Amartya Sen and the mathematics of collective choice

Bhaskar Dutta

Sen's work on the causes of famines, gender discrimination and social and economic issues dealing with upliftment of the poor is largely known. However, he is better known among professional economists for his contributions in social choice theory.

AMARTYA Sen has been a very well-known economist even amongst laypersons long before the decision of the Royal Swedish Academy to award him the 1998 Nobel Prize for Economics. His fame has been largely due to his influential work on issues which are of interest to all concerned citizens. Some examples are: the causes of famines, gender discrimination, and more generally social and economic issues dealing with the upliftment of the poor. Not for nothing has he been labelled the 'conscience-keeper' amongst economists.

The social relevance of his less technical work cannot be overestimated. However, professional economists hold him in high esteem mainly because of his theoretical contributions in social choice theory and the related areas of measurement of inequality and poverty. The purpose of this essay is to provide a brief introduction to social choice theory for the non-economist, focusing on Sen's contributions.

The origins of social choice theory can be traced back to the mathematical analysis of elections and committee decisions which started roughly two centuries ago. Various authors, including Charles Dodgeson (alias Lewis Carrol, the author of *Alice in Wonderland*) had discussed properties of different voting systems with the aim of proving the superiority of one voting system over another.

Their interest in this line of research was motivated by the following problem. Given the various individual preferences of members of a society (or committee or collective), how do we arrive at a collective or social choice between various options? Various voting methods are obviously ways of *aggregating* individual preferences into a social decision. However, it was realized long ago that the method of majority rule, which is perhaps the most common voting method, could lead to inconsistencies of the following kind.

Example: Consider a committee $N = \{1, 2, 3\}$ which has to choose one option out of the feasible set $A = \{a, b, c\}$. Suppose that the preferences of individuals in N are as follows:

1 : a b c
2 : b c a
3 : c a b

(Here, preferences are indicated by the order from left to right; so, 1 prefers a to b to c, and so on.)

If majority rule is used to derive the committee's (binary) preference relation over pairs of alternatives, then a is preferred to b since 1 and 3 prefer a to b, and only 2 prefers b to a. Similarly, the committee must prefer b to c and c to a. Hence, there is no obvious 'best' choice for the committee since the committee's preferences *cycle* over $\{a, b, c\}$.

The birth of modern social choice theory is due to Kenneth Arrow, the Nobel Laureate for Economics in 1971. In what has been rather inappropriately named the 'General Possibility Theorem', Arrow showed that *every* 'reasonable' method of aggregating individual preferences to a social preference ordering suffered from the same problem of inconsistency associated with the majority rule. Here is a brief description of what has been called the Arrowian framework.

Let N be a finite set of individuals, with $|N| = n \geq 2$, and A be a set of feasible alternatives or options with $|A| \geq 3$, I will assume that A is finite, purely for heuristic convenience. However, much of the subsequent discussion depends crucially on the assumption that N is finite. Each individual in N has a preference ordering R_i over A. A binary relation B on a set X is called an ordering if (i) for all $x, y \in X$, xBy and yBx (the connectedness condition); and (ii) for any $x, y, z \in X$, xBy and yBz implies xBz (transitivity). Let us interpret R_i as the 'weak' preference relation 'at least as good as'. With each R_i is associated a *strict* preference relation P_i, the so-called asymmetric factor of R_i, which stands for 'strictly better than'; namely xP_iy for $x, y \in A$ if xR_iy but not yR_ix. The symmetric factor of R_i, denoted I_i, stands for the relation 'is indifferent to'. So, for any $x, y \in A$, xI_iy holds if xR_iy and yR_ix.

Bhaskar Dutta is in the Indian Statistical Institute, New Delhi 110 016, India.

Let $\Re$ be the set of all possible orderings over A.

Definition: An *Arrow social welfare function* is a mapping $f: \Re^n \rightarrow R$.

The interpretation is that an Arrow social welfare function is a procedure which maps every logically possible n-tuple of individual preferences into a social preference ordering.

It will be convenient to use the following notation in the subsequent discussion. Given any n-tuple of individual preferences $(R_1, R_2, ..., R_n)$ and an Arrow social welfare function f, let us use $R \in \Re$ to denote the social preference ordering $f(R_1, R_2, ..., R_n)$. Similarly, R' will denote the social preference ordering $f(R'_1, R'_2, ..., R'_n)$. Just as P_i is used to denote the asymmetric factor of R_i, let us use P and P' to denote the asymmetric factors R and R'.

