

MODELS OF RATIONAL DECISION: FORMAL CRITERIA AND NORMS

Assistant Florin Popa, Ph.D
A.S.E., Bucharest

Abstract: *The present paper discusses the relevance of decision models developed in rational choice theory and game theory for rational decision. We try to analyze the way they can be used for assessing the rationality of decision and point out some difficulties in applying them. These difficulties are linked with specifying certain formal criteria (like completeness, reflexivity, transitivity, continuity), but mostly with the structure of the decision model as such. We outline an approach which balances the role of the formal criteria for assessing the rationality of decision with the role of norms. Norms are understood as requirements of rational adequacy which are open to adjustments. The application of these requirements to specific decision cases is not formalized, and their specification (even if incomplete and subject to further adjustments) can only be made within a specific decision context.*

Keywords: *rationality, decision theory, rational choice theory, formal criteria, norms*

*

Just like other concepts which underline a whole theory, rational decision has to obey a double standard: if its definition is too narrow, many cases (empirical or hypothetical) are unaccounted for by the theory. If its definition is too wide, the concept loses its explanatory relevance.

The present paper discusses the models of rational decision developed within rational choice theory (RCT) for the analysis of real life decisions. At the same time, we aim to evaluate to what extent RCT is able to offer a coherent and unitary conception of the rationality of decision. If there are interactive decisions that cannot be represented as strategic interactions (games) with a (theoretically possible) solution, can they be considered *rational* nevertheless?

Within the *rational actor* model, which is the standard for RCT, the substance of the rational action concept is *the maximization of individual utility*. The individual can choose between alternative courses of action, taking into account individual preferences and states of nature. The rationality of decision is evaluated by taking into account in what measure the results of the decision have succeeded in maximizing utility (that is, satisfying individual preferences) under some specific contextual constraints (states of nature).

But why *individual* utility? Because if we understand by decision not just any act of choice, but a choice on the basis of preferences and available alternatives, then only the individual can choose. He is the basic unit of analysis:

“We start from the presumption that only the individual chooses, and that rational behavior, if introduced at all, can only be discussed meaningfully in terms of individual action.”¹

Only in the case of individual decision can we assume that there are consistent and well-defined preferences. The possibility of aggregating individual preferences into collective preferences is marked by serious difficulties, with the exception of groups manifesting unanimity (even in these cases, it is still to be seen if unanimity is the expression of a real convergence of preferences or the result of various constraints). But the majority of textbook or real life cases exhibit diverging or conflicting preferences and interests.

The problem of aggregating individual into collective preferences has a classic illustration: Arrow’s theorem of impossibility (or Arrow’s paradox). The demonstration of the theorem has been offered by Kenneth Arrow in his PhD thesis, but its popularity is linked to the publication of his book *Social Choice and Individual Values* in 1951. Arrow sets to demonstrate the impossibility of developing a model of collective decision which would respect a number of rationality criteria.

Arrow’s theorem tries to link the individual orderings of a common set of preferences with a social ordering that takes into account all individual orderings (called a social choice function), which would offer an adequate method of aggregating individual preferences.

This function would have to respect the following criteria, which Arrow regards as self-evident:

1. Non-imposition – each possible ordering of preferences at collective level can be obtained on the basis of a certain combination of individual preferences. There are no possible collective results that cannot be obtained somehow.

2. Unrestricted domain - every collective ordering of preferences should be based on a set of individual preferences taking into account all members of a given group. This assumes that each member has the ability and will to make a complete ordering of preferences, and that the aggregation procedure will take all orderings into account.

3. Independence of irrelevant alternatives – the social ordering of alternatives x and y is a function of individual orderings of x and y, independently of other irrelevant alternatives (say, of x and z). If we apply a social choice function to a subset of alternatives, than the result should be consistent with the result we would obtain if we applied the function to the whole set. Possible modifications of individual preferences which entail irrelevant alternatives should not have an impact of the social ordering of the given subset.

4. Positive association of social and individual values – if a certain member of the group changes his/her preferences – say, giving priority to y over x rather than the other way around – than the social ordering should reflect this change (by giving priority to y over x or by not changing the global ordering, but in no case by giving priority to x over y).

