Formal Model in social choice theory

A formal model in social choice theory and game theory provides a mathematical structure to analyze how individual preferences or strategies aggregate into collective decisions or outcomes. It uses precise definitions, assumptions, and equations to model the behaviour of agents (voters, players, etc.), their preferences, and the mechanisms used to make group decisions.

Here is a formal model for a social choice function and the related concepts in social choice theory:

1. Basic Elements of a Formal Social Choice Model

- **Set of Agents (Voters)**: Let $N=\{1,2,\ldots,n\}$ be the set of voters, where n is the total number of voters.
- **Set of Alternatives (Outcomes)**: Let $X=\{x_1,x_2,\ldots,x_m\}$ represent the set of alternatives or outcomes, where m is the number of possible outcomes or policies.
- **Preferences of Voters**: Each voter $i\in N$ has a preference relation P_i over the set of outcomes X, where P_i represents a complete, transitive ranking of alternatives in X. For simplicity, this preference is usually expressed as a total order over X, meaning that voter iii ranks all outcomes in X in a strict order (though ties can be allowed in some models).
- **Social Welfare Function**: A **social welfare function (SWF)** aggregates the preferences of all voters into a collective ranking of the alternatives. Formally, the social welfare function f maps the profile of individual preferences $P=(P1,P2,\ldots,Pn)$ to a social preference relation P^* over X:

$f:P_1\times P_2\times\cdots\times P_n\rightarrow P^*$

The output P* represents the collective preference or ranking of the alternatives for society.

2. Social Choice Function (SCF)

A **social choice function (SCF)** directly selects one outcome based on the preferences of voters. It can be seen as a specific case of a social welfare function where the final output is a single alternative rather than a complete ranking.

- **Social Choice Function**: A function ccc maps a profile of individual preferences $P=(P_1,P_2,\ldots,P_n)$ to a single outcome $x*\in X$

c: $P1\times P2\times\cdots\times Pn\rightarrow X$

Here, $c(P)=x^*$ represents the alternative selected by the group based on the voters' preferences.

3. Properties of Social Choice Functions

Several properties can be desired in a social choice function or social welfare function:

- **Pareto Efficiency**: If every individual prefers outcome x over outcome y, then the social choice function should select x over y.

Formally, for all $i \in N$, if $x \, P_i \, y$, then $x \, P^* \, y$

- **Non-Dictatorship**: No single voter should always determine the outcome, regardless of the preferences of other voters. Formally, for any individual $i \in N$, there must be at least one profile where the selected outcome differs from i's top choice.
- **Independence of Irrelevant Alternatives (IIA)**: The group's ranking of any two alternatives x and y should depend only on how individuals rank x and y, and not on the ranking of other irrelevant alternatives.

Formally, if P_i changes for some $i \in N$ regarding some irrelevant alternatives, the ranking between x and y should remain unchanged.

- **Transitivity**: The social welfare function should produce a transitive ordering. If $x \, P^* \, y$ and $y \, P^* \, z$, then it should be that $x \, P^* \, z$.

4. Example of a Formal Model for Voting

Consider a **majority rule** social choice function for three voters and three alternatives $X=\{A,B,C\}$

- Voters 1, 2, and 3 have the following preferences:
  - P1: A>B>C
  - P2: B>C>A
  - P3: C>A>B
- The **majority rule** SCF selects an alternative that wins in pairwise comparisons:
  - Compare A vs. B:
    - Voter 1 prefers A, Voter 2 prefers B, Voter 3 prefers A.
    - Majority: A wins.
  - Compare A vs. C:
    - Voter 1 prefers A, Voter 2 prefers C, Voter 3 prefers C.
    - Majority: C wins.
  - Compare B vs. C:
    - Voter 1 prefers B, Voter 2 prefers B, Voter 3 prefers C.
    - Majority: B wins.

In this case, no alternative wins all pairwise comparisons. This could lead to a **Condorcet cycle**: A beats B, C beats A, and B beats C.

5. Arrow's Impossibility Theorem (Formal Statement)

Arrow's Impossibility Theorem states that no social welfare function f can satisfy all the following conditions simultaneously when there are three or more alternatives:

1. **Pareto Efficiency**: If all voters prefer one alternative to another, the social welfare function should reflect this.
2. **Non-Dictatorship**: No single voter should determine the outcome.
3. **Independence of Irrelevant Alternatives (IIA)**: The social ranking between two alternatives should depend only on voters' preferences between those two.
4. **Transitivity**: The social preference relation P* must be transitive.
5. **Unrestricted Domain**: The function must work for any possible set of individual preferences.

Arrow's theorem is a formal result showing that it's impossible to design a social welfare function that satisfies all of these properties, except in trivial cases (e.g., a dictatorship).

6. Formal Model in Game Theory

In game theory, a **formal model** for a two-player zero-sum game consists of:

- **Players**: $i=1,2$
- **Strategies**: Let S_1 and S_2 be the strategy sets for players 1 and 2, respectively.
- **Payoff functions**: Let $u_1(s_1,s_2)$ and $u_2(s_1,s_2)$ be the payoff functions of players 1 and 2, where $s_1 \in S_1$ and $s_2 \in$. In a zero-sum game, $u_1(s_1,s_2)=-u_2(s_1,s_2)$
- The **Nash equilibrium** is a pair of strategies (s_1^*,s_2^*) such that neither player can improve their payoff by unilaterally deviating from their strategy.