Bounded Rationality and Philosophy of Science

J. Francisco Álvarez. UNED. Madrid

12th Congress DLMPS, Oviedo 2003

The standard economic theory and specially the rational choice model are blocks shared as a theoretical background by a lot of works in philosophy of science and philosophy of technology. The logical approach to science studies is another reason of that common style in the received view. One of the issues that I mean that needs changes is the model of human being on wich those theories are built up.

There are a huge number of critics to the basis of the rational standard model, and some ot them are very well knowed. I think that Amartya Kumar Sen has developed an approach to economic studies that in its critics to rational economic man is relevant to general studies of science. Some sociologist and psycologist worked out also very pregnant remarks that is changing that standard model.

Some of the main critics to the model are related with the existence and unicity problem. Other are linked with empirical and experimental inadecuacies of the same model. In this area usually are referred the Kahneman and Tversky point of view.

Inside Psycology and Economics studies this is a common debate and there are a lot of important results. Even there are several trends that are building theoretical point of views against the received or mainstream theory. Although inside philosophyof science is not so usual this kind of critics there are already a few of scholars that are working hard on this direction. Only as an example, I referred to Adam Morton (Fools ape angels: epistemology for finite agents (For Gardenfors e-schrift) or Dov M. Gabbay and John Woods (Normative Models of Rational Agency, February 2002).

As Marcelo Dascal has said: "If extant theories have been unable to harmonize a normative rationality with description of the facts of scientific praxis, it is because they have not been able to bring them closer to each other".

Some sociologist have remarked that we need an axiological and cognitivist notion of rationality “to avoid the Charybdis of the irrational models and the Scylla of the narrow versions of rationality the Rational Choice Model endorses” (BOUDON, R.,

2001). This new notion of rationality leads to delineate a rather different models of human being than those who arised from instrumental or consecuencialist rationality. In my opinion, there are very important links between procedural rationality (Simon´s view) and the axiological cognitivist rationality (Boudon´s approach). A lot of decision procedures could be understood as frugal and simple mechanism that we use to act putting our values in action; our values are giving us reasons to act. We are instrumentally rational agents, but we also exhibit axiological rationality. These are two different notions of rationality but we act in a single communicative situation with both of them. I would like to talk about the fabric of rationality, where the expressive or axiological rationality appears as the warp and instrumental rationality as the weft of it.

Our space of values appears framing and filtering the context and signalizing towards a pragmatic notion of rationality (a synthetic notion) that arises as a minimal condition inside which the interactions may be possible. A first step, an empirical one, could be to delineate the boundaries of the space of values that participants trying to fulfil. Their target is not, or not only, to optimize some singular variable (such as truth, rhetoric force or consistence); but to try to satisface a set of values or cues that they regard as important: even own authorship or agency could be one of these values.

The main idea is that some features of the context could generate rules. Usually we are prone to ascribe these rules to cognitive capabilities of participants but those rule maybe the output of the relationships themselves. We don´t need to suppose Olympic participants with absolute and common knowledge (each one knows what the other one knows) neither in our human interactions nor in human-computer relationships; all we need is some blood and flesh human beings in contextual interactions.

These real agents could not be blurred, they always must remain at least as a parameter of the interaction. In the standard view of rationality, the olympic agents

could be eliminated because every agent is similar to each other; as every one is an "epistemic god", no one is necessary. We upgrade, as it were, to some kind of popperian third world where we can reach objective knowledge. On the contrary, I mean that we always need a concrete agent: the objectivity is not the view from nowhere, is the view from somewhere (as Amartya Sen said, (SEN, A. K., 1993)). We cannot eliminate the particular agent, we always need it at least as a parametric reference. We are rational but less than gods.

The attempts to build a purely positive economics always require a minimal commitment to an idea of human being, apparently disconnected from ethics (as a result of the pressure of the facts/values dichotomy, the naturalist fallacy, and other related issues). But the very notion of rationality involves a model of human being, whether an active agent or a mere puppet of social trends, or pure role-player with role as the normative expectations (HOLLIS, M., 1998), and no economic theory, no matter if it is said to be "positive" or not, can avoid the moral commitments involved.

When economists have proposed other ideas against both single utility and optimization, they have mainly unfolded two different views. As Selten said: "One way to model limited search without giving up the ideal of optimization is known as optimization with decision cost taken in account, also referred to as *optimization under constraints*"(GIGERENZER, G. y R. SELTEN, 2001). The other option has been "models of bounded rationality use fast and frugal stopping rules for search that do not involve optimization". The first models "become even less psychologically plausible". ("The knowledge and the computations involved can be so massive that one is forced to assume that ordinary people have the computational capabilities and statistical software of econometricians"). I mean that some movements in information theory show similar drifts (for instance, the optimization foraging information theory).

