Cambridge University Press
978-1-107-16516-8 — Theory of Social Choice on Networks
Wynn C. Stirling
Excerpt
More Information

1 *Preference*

This is called practice, but remember to first set forth the theory.

— Leonardo da Vinci
Codex Madrid I

Ever since Galileo rolled balls down an inclined plane and realized that he could use mathematics to describe that motion, mathematical models have been central to the understanding of natural phenomena – physical, biological, and social.[1] Such models are abstract representations of real phenomena that help us understand how, but not necessarily why, the phenomena occur. Social science uses mathematical models as mechanisms for the study of selected features of human social behavior. A complex social problem is defined and factors that are deemed to be relevant are encoded into mathematical expressions, while those factors considered to be irrelevant are ignored. Such models can be used to conduct systematic investigations, test theories, simulate behavior, and evaluate performance. They can also be used to design and synthesize artificial social systems that are intended to function in ways that are compatible with human social behavior.

Since the days of Condorcet, the problem of how a collective of autonomous individuals should choose from a set of distinct and mutually exclusive alternatives has been subjected to intense mathematical modeling. There are two basic ways to address this question. Either the rationale for making the choice is a direct attribute of the collective viewed as single entity, or it is derived from the desires of the individuals by some process of aggregation. To comply with democratic principles, social choice theory has adopted the latter approach, and focuses first on the individual.

[1] "The great book of nature is written in the mathematical language, . . . without whose help it is impossible to comprehend a single word of it" (Galilei, 1623, sect. 6).

The classical way to construct a social choice model is to make only minimal assumptions about the behavior of the individuals and then investigate what can be deduced about the behavior of the collective. The behavioral assumption generally used to define a social choice model is the doctrine of individual rationality: The members of a collective are primarily (some might argue exclusively) motivated by self-interest. Philosophers may argue about the veracity of such a claim; nevertheless, individual rationality continues to be at the core of much of decision theory. Tversky and Kahneman explain why individual rationality is so dominant.

> The assumption of [individual] rationality has a favored position in economics. It is accorded all of the methodological privileges of a self-evident truth, a reasonable idealization, a tautology, and a null hypothesis. Each of these interpretations either puts the hypothesis of rational action beyond question or places the burden of proof squarely on any alternative analysis of belief and choice. The advantage of the rational model is compounded because no other theory of judgment and decision can ever match it in scope, power, and simplicity.
>
> Furthermore, the assumption of rationality is protected by a formidable set of defenses in the form of bolstering assumptions that restrict the significance of any observed violation of the model. In particular, it is commonly assumed that substantial violations of the standard model are (i) restricted to insignificant choice problems, (ii) quickly eliminated by learning, or (iii) irrelevant to economics because of the corrective function of market forces (Tversky and Kahenman, 1986, p. 89).

As a general concept, individual rationality may admit several definitions, but when used as a mathematical model, it must be given a precise operational definition in terms of some mathematically expressible concept. The simplest possible concept of self-interested behavior is that an individual prefers more to less, and manifests that preference in the form of comparative evaluations between alternatives.[2] Such binary comparisons are ordinal; they do not require specifications of intensity. They are also relative; there need be no fixed standard of performance against which the alternatives are evaluated.

[2] "Among the classical economists, such as Smith and Ricardo, rationality had the limited meaning of preferring more to less" (Arrow, 1986, p. 204).

To ensure that minimal properties of consistency are maintained, the comparison model is usually assumed to be reflexive, antisymmetric, transitive, and complete, that is, it is a *linear ordering*, denoted by the symbol $\succsim$. For any two alternatives a and a', the expression $a \succsim a'$ means that one either strictly prefers a to a' ($a \succ a'$) or is indifferent ($a \sim a'$).[3] This bare bones model is stripped of all irrelevant considerations and contextual issues. It is a model for the most elementary notion of individual rationality.

This book expands beyond the narrow confines of this classical model. It keeps some meat on the bone by not completely removing all context from an individual's preference model. It develops an expanded operational definition of individual rationality that is designed to characterize the behavior of societies where its members are interconnected by explicit social linkages. Such a collective is a *network* if it can be expressed by a graph whose vertices represent the individual members and whose edges represent the connecting linkages. In this study we focus on networks with linkages that enable the individuals to influence each other. A special case of a network is a graph with no edges – a trivial network. Since there are no explicit influence linkages, each individual is confined to consideration of its own welfare as expressed through its linear preference ordering. This egocentric structure severely limits an individual's ability to expand its sphere of interest beyond its own narrowly construed concerns.