5. Non-dictatorship – the social choice function cannot reflect the preference ordering of a single member of the group. It has to reflect the individual orderings of a multitude of members, each ordering having equal value.

Arrow’s theorem states that, in all choice situations with at least two members and three alternatives, it is impossible to define a social choice function fulfilling

¹ J. Buchanan & G. Tullock, (1999), 3.4.4

simultaneously all five criteria. In particular, it is demonstrated that if we build a social choice function which fulfills the first four criteria, than there is at least a preference ordering contradicting the last criterion (non-dictatorship).

There are ways of getting around the paradox, although their success is disputed. An alternative is to weaken or eliminate one or several criteria. This way, a consistent social choice function can be built. However, Arrow considered the five criteria *minimal conditions* of rational aggregation of individual into social preferences. For instance, we can weaken the last criterion, assuming that individual orderings of preferences can be determined by a single member of the group, but then the function cannot be called *social* anymore, since it is not based on a multitude of *independent* individual orderings. On the other hand, if we restrict the domain of the social choice function (eliminate the second criterion) assuming a „single peaked preference” (a predetermined ordering of preferences by reference to which all individual preferences are evaluated), then majority rule can meet all of Arrow’s criteria. But it is not clear how individual orderings of preferences could be represented on a single scale without retorting to the use of a common denominator (as in Mill’s or Bentham’s concept of utility). Other authors have proposed a weaker version of the criterion concerning the independence of irrelevant alternatives, arguing that its strong version rules out most real-life contexts of decision. In particular, they propose replacing transitivity with acyclicity (if $a > b$ and $b > c$, then it is not the case that $c > a$). All this proposals discuss the status of one or more formal criteria of rational adequacy, without debating if formal criteria alone are sufficient in drawing the borders of rational decision-making.

What is the relevance of Arrow’s theorem? First of all, it shows that there is no „right” or „correct” method of aggregating individual preferences (be it majority rule or any other) into collective preferences, if we accept the necessity of the five criteria mentioned above. Its relevance extends to any procedure of collective decision, including voting systems. This has been sometimes expressed in simplifying formulas, like „no voting method is completely fair”. Secondly, Arrow’s paradox shows the difficulties of any attempt to evaluate the rationality of decisions *exclusively* on the basis on formal criteria.

The rejection of the whole idea of *collective interest* (as a separate and autonomous type of interest, above the variety of individual interests) and *collective preference* does not have to do only with the difficulties of aggregating individual preferences, but also with the analysis of individual behavior within groups. The study of real life cases leads to some counter-intuitive conclusions, that cast doubts on the hypothesis of the rational maximizing actor. Discussing collective actions which offer collective benefits, Mancur Olson considers that:

“...it is *not* in fact true that the idea that groups will act in their self-interest follows logically from the premise of rational and self-interested behavior. It does *not* follow, because all of the individuals in a group would gain if they achieved their group objective, that they would act to achieve that objective, even if they were all rational and self-interested. Indeed unless the number of individuals in a group is quite small, or unless there is coercion or some other special device to make individuals act in their common interest, *rational, self-interested individuals will not act to achieve their common or group interests.*”²

² M. Olson, (1965), p. 2

What makes individuals depart from self-interest in situations in which, according to the rational actor approach, they would have obvious reasons not to do so? A preliminary answer, which will be developed subsequently, is that the rational actor assumes a model of rationality based on applying *formal criteria* (which do not make reference to the nature or content of individual preferences, or to specific features of decision contexts). But such a model cannot rely exclusively on formal criteria: it has to take into account *norms of rational adequacy* and *contextual constraints*.

Actions, preferences, states of nature and outcomes – these are the basic elements of any decision theory. Methodological individualism is an essential assumption (in rational choice theory but not only) – for the decision maker, the external world (including other agents whose decisions and actions may influence his own decisions) represent – as far as the rationality of decision is concerned – a set of opportunities and constraints.

