met in Lecture 16), which reopens the door to much stronger positive results.*

15.4 Voting and Scoring Rules

Voting is the most familiar instance of social choice. Consider four shortlisted models $A = \{a, b, c, d\}$ and four committee members with the profile

M1	M2	M3	M4
a	a	c	d
b	b	b	c
c	c	d	b
d	d	a	a

Definition 15.7 (Scoring rule) *A **scoring rule** fixes a score vector $(s_1, \ldots, s_m)$ with $s_i \geqslant s_{i+1}$, awards s_k points to whichever alternative sits in position k of a voter's ranking, sums each alternative's points across voters, and declares the highest total the winner.*

Common instances: **plurality** $(1, 0, \ldots, 0)$, **veto** $(1, 1, \ldots, 1, 0)$, **Borda count** $(m-1, m-2, \ldots, 1, 0)$, the **harmonic rule** $\left(1, \frac{1}{2}, \ldots, \frac{1}{m}\right)$, and k**-approval** (top k positions get 1, rest get 0).

Not every rule of interest is a scoring rule. **Plurality with runoff** (the two-round system, used in French presidential elections) holds a plurality vote, keeps only the top two, and re-runs plurality between them. The **maximin** rule scores each alternative by its worst pairwise performance,

$$\text{score}(a) = \min_{y \in A \setminus \{a\}} |\{i : a \succ_i y\}|,$$

and **Copeland** scores each alternative by how many pairwise contests it wins outright.

Example 15.8 (Maximin on the model profile) *On the profile above,*

$$\begin{aligned} \textit{score}(a) &= \min\{2,2,2\} = 2, & \textit{score}(b) &= \min\{2,2,3\} = 2, \\ \textit{score}(c) &= \min\{2,2,3\} = 2, & \textit{score}(d) &= \min\{2,1,1\} = 1. \end{aligned}$$

Alternatives a, b, c tie at the top; d is weakest because it loses badly to c head-to-head.

15.5 Condorcet Consistency

Definition 15.9 (Condorcet winner and consistency) *Alternative x is a **Condorcet winner** if it beats every other alternative in a pairwise majority vote: $\forall y \neq x$, $|\{i : x \succ_i y\}| > |\{i : y \succ_i x\}|$. An SCF is **Condorcet consistent** if it always selects the Condorcet winner whenever one exists.*

A Condorcet winner need not exist. Three members with $a \succ b \succ c$, $b \succ c \succ a$, $c \succ a \succ b$ produce a cycle: a beats b, b beats c, c beats a, each 2–1. No alternative survives every pairwise contest.

Example 15.10 (A Condorcet winner that plurality misses) *The org polls 100 engineers. Their rankings, as shares:* 30% *report $a \succ b \succ c$,* 30% *report $b \succ a \succ c$,* 40% *report $c \succ b \succ a$. Pairwise, b beats a (*70–30*), a beats c (*60–40*), and b beats c (*60–40*) – so b is the Condorcet winner. Yet plurality, counting only first-place votes, gives a and b* 30% *each and c a plurality* 40%*, so* plurality selects *c: the candidate that loses head-to-head to b. **No scoring rule is Condorcet consistent in general.** Copeland and maximin, by contrast, both correctly select b here.*

Remark 15.11 (The story does not change the mathematics) *The same numbers describe a city council voting on competing pricing policies, or any other three-way contest – the cycle and the* 30/30/40 *split behave identically; only the labels on a, b, c change.*

15.6 Pareto Efficiency and Unanimity

Definition 15.12 (Pareto domination and efficiency) *Alternative a is **Pareto dominated** by b at profile P if $b \succ_i a$ for every agent i. SCF f is **Pareto efficient (PE)** if it never selects a Pareto dominated alternative.*

Definition 15.13 (Unanimity) *SCF f is **unanimous (UN)** if whenever every agent's top choice is a, then $f(P) = a$.*

Proposition 15.14 *PE $\implies$ UN.*

Proof: If every agent ranks a first, every other alternative $b \neq a$ is Pareto dominated by a. PE then forbids $f(P) = b$ for any such b, forcing $f(P) = a$. ■