Applying social choice theory in a nontrivial network environment suggests the need for a critical examination of the way individual preferences are expressed. Is a linear order the only mathematically precise way to express the concept of individually rational behavior? In other words, does adherence to the doctrine of individual rationality automatically imply that members of a collective care about and (at least ostensibly), only about, their own narrow self-interest without regard for the welfare of others? Or does individual rationality allow space for individuals to incorporate the interests of others into their own interests? And if the latter posit is allowed, how can such an

[3] Some social choice theorists deny that transitivity applies to indifference, and therefore focus on strict partial orders. Such distinctions, however, are not central to the topics of this treatment, and will not be pursued.

expanded notion of individual rationality be expressed operationally in an individual's preference model? Once that question is answered, the next one emerges: Can these more complex expressions of individual preference be used to define an operational notion of rational social behavior?

These are the questions that must be considered when applying social choice theory to nontrivial networks. Indeed, it is in this more structured social context that social choice theory has extended beyond the confines of social science, its traditional purview. Social models are increasingly being applied in the computer science and engineering disciplines as a means of designing and synthesizing artificial social systems. (Genesereth et al., 1986; Weiss, 1999; Parsons and Wooldridge, 2002; Russell and Norvig, 2003; Goyal, 2007; Nisan et al., 2007; Shamma, 2007; Vlassis, 2007; Jackson, 2008; Shoham and Leyton-Brown, 2009; Easley and Kleinberg, 2010). Although the social science and computer science and engineering domains are distinct, the models and mathematical techniques they use have much in common. Social science disciplines use models primarily for analysis; that is, to explain, predict, justify, and recommend human behavior. In this context, the models are idealized approximations of reality, but they are not causal. Computer science and engineering disciplines use models for synthesis; that is, to design and construct artificial entities whose behavior is governed by the models. In this context, the models are used to control behavior – they are causal. Put another way, the difference between analysis and synthesis applications is that the former uses models to reduce a reality to an abstraction, while the latter uses them to create a reality from an abstraction. This distinction is important. With analysis, simulated decisions induced by a model can be interpreted as socially motivated without actually endowing the individuals with any specific social attributes or with the ability to act situationally. Social context can then be overlaid on the mathematical model through the solution concept, that is, by the concept of aggregation that is applied. Such expressions of social behavior are exogenous – arising from some source other than the model. With synthesis, however, social attributes must be explicitly incorporated into the model and the agents must possess the ability to respond dynamically to specific social situations. Such expressions of social behavior are endogenous – arising from the model.

Cambridge University Press
978-1-107-16516-8 — Theory of Social Choice on Networks
Wynn C. Stirling
Excerpt
More Information

1.1 Categorical Preferences

In this field, there are many conflicts and many dilemmas.

— Amartya Sen
Collective Choice and Social Welfare (North-Holland, 1979)

Given an alternative set $\mathcal{A}$ and a collective $\{X_1, \ldots, X_n\}$ of individuals, classical social choice theory is based on the assumption that each X_i possesses a linear ordering over $\mathcal{A}$, denoted $\succsim_i$. A theoretically ideal approach would be to identify a *social welfare function*, an aggregation process by which an arbitrary preference profile $\{\succsim_1, \ldots, \succsim_n\}$ would generate a linear *social order* over $\mathcal{A}$, denoted $\succsim_s$, for the collective. Perhaps the first to recognize the inherent difficulties of such a pursuit was Condorcet (1785) who showed that non-transitivity can arise with social preference orderings, even though individual orderings are transitive. Subsequently, Arrow (1951) showed that such dilemmas are impossible to eliminate without violating a set of arguably desirable and reasonable properties.[4] Without doubt, the most heavily studied issue of social choice theory is how to respond to this conundrum.

As astutely noted by Shubik (1982, p. 11), “A model is defined by its boundaries.” A natural way to deal with boundary problems is to minimize the effect that un-modeled phenomena have on the model. One way to do this is simply to eliminate the need for a concept of social rationality and argue that, although it must be conceded that groups indeed do make decisions, ascribing rational behavior to a group is nothing more than an anthropomorphic trap. According to Shubik (1982, p. 124):

It may be meaningful, in a given setting, to say that a group ‘chooses’ or ‘decides’ something. It is rather less likely to be meaningful to say that the group ‘wants’ or ‘prefers’ something.

[4] These properties are *monotonicity*, the property that if the rank of an alternative changes for one individual, then the social rank of that alternative changes in the same direction or remains unchanged; *independence of irrelevant alternatives*, the property that if some alternatives are deleted from $\mathcal{A}$, then the social ranking of the remaining alternatives does not depend on the individual rankings of the deleted alternatives; *unanimity*, the property that if one alternative is ranked higher than another by all individuals, then the social ordering also preserves that ordering; and *non-dictatorship*, the property that the ranking of no single individual unilaterally induces the social ranking.