However, there is a domain in which both individual decision and collective outcomes are essentially influenced by other people's choices. Individuals with partly conflicting interests interact, choosing between strategies and sets of strategies. Instead of assuming a rational actor being faced with an 'impersonal' and essentialized reality, reduced to opportunities and constraints, game theory deals with strategic interactions between interdependent agents. In this perspective, *the other* is invested with the same rationality features as *myself*; consequently, his decisional procedures became transparent (although by no means predictable): I can, in principle, evaluate his alternative strategies and their respective consequences.

Many classic developments and applications of game theory deal with economics. If in neoclassical economics the analysis focuses on rational decision under certain market conditions (contextual constraints), game theory analyses direct interactions - that do not need an impersonal mechanism for aggregating interests and allocating resources - as the market. In the first case, rational choice involves maximizing individual utility under certain constraints. In the second case, priority is given to maximizing the utilities of a group of interacting individuals. This opens interesting possibilities for analyzing, with game theory instruments, the relation between individual and collective rationality.

The models of rational decision influenced by game theory are based on a concept of rationality which sets criteria regarding the consistency and coherence of individual behavior, not regarding its content. It is about *how*, not about *what*. The model includes only requirements about how the agent is managing his preferences or about the relation between preferences, estimated benefits and decisions, not about their specific content.

These axiomatic models were largely developed as part of studies on consumer behavior. The analysis of types of constraints which determine consumer choices is relevant for any decision process, since the structure of decision-making is the same, no matter what sort of preferences or benefits we are dealing with.

Formally, individual decision can be regarded as rational if the ordering of preferences is:

- Complete – the agent is able to define and order his preferences for each possible pair of elements from the list of alternatives,

For any $\{x,y\}$, $x \geq y$ or $y \geq x$ ³

- Reflexive – for any x , $x \geq x$ (there can be no strict preference of x over x)
- Transitive – « x is preferred to y » and « y is preferred to z » imply « x is preferred to z »

For any $\{x, y, z\}$, $x \geq y$ and $y \geq z$ imply $x \geq z$
and

- Continuous – if $x \geq y$ and z is close enough to x , then $z \geq y$.

Does this concept of formal rationality succeed in offering a better description of how individuals take decisions and does it manage to include borderline cases which, from the perspective of *rational actor* models, are discarded as irrational?

Before discussing the adequacy of these criteria, a brief note: when discussing the rationality of decision, we refer to practical rationality (arguments which determine our decisions) and not simply theoretical rationality (arguments having to do with our opinions and beliefs). This puts things in a different perspective: it would be an error to consider practical reasoning a species or extension of theoretical reasoning. It would also be an error to consider that our preferences, interests or intentions are related to our decisions just like premises are related to our conclusions in theoretical rationality. In other words, if:

I believe p and if p then q – this implies I believe q

But in the case of a practical reason that deals with *adequacy of means to ends*:

I want that p and if p then q – this does not imply that I want q .

The reasons why theoretical reasoning and practical reasoning operate differently will be developed in the following paragraphs. They have to do with what John Searle calls *the gap* (or interval) between the set of motivations, opinions and beliefs which represent the antecedents of a decision and the actual decision-making.

“The gap can be given two equivalent descriptions, one forward-looking, one backward. Forward: the gap is that feature of our conscious decision making and acting where we sense alternative future decisions and actions as causally open to us. Backward: the gap is that feature of conscious decision making and acting whereby the reasons preceding the decisions and the actions are not experienced by the agent as setting causally sufficient conditions for the decisions and actions.”⁴

This gap is *par excellence* the space of practical rationality – which entails *conciliation, adjustment and compromise between diverging or conflicting interests and motivations*:

“This points to the following conclusion: even if we confine our discussion of practical reasoning to means-ends cases, it turns out that practical reason essentially involves the adjudication of conflicting desires and other sorts of conflicting motivations (i.e., factitives with upward direction -of-fit) in a way that theoretical reason does not essentially involve the adjudication of conflicting beliefs.”⁵

³ $>$ amounts to *strong preference*: x is strictly preferred to y , meaning that x is not indifferent in relation to y and y is not preferred to x

\geq amounts to *weak preference*: x is preferred to y or is indifferent in relation to y

\sim amounts to *indifference*: x can be replaced with y and y with x , having equivalent utilities

⁴ J. Searle, (2002), p. 62.

⁵ J. Searle, (2002), p. 253.