Arrow imposed the following relatively mild conditions on f.

Condition I (Independence of irrelevant alternatives): For any two n-tuples $(R_1, R_2, ..., R_n)$ and $(R'_1, R'_2, ..., R'_n)$ in the domain of f, and any pair of options x and y such that xR_iy if and only if xR'_iy, we must have xRy if and only if $xR'y$.

Condition I says that the social preference between any pair x and y should depend only on the individual preferences between x and y.

Condition D (Non-dictatorship): There is no $i \in N$ such that for all $(R_1, R_2, ..., R_n) \in \Re^n$, for all pair x and y in A, xPy holds whenever xP_iy.

Condition D rules out the existence of an individual whose preference dictates social preference between every pair of alternatives.

Condition P (Pareto): For any $(R_1, R_2, ..., R_n) \in \Re^n$ and $x, y \in A$ if xP_iy for all i then xPy.

Arrow Impossibility Theorem: There is no Arrow social welfare function satisfying conditions I, D and P.

Remark: Notice that the restriction that A contains at least *three* elements is crucial. For if A contains only two elements, then majority rule satisfies all of Arrow's conditions. Indeed, all sensible voting rules will yield the same social preference relation in this case. The reason why the impossibility theorem requires A to contain at least three alternatives is that *transitivity* becomes a vacuous concept if there are only two elements.

This theorem posed a *formal* challenge to the possibility of arriving at consistent social decisions in a democracy. A 'democratic' government should be responsive to the common will of its citizens. Hence, it will want to evaluate various policy options from the standpoint of the 'representative' citizen, where the representative citizen's preference ordering combines or aggregates the views of all the citizens in the society. Unfortunately, one interpretation of the Arrow theorem is that the representative citizen's preferences cannot be represented as an *ordering*.

Much of modern social choice theory can be viewed as attempts to see whether consistent decision-making is possible if the original conditions are relaxed. Sen has been one of the leading researchers in this area.

For instance, it can be argued that consistent decision-making is possible even when the social preference relation is not an ordering. What is required for consistent social choice is the existence of an alternative which is the 'best' in the feasible set in terms of the social preference relation. So, let R be any social preference relation, and $A' \subseteq A$. Define the *choice set* $C(A', R) = \{x \in A' | xRy \text{ for all } y \in A'\}$. That is, the choice set consists of the pairwise-best elements in A' according to the binary relation R.

Now, consider the notion of 'quasi-transitivity', which means that the *strict* preference relation is transitive. Formally,

Definition: The binary relation R is *quasi-transitive* over A if for $x, y, z \in A$, xPy and yPz implies xPz.

Now, the set of socially optimal outcomes can be taken to be the choice set corresponding to the social preference relation R, and this set will be non-empty if R is connected and quasi-transitive.

It can be seen that given an n-tuple $(R_1, ..., R_n)$ of individual preference orderings of the n members, we can define a quasi-transitive social preference relation by the so-called Pareto extension rule: given $x, y \in A$, xPy if and only if xP_iy for all $i = 1, ..., n$.

The Pareto-extension rule declares the alternative x to be socially preferred to y only if every one in the society prefers x to y. This rule, and more generally relaxation of the requirement of transitivity of the social preference to quasi-transitivity, provides a route to escape the Arrow impossibility result. However, note that the Pareto-extension rule gives everyone a 'veto' – if any individual prefers x to y, then x is socially at least as good as y. The existence of an individual with a veto does not represent an asymmetric distribution of power whereas a dictatorship clearly means that power is distributed very unevenly. However, under rules such as the Pareto extension rule the social preferences are typically indecisive, since it requires unanimous consent for x to be declared strictly better than y. Unfortunately, it was subsequently shown that the existence of a vetoer is an *unavoidable* consequence of any attempt to resolve the Arrow problem by weakening the transitivity of social preference to quasi-transitivity.

Another route out of the impossibility result is to *restrict* the domain of the Arrow social welfare function. It can be argued that in many practical applications, all logically possible orderings cannot possibly be meaningful individual preferences. To take an economic example, suppose A is the set of all possible allocations of a fixed endowment of a certain number of goods. It makes sense to assume that individuals only care about their own component of the allocation, and that they prefer more of any good to less. These plausible assumptions can then be used to restrict the domain of the aggregation procedure.