Elster (1986, p. 3) appeals to the principle of methodological individualism to assert that "there do not exist collective desires or collective beliefs. A family may, after some discussion, decide on a way of spending its income, but the decision is not based on 'its' goals and 'its' beliefs, since there are no such things." Another approach is to argue, as does Arrow (1974, p. 17), that

> A truly rational discussion of collective action in general or in specific contexts is necessarily complex, and what is even worse, it is necessarily incomplete and unresolved.

Luce and Raiffa (1957, p. 196) simply concede that adequately characterizing social rationality may be beyond the scope of individual rationality: "it may be too much to ask that any sociology be derived from the single assumption of individual rationality."

As a result of these positions, it may be tempting to conclude that individual rationality and social rationality cannot coexist. But that is not the content of the impossibility theorem, which states only that a particular model of expressing individual preference ordering is not compatible with the same preference ordering model for society. Limitations of the *model*, however, do not imply limitations of the *concept*. There is no obvious logical or intrinsic contradiction between individually rational behavior and socially rational behavior. On the contrary, there is much empirical evidence that they can and do coexist. And if in the *practice* of making social choices the two concepts can be compatible, it would seem to be important also to establish their compatibility in *theory*.

The validity of using the binary relation $\succsim$ to express individual preferences rests on the assumption that each individual's preference ordering is fixed, immutable, and acontextual – it is *categorical*. Furthermore, it is *static* and cannot adapt to a dynamic social environment as the members of the collective interact. All individuals are assumed to come to the decision problem with their preferences already defined *ex ante*, and they are impervious to making any changes for any reason. This is a strong assumption, but it is one that traditional social choice theory makes without apology. As Arrow (1951, p. 17) puts it, "It is assumed that each individual in the community has a definite ordering of all conceivable social states, in terms of their desirability to him. . . . It is simply assumed that the individual orders all social states by whatever standards he deems relevant." Friedman (1962, p. 13) even argues that one does not need to know how the preferences are formed to arrive at a solution: "The economist has little to say about the formation of wants; this is the province of the psychologist. The economist's task is to trace the consequences of any given set of wants." Johnson (1998, p. 4) also justifies reliance on this assumption: "In social choice theory, as in the broader field of rational choice, individual goals are typically taken as 'givens,' part of the data provided by a study of a particular situation. This is a practical decision, based in large part on the need to keep research projets manageable." The consequence of this practice, however, is that any history or explanation regarding the origin or justification of individual preferences is considered to be irrelevant to the decision-making process.

Since Arrow's pioneering work, many developments, extensions, and refinements have been introduced to account for social issues such as fairness, justice, welfare, resource allocation, coalition formation, etc. Also, the interdisciplinary field of computational social choice is becoming increasingly important as an interface of social science and computer science. Notwithstanding these developments, the underlying categorical preference ordering mechanism has remained unchanged. It has not, however, remained unchallenged. Sen notably threw down the gauntlet long ago:

> A person is given *one* preference ordering, and as and when the need arises this is supposed to reflect his interests, represent his welfare, summarize his idea of what should be done, and describe his actual choices and behavior. Can one preference ordering do all these things? A person thus described may be "rational" in the limited sense of revealing no inconsistencies in his choice behavior, but if he has no use for these distinctions between quite different concepts, he must be a bit of a fool. The *purely* economic man is indeed close to being a social moron. Economic theory has been much preoccupied with this rational fool decked in the glory of his *one* all-purpose ordering. To make room for the different concepts related to his behavior **we need a more elaborate structure** [italic emphasis in original, bold emphasis added] (Sen, 1977, pp. 335–336).

Sober and Wilson add to this concern by arguing that pure selfishness as an explanation of human behavior, and which is the assumption (at least implicitly) upon which reliance on categorical preferences is based, has yet to be conclusively established.

> Psychological egoism is hard to disprove, but it also is hard to prove. Even if a purely selfish explanation can be imagined for every act of helping, this doesn't mean that egoism is correct. After all, human behavior also is consistent with the contrary hypothesis—that some of our ultimate goals are altruistic. Psychologists have been working on this problem for decades and philosophers for centuries. The result, we believe, is an impasse—the problem of psychological egoism and altruism remains unsolved (Sober and Wilson, 1998, pp. 2–3).