Let us come back to the criteria of rational adequacy of the classical model: completeness, reflexivity, transitivity, continuity. Can they be applied non-ambiguously and consistently? We will discuss two of them.

I. Completeness

Do we have complete orderings of individual preferences in most real-life cases? Would it be possible to have complete orderings, in principle? This would imply that, given the set of available options $\{x,y,z,\dots\}$, an agent would be in principle able to establish a relation of $x \geq y$ or $y \geq x$ between any other two elements of the set. However:

a. The majority of real-life decision contexts present some sort of difficulties in processing information: the agent can be indifferent between two options or can be unable to order them – not because they have equivalent utilities, but because each of them ranks higher on certain criteria of utility and lower on other, and there is no straightforward way of ordering the utility criteria themselves. Complex decisions involve large quantities of information and available options, under limited capacities of processing. The agent resorts to processing shortcuts and heuristics. The model of *bounded rationality* assumes that preferences are not mechanically determined by individual tastes or interests that remain constant, but the result of an interaction between tastes and interests, contextual constraints and norms. It is an open-ended process, where preferences not only influence decisions, but are also influenced (adjusted) by the decision context.

This also means that the agent can obtain slightly different orderings of preferences, in cases where the contextual constraints on decisions are radically different.

b. Apart from constraints on information collection and processing, our preferences – unlike opinions – can be divergent or conflicting, and this does not entail irrationality. Firstly, there are various situations in which individual preferences can be conflicting. For instance, this moment I can wish to continue working on this paper, but I can also desire to go hiking in the mountains. These are conflicting desires, because following one of them means abandoning the other. I wish to finish the paper, because I have made a commitment (to myself and to others), because I have a deadline to respect, because I am interested in the subject or I have the feeling I am inspired. At the same time, I wish to go to the mountains because it is a beautiful autumn day, I have not been there in a while etc. Both wishes have strong motivations behind them. I will probably decide to continue working on the paper, meaning that the first preference will prevail. But it is important to point out that this decision was taken (just like many others) in the context of several conflicting preferences.

Secondly, even if we start from a logically consistent set of preferences or opinions, „secondary“ preferences (those obtained as a result of practical reasoning) *can* be inconsistent (Searle, pp. 248-252) *within the same context* (in other words, assuming that the ordering of preferences does not change).

The capacity of obtaining a complete ordering of individual preferences is limited to a minority of cases, most of which are voluntarily simplified. This reduces significantly their practical relevance. Apart from the previously mentioned problems, the difficulty of obtaining complete orderings has to do with the basic assumptions of utility theory – especially the attempt to generalize the way individuals manage their preferences related to purchase and consumption to the whole range of human preferences.

II. Reflexivity

We have already pointed out that preferences and preference orderings are not simply *given*; they are influenced to some degree by contextual constraints and can modify.

a. individuals „manage” their preferences (have an active role in their development or adjustment, can avoid or suppress them, can confound them etc). This seems to contradict the reflexivity requirement, which involves stability of preferences.

The modification of preferences occurs in many economic contexts. Some cases in which individuals exhibited an apparently irrational behavior could be nevertheless integrated in the general model of economic rationality, adding the assumption that preferences do modify in time. The standard model of rational choice theory assumes that individual preferences and tastes are relatively constant, and that agents have a complete knowledge of them. But modifications of preferences can be explained, for instance, as gradual accumulation of knowledge and expertise. Attending a arts history course, a wine tasting workshop or going to concerts or exhibitions *can* modify individual preferences, because the individual becomes more efficient, more „productive” in generating utility out of certain activities.

On the other hand, the modification of preferences can be minimal or inexistent, but their individual ordering can change according to specific accumulations of expertise. Technological change can also influence preference orderings: for example, the large-scale use of online payments for goods or services shifts the attention of consumers to those categories of goods and services which are more likely to be purchased online (electronic books and audio, access to online libraries and databases etc).

b. since we are talking about preferences and tastes, not opinions, p and non-p can coexist. We can have conflicting preferences without being irrational. The fact that there are *some* reasons why I prefer finishing this paper and *other* reasons why the perspective of a mountain trip is clearly preferable simply shows that, in most real-life decisions, there is no common standard of evaluating the utility of each alternative. Moreover, the complexity of the decision context can be too high to consider every possible alternative, evaluate their utility and order them.