Remark 15.15 (UN does not imply PE) *The converse of Proposition 15.14 fails. Consider the SCF that outputs the unanimous top choice when the profile is unanimous, and outputs a fixed alternative b otherwise. This rule is unanimous by construction. But at a non-unanimous profile where every agent prefers c to b, the rule still returns b – a Pareto dominated outcome. So UN does not imply PE.*

Definition 15.16 (Onto) *SCF f is **onto (ONTO)** if every $a \in A$ is $f(P)$ for some profile P.*

Since unanimously ranking any a first forces $f(P) = a$, we get UN $\implies$ ONTO immediately.

15.7 Manipulability

Definition 15.17 (Manipulability) *SCF f is **manipulable** if some agent i and profile P admit a misreport P_i' with $f(P_i', P_{-i}) \succ_i f(P_i, P_{-i})$.*

Example 15.18 (Plurality is manipulable) *Three alternatives $\{a, b, c\}$, fixed tie-break $a \succ b \succ c$, nine voters:*

4	4	1
a	b	c
b	a	b
c	c	a

*Sincerely, a and b tie at 4 votes and the tie-break gives a. The lone voter who truly prefers c switches their ballot to $b \succ c \succ a$, giving b a clean majority of 5. They never wanted b most, but they strictly prefer b to a, so misreporting pays off. **Plurality with fixed tie-breaking is manipulable.***

Example 15.19 (Copeland is manipulable) *Three alternatives, tie-break $a \succ b \succ c$, three voters forming a perfect cycle $a \succ b \succ c$, $b \succ c \succ a$, $c \succ a \succ b$. Every alternative wins exactly one pairwise contest, so all Copeland scores tie at 1 and the tie-break gives a. A single voter can misreport to change the pairwise results and hence the Copeland scores, again profitably. **Even Copeland with fixed tie-breaking is manipulable.***

Remark 15.20 *Examples 15.18 and 15.19 are stated with abstract voters, but the same 4/4/1 and 1/1/1 numbers describe, respectively, a ridesharing city-council pricing vote and the model selection committee – the vulnerability to manipulation is a property of the rule and the numbers, not of the cover story.*

15.8 Strategyproofness

Definition 15.21 (Strategyproofness) *SCF f is **strategyproof (SP)** if it is not manipulable by any agent at any profile: for every i, every $P_i, P_i' \in \mathcal{P}$, every P_{-i},*

$$f(P_i, P_{-i}) \succsim_i f(P_i', P_{-i}).$$

Definition 15.22 (Dominated set) *For alternative a and preference P_i, the **dominated set** is $D(a, P_i) := \{b \in A : a \succ_i b\}$ – everything i ranks below a.*

Definition 15.23 (Monotonicity) *SCF f is **monotone (MONO)** if whenever $f(P) = a$ and $D(a, P_i) \subseteq D(a, P_i')$ for every agent i, then $f(P') = a$: if a already wins and every agent moves a up (weakly), a still wins.*

Theorem 15.24 *An SCF f is strategyproof if and only if it is monotone.*

Strategyproofness quantifies over every possible misreport, which makes it hard to check directly; monotonicity is a purely structural condition on the rule, which is usually far easier to verify. Theorem 15.24 lets us check the easy one.

Proof: (SP $\implies$ MONO.) Let $f(P) = a$ and $D(a, P_i) \subseteq D(a, P_i')$ $\forall i$. Move from P to P' one agent at a time via intermediate profiles $P^{(0)} = P, P^{(1)}, \ldots, P^{(n)} = P'$, changing one agent's report per step. Suppose for contradiction $f(P^{(k-1)}) = a$ but $f(P^{(k)}) = b \neq a$, where only agent k's report changed from P_k to P_k'. Three cases are possible:

1. $a \succ_k b$ and $a \succ_k' b$: agent k profitably misreports $P_k' \to P_k$ at profile $P^{(k)}$, changing the outcome from b to a, which k strictly prefers.
2. $b \succ_k a$ and $b \succ_k' a$: agent k profitably misreports $P_k \to P_k'$ at profile $P^{(k-1)}$, changing the outcome from a to b, which k strictly prefers.
3. $b \succ_k a$ and $a \succ_k' b$: agent k misreports in *either* direction depending on which profile is taken as true, so both $P^{(k-1)}$ and $P^{(k)}$ admit a profitable deviation.