Regardless of its structure, a preference model is simply an abstract mathematical characterization of social behavior and, as Friedman (1962, p. 13) observes: "The legitimacy of any justification for this abstraction must rest ultimately, in this case as with any other abstraction, on the light that is shed and the power to predict that is yielded by the abstraction." Since the introduction of social choice as a formal theory, much of the modeling light has been focused by the lens of narrow self-interest as expressed with categorical preference orderings. But as is true in optics, the wider the lens, the sharper can be the focus. Expanding the lens of individual rationality beyond myopic self-interest may permit greater focus, or precision, for modeling social behavior.[5] Shubik (1982, p. 1) put it this way: "The usefulness of mathematical methods – game theory or not – depends upon precision in model, and in economics as elsewhere, precise modeling implies a careful and critical selectivity." Arrow further amplified the need for careful and critical selectivity with regard to the appropriateness of relying models that are restricted to narrow self-interest.

> Rationality in application is not merely a property of the individual. Its useful and powerful implications derive from the conjunction of individual rationality and other basic concepts of neoclassical theory – equilibrium, competition, and completeness of markets. . . . When these assumptions fail, the very concept of rationality becomes threatened, because perceptions of others and, in particular, their rationality become part of one's own rationality (Arrow, 1986, p. 203).

Individually rational behavior does not prohibit an individual from incorporating the interests of others into its own self-interest – in fact, it can require it.

[5] An individual's interests are myopic if it does not take into consideration all of the material and social consequences of its choice.

Classical social choice theory is, and has been, a successful tool for the analysis of society. It is settled theory. But being settled makes it a candidate for an observation by Dewey (1938, p. 8, 9): "In scientific inquiry, the criterion of what is taken to be settled, or to be knowledge, is being *so* settled that it is available as a resource in further inquiry [emphasis in original]." A fundamental issue that deserves further inquiry is how to define a preference model structure for an expanded concept of individual rationality that allows the incorporation of the rationality of others as part of one's own rationality. In particular, an expanded model should enable the accommodation of such social concepts as coordination, cooperation, compromise, and altruism. Accounting for such concepts, however, requires the individual to define its preferences according to specific situations. The preference relations must be context dependent, allowing the individual to change its preference ordering in response to given situations. Categorical preferences simply do not provide that flexibility. The only way complex social behavior can emerge from the association of a group of individuals is if social relationships among the individuals are explicitly modeled.

Simply put, a categorical preference model is too blunt an instrument to provide the precision necessary to characterize the behavior of individuals in a network. Any new instrument must extend precision in two ways. First, it must provide an explicit mechanism by which influence can be exerted between individuals. Second, it must accommodate an expanded concept of individually rational behavior that extends beyond narrow and myopic self-interest and incorporates the interests of others as part of ones own interest. A model of individually rational behavior in a complex social context must be true to the Latin root for *social*, namely *socialis* (meaning "of companionship" or "of allies"). Companions and allies do not operate in social vacuums; they operate in context.

Arrow's impossibility theorem is simply the mathematical confirmation of a fundamental truth; namely, a group of asocial individuals cannot generate a society with complex attributes. Nevertheless, many approaches have been offered to modify or bypass Arrow's theorem, while retaining the categorical preference model structure. These approaches include restricting the domain, limiting the number of alternatives, relaxing the independence of irrelevant alternatives assumption, and substituting social choice for social preference. Another approach is to overlay the model with a veneer of psychological features such as the reputation of being "nice" that is employed by Axelrod (1984) to describe the behavior of individuals in the context of repeated play of the Prisoner's Dilemma game. Such features, however, are not part of the mathematical model; they only affect the solution concept that is either imposed on the individuals or is learned as a result of experience or as the end result of social evolution. Although such approaches may achieve limited success in an analysis context, there is no way to overlay the model with psychological features or other assumptions or constraints that are not part of the model in a synthesis context, where the goal is to design and construct artificially intelligent decision-making societies such as multiagent systems and distributed control systems that must function autonomously.

1.2 Reactive vis-à-vis Responsive Models

Homans (1961) offers three criteria for behavior to qualify as *social*. First, an individual's actions must elicit some form of reward or punishment as a result of behavior by another individual. Second, behavior toward another individual must result in reward or punishment from that individual, not just a third party. Third, the behavior must be actual behavior, not just a norm of behavior. A natural way to categorize societies is to order them according to the sophistication of their social relationships.

Social Framework I – Anarchies: Perhaps the least structured social framework is an anarchy, comprising individuals who are constrained by no sense of order or purpose for the society, each is a law unto itself with no controlling rules or principles to give order. Individuals may or may not have preferences, and even if they do, their preferences may or may not govern their behavior.

Social Framework II – Collectives: Evaluating alternatives by means of categorical preference orderings is consistent with a particular economic theory termed the *price system* (Hayek, 1945; Friedman, 1962). Prices constitute the information that guides both users and providers of products as they make decisions regarding the various transactions they undertake. Prices are attached to all products, which can be bought and sold, thereby creating an efficient and standardized