It is true that the agent can form erroneous practical reasons or can act irrationally on the basis of conflicting preferences. But the simple existence of conflicting preferences does not necessarily imply irrationality.

What is the relevance of these difficulties in applying formal criteria to individual decisions? Is there a problem of choosing certain criteria (and thus assuming that other criteria would operate better) or is there a problem with applying formal criteria *in principle* to evaluate the rationality of decisions? Most probably none. These difficulties do not deny the role of formal criteria in rational decision-making, but specify their status and their limits of applicability. It specifies their status – in the sense that any set of formal criteria will be, by itself, insufficient in order to evaluate the rationality of decisions. And it specifies their limits of applicability – by „relaxing” the condition of universal applicability: it is rational to assume completeness (processing all relevant information and alternatives), but taking into account constraints that have to do with the limits of cognitive capacities and external constraints. It seems not only unrealistic but also unreasonable to assume context-independent completeness, be it only because some alternatives are clearly irrelevant.

The first question would be if not respecting formal criteria of rational adequacy implies irrationality. The answer is affirmative: no matter what the decision context is, logical inconsistency, inadequacy of means to ends or voluntary ignorance of relevant information are signs of irrationality, because „Constraints of rationality are universal

and built into the structure of mind and language, specifically into the structures of intentionality and speech acts.”⁶

Evaluating pros and cons in decision-making, or the alternative courses of action for achieving a certain aim, we are already subjecting ourselves to implicit rational constraints. Respecting formal criteria is a *necessary* condition for the rationality of decision.

The second question would be if formal criteria are *sufficient* in evaluating the rationality of decision. Let us begin by noticing that the evaluation of decision rationality can imply two approaches:

- A judgement on the adequacy of means to ends (internal adequacy); „in this respect – says Newton-Smith – rationality, which is sometimes called instrumental rationality, and explanatory success do not depend on the reasonableness of aim, nor on the truth or falsity of the beliefs involved, nor on their being reasonable or not.”⁷ This is what Newton-Smith calls a *minirat* evaluation of rationality. On the basis of *minirat*, we can evaluate the reasonableness of a decision in its own context – that is, without reference to the truth or reasonableness of its aim or of the beliefs guiding it.

- At the same time, we can adopt a *maxirat* approach, „which involves a positive justification of the given aim and/or beliefs”. Newton-Smith goes on saying: „considering a belief as irrational, we claim that the individual failed to make rational steps in order to obtain reliable data or has not ensured the reliability of data in a satisfactory way. Moreover, when applying the term „rational” to beliefs, we also apply it to aims. To consider that an aim is rational is to put in balance pros and cons of the agent adopting this aim”.⁸

Searle’s universal *constraints of rationality* cannot be used as a failproof guide to evaluating the rationality of actual decisions. All we can say is that when they are not respected our decisions cannot be called rational.

The *minirat* approach is bound to take into account the context of decision in the evaluation of its rationality. We may well have situations in which the formal criteria of rational adequacy are respected, but whose overall rationality is problematic when considered in a larger context. A terrorist can justify violent actions towards a certain group by making reference to specific context (previous violent acts committed by that group, unjustified privileges, exploitation). The adequacy of means to ends is out of the question, yet it would be hard to qualify these actions as rational. Their irrationality has to do with the inevitably limited frame of reference considered by the agent, who does not evaluate the reasonableness on his aims in the wider context of ethically and politically reasonable aims and taking into account all the relevant contextual constraints. In other words, there can be no *local* rationality (as there can be no local morality). An overall evaluation of rationality cannot be confined to means-ends adequacy.

In the case of *maxirat* approach, our aims and beliefs cannot be evaluated by applying formal criteria, because they represent preconditions of decision-making, while formal criteria apply when evaluating decision-making itself. We cannot evaluate the rationality of aims by appealing to consistency or completeness of reasons or preferences behind it. Selecting an aim means appealing to values and norms, which are substantial (they make reference to content, to the reasonableness of *what* is being valued or preferred and *why*) – unlike formal criteria.