Another, more radical, option is linked to Herbert Simon´s idea of bounded rationality. Simon used the metaphor of a pair of scissors. Although I prefer the loom and weave metaphor. In Simon´s one of the blades is the "cognitive limitations" of human beings and the other one is the "structure of the environment". Cognitive rationality and Ecological rationality as Gigerenzer calls them nowadays. The very importante thing is that "Minds with limited time, knowledge, and other resources can be nevertheless successful by exploiting structures in their environments"(Selten).

To increase the complexity of a task does not imply a corresponding complexity of individuals. Sometimes a better comprehension of the environment could do the task. A system of relationships sometimes allows that the adoption of some fast and frugal mechanism produces better results than those supposed by optimal rationality that appears with a high computational complexity.

The cost-benefit analysis has been a common feature in information theory, maybe that new approaches as foraging information theory (PIROLLI, P. y S. CARD) could be more relevant from a philosophical point of view as soon as it is built upon a new kind of human being model. Interactive information systems have been developed from cognitive and computing science and are enclosed in user time optimization: "an attempt to increase relevant information gained per unit time expended". But "for task analysis, design exploitation, and evaluation of information systems, a more developed theory is needed". My claim is that although Pirolli himself says: "A more successful hypothesis about humans is that they exhibit bounded rationality or make choices based on satisficing", it is not a good step to adopt that "satisfying (satisficing in Simon´s terms) can often be characterized as localized optimization".

It is much the same thing with the so-called *economics of science*: we apply economic models to analyse scientific activity, widening the economic patterns of

action, but we cannot avoid reflecting on the individual behind scientific constructions either. That is the case as well when economics is applied to ethics: on the one hand, ethical concepts are often refined or improved; on the other, sometimes we run into slippery slope conclusions as a result of the restricted scope of the models of human being applied. Some current paradoxes and dilemmas most often show the informational weakness of the models, more than the paradoxical nature of the modelled subject.

In short, I propose to study different models of human being, since we will obtain on our choice both different theories and different conclusions. Against Laudan´s normative naturalism(LAUDAN, L., 1998), I argue for a naturalized normativism(ÁLVAREZ Á., J. F., 1999). Our sets of shared values, our theoretical commitments (always subject to debate) are the starting point for our models.

There are some contributions on this topic, I like named it as a second order realism. Game-theoretic models but with individuals that exhibit bounded rationality could be an improvement to this understanding. For example, Osborne and Rubinstein in "Games with procedurally rational players" (*The American Economic Review*, sep. 98, pp. 834-79).

Some problems arise in other fields. A huge part of contributions economist of science have made to renewal some epistemological topics does share a model of human being originated in economics. Whichever it is our conception of the various duties entrusted to the contemporary philosophy of the social sciences, even when we adopt a second order attitude, we must be conscious of the model of human being assumed, in order to avoid the unnoticed assumption of the one involved in the particular science we deal with (or we are inspired by). As it happens, this is quite often the case among philosophers of economics and economist themselves.

Understanding the aims of science and its nature, like a variety of human action requires an explicit discussion of the model of human being involved in its analysis.

Even if most of the usual critiques addresses to rational choice theory (RCT) cannot stand up to a thorough philosophical scrutiny, however, it is true that some of these considerations have contributed to improved current assessment of RCT. Furthermore, authors such as Martin Hollis, Shaun Hargreaves Heap or Ken Binmore have provided us with a bunch of reflections on economics models of human being that do not simply discredit RCT. I think that philosophy of economics can go further in assessing these models if the point of view of sceptic *realism* is assumed including the idea of positional objectivity (A. Sen). My proposal is related to the modelling of imperfect rationality, not a simple move to escape from RCT difficulties, but a serious attempt to develop a more accurate representation of the informational structure of real rationality (See James Johnson, "How not criticize rational choice theory. Pathologies of 'common sense'",*Philosophy of social sciences* 26:1 (1996), pp. 77-91).

How much should we trust in those expert systems based on a restricted conception of rationality that are shaping, to a great extent, our society? Our degree of confidence is certainly a result of our intelligence, which is not only related to science, but also to a reflection on science itself and on other projects originated in our expressive rationality.

It is quite obvious that we could not live in any advanced society if we were perfectly rational: we need a certain degree of confidence in its ruling expert systems (vid. A. Giddens). That degree of confidence is therefore a criterion in our own decision-making processes, and so we may speak of bounded rationality, in Simon´s sense, as long as our actions are made easier indeed by social norms. Let´s recall that to accept and obey those *nomoi* was an advice of ancient scepticism.