One preference restriction which has been used in a number of applications is called *singlepeakedness*. Assume that the set of alternatives is some subset of $\Re^1$, the real line. For example, the alternatives could be different political parties, arranged from left to right in terms of their political ideology. Alternatively,. the elements of A could represent tax rates corresponding to various levels of government spending.

A preference ordering R is singlepeaked if there exists $a^* \in A$ such that for all $a, b \in A$, a^*PbPa whenever $(a^* < b < a)$ or $(a < b < a^*)$. In other words, preferences are singlepeaked if there is a most-preferred outcome a^*, and preferences decline the further away the alternative is from a^*. Now, if individual preferences are singlepeaked and if the number of individuals is odd, then the method of majority rule will yield transitive social preferences.

An alternative interpretation of the singlepeakedness is that in *every* triple of alternatives $\{x, y, z\}$, all individuals agree that some outcome is *not the worst*. For instance, if x, y, z is the order in which the alternatives are arranged, then for i for whom xR_iy, it must be the case that yP_iz. So, this condition is equivalent to demanding that for all i, xR_iy and zR_iy do not both hold simultaneously. In other words, y is not the worst outcome in $\{x, y, z\}$ for any individual i.

Sen showed that the sufficient condition for transitivity of social preferences under majority rule (for an odd number of individuals) could be weakened considerably to one of *value restriction*. This essentially requires that in any triple of alternatives $\{x, y, z\}$, there is at least one alternative, say y, such that everyone agrees that y is 'not the worst' or 'not the best' or 'not the middle'. Further, Sen showed that value restriction is sufficient to ensure that the social preference relation is quasi-transitive irrespective of whether the number of individuals is odd or even. In joint work with Prasanta Pattanaik, Sen went on to characterize the domains of preferences under which (i) the majority rule will yield transitive social preferences, and (ii) the majority rule will yield social preferences which are *acyclic*. (Acyclicity means that there is *no* set $\{x_1, ..., x_k\}$ such that $x_1P_ix_2 ... P_ix_k$ and $x_kP_ix_1$.) So if preferences are acyclic over A, then the choice set will be non-empty.

Unfortunately, subsequent work has shown that for subsets of higher dimensional Euclidean spaces these sufficient conditions for transitivity or even acyclicity are extremely stringent. This implies that the Arrow impossibility theorem cannot be evaded by seeking for plausible restrictions on individual preferences.

Perhaps, this motivated Sen and others to pursue another route of escape for the straightjacket imposed by the Arrow theorem. This is to enrich the informational base of social choice by making the social preference relation R, a function in n-tuples of individual *utility* functions, instead of as a function of n-tuples of individual orderings. The use of utility functions permits the introduction of various measurability and comparability assumptions.

Two different assumptions about the nature of individual utilities are made in economics. If individual utility is *ordinal*, then an individual can only compare the *levels* of utility (or satisfaction) associated with different alternatives. In other words, the individual's utility function can be viewed as a real-valued representation of his preference ordering. So, if u_i represents individual i's utility function, then so does *any* positive *monotonic* transformation of u_i. *Cardinal* utility is informationally richer. If utility is cardinal, then an individual can compare the difference in utility between alternatives x and y with those between z and w. So, if individual utility is *cardinal*, then any positive *affine* transformation of u_i represents the same underlying tastes and preferences.

So, the utility information which is to be used to generate a given social ordering R is not a single n-tuple of utility functions, but a *set* of n-tuples of individual utilities which are informationally identical. Of course, the analysis also has to specify the extent to which utilities are comparable *across* individuals. The measurability (that is, whether utilities are ordinal or cardinal) and comparability assumptions can be incorporated by imposing a class of *invariance* requirements which demand the same social ordering for each of the n-tuples utility functions that reflect the same underlying reality.

A *social welfare functional* specifies exactly one social ordering over A for any given n-tuple $(u_1(\cdot), ..., u_n(\cdot))$ of individual utility functions each defined over A. Let L be the set of all possible functions $u : A \rightarrow \Re^1$. So, L is the set of possible utility functions. Let $\mathcal{L} = L^n$ be the set of all possible n-tuples of utility functions. Then a social welfare functional is a mapping $F : \mathcal{L} \rightarrow \Re$.