⁶ J. Searle, (2002), p. xiv

⁷ W. H. Newton-Smith, (1984), p. 297.

⁸ Ibidem, p. 299.

Therefore, respecting formal criteria is a necessary but not sufficient condition for the rationality of decisions.

The fact that formal criteria intertwine with different types of constraints in shaping decision-making does not imply relativism. The only questionable aspect here is the assumption that preferences are by default well defined, constant and context-independent. For instance, bounded rationality models are based on the assumption of limited capacities for obtaining relevant information and adequately processing it. Any rational behavior is, in this perspective, behavior under constraints. These constraints do not limit individual rational capacities, but help specify or define the actual limits of rational decision and action. Herbert A. Simon pointed out that models of rational decision should not focus exclusively on finding optimal strategies, but should also include psychological and cognitive features of real decision-makers, that limit their information-processing and problem-solving abilities. The rational agent is always deciding and acting in an environment which poses specific contextual constraints on actual decision-making.

Therefore, decision-making is not all about maximizing utility or choosing the optimal means for a given end; more likely, it involves adjudication between diverging and sometimes conflicting desires, interests, preferences and aims. The reasonability of a given decision is to be judged according to specific contextual constraints, not only according to pre-defined decision criteria. Instead of finding the optimal strategy, bounded rationality is concerned with finding and selecting *reasonable* options.

A frequent criticism regarding bounded rationality is that it apparently implies relativism. Since any evaluation of practical rationality should take into account local constraints, does it make sense to speak of rationality after all? Do we still have a common understanding of rationality underlying human decision and action, or just local, contextual rationalities?

To a large degree, this criticism is based on a misunderstanding of the basic assumptions of bounded rationality. What this model opposes is the assumption of universal and context-independent applicability of a predefined set of rationality criteria. However, it does not oppose the necessity of having a minimal set of criteria with general applicability, nor the relevance of criteria in general in evaluating the rationality of decisions.

The exclusive choice between a concept of rationality based exclusively on applying context-independent criteria and radical relativism is misleading. The first option cannot account for a large number of decision cases which do not follow a given set of criteria but nevertheless cannot be discarded as simply irrational. The second choice makes the whole rational/irrational distinction irrelevant, since it can justify in principle most decisions on contextual grounds, without appealing to an inter-subjectively testable and generalizable procedure.

We have already stated that developing a model for evaluating the rationality of decision means, among others, specifying formal criteria, which do not say anything about the content of individual interests, preferences, utility functions or aims. On the other hand, we attempted to show that applying a set of formal criteria is a necessary but not sufficient condition of rational adequacy. An overall evaluation of the rationality of decision involves an evaluation of aims, as well as beliefs, interests and preferences.

If applying formal requirements of rational adequacy is not sufficient in order to evaluate the rationality of actual decisions, it is possible to develop a functional model of rational decision? If it is, what exactly would ensure its functionality, apart from the specification of formal criteria?

We can approach a preliminary answer by noticing that in practice there are very few situations in which we cannot establish the irrationality of decisions. In most real life situations we can offer grounds for considering a certain decision irrational, although it is less likely that we can establish its rationality. This is because the procedures for establishing irrationality or rationality of decisions are different: the former is based on necessary and sufficient conditions (for instance, contradictory beliefs or opinions), while the latter cannot be reduced to a set of predefined conditions or rules.

Rationality is not entirely about respecting formal criteria or about integrating contextual constraints into the decision-making process. It also involves the application of *norms*.

Norms can be characterized by reference to concepts such as *adaptive rationality*⁹ and *communicative rationality*¹⁰. This perspective does away with the idea of a fully-determined concept of rationality which can be readily applied to decisions and actions without being influenced by the context of its application. Just like knowledge, rationality develops in time through adjustments and corrections, taking into account feedback from actual decisions, situational constraints, as well as consistency and integration with other relevant knowledge.