The fourth logically possible combination, $a \succ_k b$ and $b \succ_k' a$, cannot occur: it would move b from outside $D(a, P_k)$ to inside $D(a, P_k')$, contradicting $D(a, P_k) \subseteq D(a, P_k')$. Every remaining case contradicts SP, so $f(P^{(k)}) = a$ throughout, giving $f(P') = a$.

(MONO $\implies$ SP, via contrapositive $\neg$SP $\implies$ $\neg$MONO.) Suppose agent i profitably misreports $P_i \to P_i'$, moving the outcome from $a = f(P_i, P_{-i})$ to $b = f(P_i', P_{-i})$ with $b \succ_i a$. Construct P_i'' with b ranked first and a second. From (P_i, P_{-i}), $D(a, P_i) \subseteq D(a, P_i'')$, so MONO would force $f(P_i'', P_{-i}) = a$. From (P_i', P_{-i}), $D(b, P_i') \subseteq D(b, P_i'')$ (b is now top-ranked), so MONO would force $f(P_i'', P_{-i}) = b$. Since $a \neq b$, both cannot hold: f cannot be monotone. ■

Lemma 15.25 *If SCF f is MONO and ONTO, then f is PE.*

Proof: Suppose not: f is MONO and ONTO but $\exists\, a, b, P$ with $b \succ_i a$ $\forall i$, yet $f(P) = a$. Since f is ONTO, some P' has $f(P') = b$. Construct P'' with b ranked first, a second, for every agent. Then $D(b, P_i') \subseteq D(b, P_i'')$ $\forall i$ forces $f(P'') = b$ by MONO, while $D(a, P_i) \subseteq D(a, P_i'')$ $\forall i$ forces $f(P'') = a$ by MONO. Contradiction. ■

Corollary 15.26 *For a strategyproof SCF f: PE $\iff$ UN $\iff$ ONTO.*

Proof: PE $\implies$ UN $\implies$ ONTO holds unconditionally (Proposition 15.14 and the remark following it). Lemma 15.25, combined with Theorem 15.24 (SP $\iff$ MONO), gives ONTO $\implies$ PE whenever f is SP. The cycle closes. ■

15.9 Summary and Preview

We set up the social choice function $f : \mathcal{P}^n \to A$ and saw that Arrow's impossibility theorem (Theorem 15.5) makes a full social *ranking* unattainable under mild conditions once $|A| \geqslant 3$; this motivates studying SCFs, which need only pick a winner. We surveyed scoring and non-scoring voting rules and found that only some (Copeland, maximin) are Condorcet consistent, while scoring rules in general are not (Example 15.10). We defined PE, UN, ONTO, and manipulability, and showed plurality and Copeland are both manipulable under fixed tie-breaking even though they satisfy other desirable properties. Finally, the equivalence of strategyproofness and monotonicity (Theorem 15.24) let us prove that under SP, the properties PE, UN, and ONTO all coincide (Corollary 15.26).

Next lecture: for $|A| \geqslant 3$, is there *any* strategyproof and onto rule that isn't a dictatorship? The Gibbard–Satterthwaite theorem answers this, and we build the full proof from Lemma 15.25's style of monotonicity argument.

References

[AR51] K. Arrow, "Social Choice and Individual Values," Wiley, 1951.

[GS73] A. Gibbard, "Manipulation of voting schemes: a general result," *Econometrica*, 41(4), 1973, pp. 587–601.

[Sat75] M. Satterthwaite, "Strategy-proofness and Arrow's conditions," *Journal of Economic Theory*, 10(2), 1975, pp. 187–217.