The invariance requirement takes the general form of specifying that for any two n-tuples in the same comparability set $\bar{L}$, reflecting the assumptions of measurability and comparability, the social ordering R must be the same. I will give below *three* types of comparability sets.

First, note that the Arrow framework assumes that individual utilities are ordinal and that there is *no* possibility of interpersonal comparability. Let us call this ordinal non comparability (ONC). Second, suppose utilities are ordinal, but the *levels* of individual utilities are comparable. Ordinal level comparability (OLC) implies that statements of the following kind are meaningful – individual i in state x is better off than individual j in the state y. Finally, suppose individual utility is cardinal unit comparable (CUC); this implies that it is possible to compare the change in individual i's utility with that of individual j's in moving from x to y. Then, the comparability sets $\bar{L}$ corresponding to ONC, OLC and CUC are specified in the following definition.

Definition: For any $(u^*_1, \ldots, u^*_n) \in \bar{L}$, it is required that $\bar{L}$ consists of all n-tuples $(u_1, \ldots, u_n)$ such that $u_i = \psi_i(u^*_i)$ for all i, for some n-tuple of transformations $(\psi_1, \ldots, \psi_n)$ satisfying the following alternative restrictions.

ONC: Each ψ_i is a positive, monotonic transformation.

OLC: $\psi_i = \psi$ for all i, where ψ is a positive, monotonic transformation.

CUC: Each ψ_i is a positive, affine transformation where $\psi_i(\cdot) = a_i + b(\cdot)$ with $b > 0$, the same for all i.

In the case of ONC, the comparability set is 'large' because the absence of interpersonal comparability means that *different* transformations can be applied to each individual utility function. Moreover, since individual utilities are assumed to be ordinal, all monotonic transformations are permissible. An alternative framework would be to assume that individual utilities are cardinal but that there is no interpersonal comparability of utilities. Sen actually showed that this would not make any difference to the Arrow theorem.

Interpersonal comparability is incorporated into the framework by imposing the restriction that the comparability set only include transformed n-tuples of utility functions if the individual transformations are restricted in specific ways. So, when *levels* of utilities are comparable and utilities are assumed to be ordinal (the case of OLC), the same monotonic transformation has to be applied. This ensures level comparability since $u_i(x) > u_j(y)$ implies $\psi(u_i(x)) > \psi(u_j(y))$.

Notice that the invariance requirement ONC is the strongest. This has often been 'blamed' for precipitating the impossibility theorem. It is worth pointing out that in the presence of OLC, the Rawlsian rule of judging x to be socially at least as good as y if and only if the worst-off individual in x is at least as well-off as the worst-off individual in y, satisfies appropriate modifications of the other Arrow axioms. Similarly, CUC permits consideration of classic utilitarianism in which xRy if and only if

$$\sum_{i \in N} u_i(x) \geq \sum_{i \in N} u_i(y).$$

Various illuminating characterizations of these and other rules are possible in this richer informational framework. Amartya Sen has been the leading advocate of the use of interpersonal comparability in social welfare judgements, and has also contributed substantially to the characterization results.

A theorem of Sen which has generated a huge amount of literature is *Impossibility of the Paretian Liberal.* Sen's purpose was to show that even a very mild requirement of individual rights or liberty can conflict with the Pareto principle. The particular form of liberty is based on identifying certain types of choice as being in a person's 'protected sphere', and then allowing each person to determine the social preference over pairs of alternatives in the person's protected sphere. Define a person i as *strongly decisive* over (x, y) if xPy holds whenever xP_iy holds.

Minimal liberty condition: At least two persons are strongly decisive over one pair of social states each.

Sen went on to prove that there is no choice rule satisfying Arrow's independence of irrelevant alternatives, the Pareto condition, minimal liberty and yielding a social preference relation which is quasi-transitive. Although this result was proved as early as 1970, there are attempts even today to resolve the conflict between minimal liberty and the Pareto principle.

Finally, no account of Sen's contribution to the theory of social choice is complete without mentioning his classic book *Collective Choice and Social Welfare,* written in 1970. Even today, it serves as the best introduction to the subject.

Received 15 February 1999; accepted 22 February 1999