The subject-centered model of rationality (which is well illustrated by Descartes' ego that can independently define and apply rationality, without reference to other egos) is discarded. Rationality is essentially social; this is not to say that rationality is simply a social construct, but that it involves communication, it cannot be conceived of outside social interaction. Rationality is potentially present in any act of communication. Individuals engaged in an undistorted dialogue not only apply rationality, but also help in defining and structuring it in the very process of interaction. Habermas considers that the ultimate basis of validity for the rationality of our decisions and actions is not the correspondence with predefined rules and procedures, but the success of interaction. The communicative action cannot be reduced to instrumental action (choosing means for given ends) and practical problems cannot be reduced to technical problems. The requirements of rational adequacy are not *given*; they are largely generated as part of the communicative effort of a community, as part of an *emerging agreement* based on a common (or convergent) interpretation and understanding of the decision context.

Norms are not completely explicit and they are not necessarily applied on the basis of formal procedures. Moreover and more importantly, norms are not completely and rigidly specified; they are open to adjustments and corrections. While criteria refer exclusively to means and ends (they are instrumental), norms appeal to commonly accepted values and standards. This common acceptance is not definitive and may well be partial; it is the preliminary result of successful communication.

This distinction between a model of rationality based on criteria and a model of rationality based on norms evokes the Weberian distinction between formal and substantial rationality. The former refers to decisions and actions on the basis of a utility-maximizing procedure of achieving predefined ends; the latter is reasonable action oriented by values.

⁹ J. March in Elster J., (1986), p. 146.

¹⁰ See J. Habermas, (2000).

Unlike the former, which can be associated with a top-bottom approach (the rules of operation are predefined and can generate solutions on the basis of an algorithm – a formalized procedure), *rationality based on norms* appeals to experience accumulated in time and can be associated with a bottom-up approach. The latter is characteristic for adaptive systems in which operation rules are continuously subjected to local adjustments. Rules of rational adequacy act upon input data (characteristics of the decision context) but at the same time are adjusted by the output data of their own action (the feedback from the results of decisions).

We can now come back to the issue of applying formal criteria in the evaluation of decision rationality. We have asked whether the difficulties in applying criteria of completeness, reflexivity, transitivity and continuity have to do with the criteria themselves *or* the status of formal criteria in the evaluation of decision rationality *in principle*. We have conceded that formal criteria cannot offer a complete specification of rational adequacy, that rationality cannot be described only in terms of following a set of rules, no matter what those rules are. Norms offer a way forward, by taking into account the context of the social interaction (the emergent agreement of the community), cognitive limits on information processing and decision-making, as well as feedback from past decisions. Norms cannot be specified outside a social and situational context, but this does not imply relativism; their formulation is always preliminary and open to adjustment, but they are not context-dependent, insofar as they remain open to further interaction with different or larger communities – and therefore confrontation with different or diverging norms.

The perspective proposed here assumes an open-ended concept of rationality, based on a „core” of necessary criteria of rational adequacy and a set of norms subjected to adaptive adjustments. It discards the idea of subject-oriented and context-independent rationality, while it keeps the requirement of trans-individual validity and grounds it in the success of communicative action.

REFERENCES

- [1] Buchanan, J. M. & Tullock, G., (1999), *The Calculus of Consent: Logical Foundations of Constitutional Democracy*, Liberty Fund, Inc. Library of Economics and Liberty.
- [2] Elster, J. (ed.), (1986), *Rational Choice*, Oxford, Basil Blackwell
- [3] Fudenberg, D. & Tirole, J., (1998), *The Theory of Learning in Games*, Cambridge, MIT Press.
- [4] Habermas, J., (2000), *Discursul filosofic al modernității. 12 prelegeri*, All Educational, București.
- [5] McCain, R., (2003), *Game Theory: A Non-Technical Introduction to the Analysis of Strategy*, South-Western College Pub.
- [6] Newton-Smith, W., H., (1984), *Raționalitatea științei*, București, Editura Științifică.
- [7] Olson, M., (1965), *The Logic of Collective Action*, Harvard University Press.
- [8] Searle, J., (2002), *Rationality in Action*, Cambridge, Mass., MIT Press.
- [9] Stevenson, W., J., (1998), *Introduction to Management Science*, McGraw-Hill/Irwin, 1st edition.
- [10] Von Neumann, J. & Morgenstern, O., (1943), *The Theory of Games and Economic Behavior*, Princeton, Princeton University Press.