Cross-fertilisation between rational-actor model and sociological role-theory, as proposed by Boudon or by Hollis, could shift our present conception of rationality towards a proper understanding of human action (including scientific cognition). Martin Hollis, for instance, seems to share this opinion: "In sum, institutional power is a real feature of social rules, and rational actors allow for it in their calculations [...] we shall not know how the people win or lose it until we have at last connected the loyal role-player with the rational actor. Autobiographies are incontrovertible evidence that the problem is solved in practice. They are written in the active voice, not in the passive, and are suffused with the authors´conviction that it matters not only what cards one is dealt but how well one plays them" (*The cunning of Reason*, 1987, p. 164).

An individual may be considered rational, but if it means that he is always obliged to maximize his own set of goals, he is prevent from saying "no" on making his final decision (Dostoievski). In fact, his decision could be to doubt at the very last moment and in his own right is to decide with trembling hands (Binmore).

As Amartya Sen pointed out, what we see depends on our relative position regarding the observed objects; our beliefs are influenced by what we see and, in turn, our actions are related to our beliefs: "Positionally dependent observations, beliefs, and actions are central to our knowledge and practical reason. The nature of objectivity in epistemology, decision theory, and ethics has to take adequate note of the parametric dependence of observation and inference on the position of the observer" (SEN, A. K., 1993), (p. 126).

This positional standpoint challenges an objectivist tradition in which objectivity would be a kind of invariant in respect of individual observers and their positions (a view from nowhere, as stated by Thomas Nagel). Besides, if we are more than a functional parameter, objectivity should be based upon an strategic interaction between

observers. The context of an objective statement could be the way an object appears from a certain viewpoint.

“The objectivity of observations must be a position-dependent characteristic: not a “view from nowhere” but one “from a delineated somewhere”” (Sen, *op. cit.*, p. 127).

I propose a joint interpretation of Amartya Sen´s thesis and those of ancient scepticism, in order to obtain a set of epistemic conditions that would help us in understanding the minimal requirements to “share” an argument. Positional-dependence does not apply only to observational statements, but also to decisions on actions and beliefs.

Conceptual nets are actual informational filters, concepts are indeed informational filters, or, as it were, they conform a semipermeable membrane: our descriptions can only be constructed according to those filters, they only make sense regarding those filters.

Until recently, the argumentative discourse was most often taken for a sort of mechanism, an algorithm ruled by logical patterns of inference. Argumentation may also be seen as a generating information process, required for its own unfolding. Argumentation is not a signpost road; it is not only a logical grid, not even the content of our discourse, but a fabric that is being knitted in deliberation. The praxis of argumentation implies thus an ultimate variety of cognitive elements, different from the logical patterns (vid. E. Bustos).

As Pearl said: “The art of reasoning under uncertainty equals to show and process resumes of exceptions”. An abundance of information to handle is, for sure, a problem that usually arises in those particular cases. In order to face it, I think we should pay attention to those decision rules related to informational constraints and both way and path that we use foraging information

Bibliography

ALVAREZ Á., J. F. (1992), "¿Es inteligente ser racional?" *Sistema,* 109**,** Madrid, 73-93.

ÁLVAREZ Á., J. F. (1999), "Los problemas de Larry Laudan", *Crítica***,** México, 93-97.

ÁLVAREZ Á., J. F. (2000), "Racionalidad, modelos humanos y economía normativa", *Argumentos de razón técnica*, 3**,** 93-114.

ÁLVAREZ Á., J. F. (2001), "Una aproximación al espacio de las capacidades potenciales" en ÁVILA, A., W. GONZÁLEZ y G. MARQUÉS (comps.), FCE, Madrid, pp. 175-195.

AXELROD, R. M. (1997), *The complexity of cooperation: agent-based models of competition and collaboration*, Princeton University Press, Princeton, N.J.

BOUDON, R. (2001), *The Origin of Values. Sociology and Philosophy of Beliefs*, Transaction Books, New Brunswick.

GIERE, R. N. (2002), "Scienfic cognition as distributed cognition" en CARRUTHERS, P., S. STICH y M. SIEGAL (comps.).

GABBAY, D. M. y J. WOODS (2002), "Normative Models of Rational Agency", International Conference and Research Center for Computer Science Dagstuhl-Seminar on The Logic of Rational Agency,

GIGERENZER, G. y R. SELTEN (comps.) (2001), *Bounded Rationality. The Adaptive Toolbox*, The MIT Press, Cambridge, MA.

HOLLIS, M. (1998), *Trust within reason*, Cambridge University Press, Cambridge.

LAUDAN, L. (1998), "Naturalismo normativo y el progreso de la filosofía" en GONZÁLEZ, W. J. (comp.), pp. 105-116.

MORTON, A. (2001), Fools ape angels: epistemology for finite agents, Página elaborada por.

PIROLLI, P. y S. CARD Information Foraging in Information Access Environments, www.acm.org/sigchi/chi95/Electronic/documnts/paper/ppp_bdy.htm, Página elaborada por.

SEN, A. K. (1993), "Positional Objectivity", *Philosophy and Public Affairs*, 22**,** 126-145.