COGNITION, INSTITUTIONS, NEAR DECOMPOSABILITY: rethinking Herbert Simon’s contribution

Massimo EGIDI\textsuperscript{a} and Luigi MARENGO\textsuperscript{b}

a. CEEL, University of Trento, Via Inama 5, 38100 Trento, Italy, e-mail: massimo.egidi@unitn.it
b. DSGSS, University of Teramo, Loc. Colleparco, 64100 Teramo, Italy, e-mail: marengo@unite.it

1 Rationality within economic institutions: the origins

One of the more illuminating ideas underlying bounded rationality is that the nature of organizations is deeply rooted in the limits of human intelligence and rationality. It is precisely because individual rationality and knowledge are limited that a social division of knowledge and competences is essential for human progress. This principle was clearly stated by the mathematician Alfred Whitehead

“It is a profoundly erroneous truism, repeated by all copy-books and by eminent people when they are making speeches, that we should cultivate the habit of thinking what we are doing. The precise opposite is the case. Civilization advances by extending the number of important operations which we can perform without thinking about them.”

Hayek expressed similar ideas, when he tried to explain the reasons of the economic institutions:

“Clearly there is here a problem of division of knowledge which is quite analogous to, and at least as important as, the problem of the division of labour. But, while the latter has been one of the main subjects of investigation ever since the beginning of our science, the former has been as completely neglected, although it seems to me the really central problem of economics as a social science. The problem which we pretend to solve is how the spontaneous interaction of a number of people, each possessing only bits of knowledge, brings about a state of affairs in which prices correspond to costs, etc., and which could be brought about by deliberate decision only by
Herbert Simon acknowledged in a conversation with one of the present writers, the close connection between the principle of bounded rationality and Hayekian assumptions on knowledge. It may thus be useful to focus here on the differences and the parallelisms between the two authors’ views.

While both Simon and Hayek consider the analysis of institutions to be coessential with a theory of the human mind, the main difference between them is that Hayek considered the market as the sole institution able to coordinate the decisions of distinct individuals endowed with fragments of knowledge, while Simon viewed the division of knowledge and the coordination as complementary processes that characterize the evolution of markets and organizations.

The road to awareness that the human ability to solve problems with bounded rationality characterizes the nature of both markets and organizations, was rugged and required many analytical steps. When Herbert Simon first published *Administrative Behavior*, in 1947, organizational studies were considered to be a field of inquiry entirely distinct from microeconomics, and the scientific approach to the study of the internal mechanisms of organizations was not directly related to the rational behavior of economic agents.

Decision making theory supposedly applied only to the behavior of independent agents in markets, and since people in organizations are by definition closely interdependent in their choice-making, the study of human behavior within organizations was relegated to the sidelines. *Administrative Behavior* opened up the ‘black box’ of an organization’s internal mechanisms. By radically criticizing the existing theories of administration, Simon identified decision making in conditions of uncertainty and interdependence as the basis on which administrative theory could be rebuilt, and thereby created the premises for a new vision of human activities within organizations. This was a view that, by developing over many decades the idea of bounded rationality as problem solving activity, gave rise to a unified approach to economic behavior in markets and other economic institutions.

Bounded rationality and division of knowledge were the core principles on which organizational studies had to be grounded and to which Simon devoted most of his research in the following decades. The former is certainly most well known and underlies a much larger proportion of Simon’s impressive scientific production, but, as we have seen, the latter has always been considered equally important by him: Simon himself strongly emphasized in various letters and at a meeting with one of us just a year before his death that the twin notions of interdependence and decomposability related to the process of division of knowledge, were not only a necessary
complement to his theory of bounded rationality, human problem solving and organizational behavior, but a sort of general unifying principle underlying all viable organized systems: human, biological or artificial.

As said, awareness that the two principles can lead to an unified approach required time: Simon built his theory of bounded rationality on close observation of the behavior of employees and managers in large organizations. During the 1950s and early 1960s, he took part in numerous collaborations and research projects at the Graduate School of Industrial Administration of Carnegie Mellon, among them a study of decision making under uncertainty conducted jointly with Charles Holt, Franco Modigliani and John F. Muth. The aim of the study was to develop mathematical tools to improve inventory control systems for production planning at a plant of the Pittsburgh Plate Glass Company. It was in this context of the concrete study of empirical data that Simon developed his early notions of routines, and satisficing.

But the milestone in the founding of organizational studies on bounded rationality was Organizations, written with James G. March and published in 1958. Here all the major issues connected with bounded rationality and its consequences in terms of adaptive and evolutionary economics were debated.

This now classic book moved forward from the notion of problem solving as individual activity to the notion of organizational problem solving, with clear recognition of the evolutionary processes of organizational adaptation and organizational learning within business corporations. Identification of these processes proceeded in parallel with the discovery that the division of labor can be considered a problem solving activity, and that the recursive division of problems into sub-problems is a property of both organizations and computer programs.

The development of a deeper theory of problem-solving became crucial for explanation of organizational routines and procedures within business firms, and their evolution. Progress in one of the sciences of the artificial - the theory of computation - became of fundamental importance for progress in the other - the theory of organizational learning. It is probably for this reason that Simon's interests moved to computation theories as natural candidates for explanation of human problem solving and discovery.
2 From decision making to cognition

About the time he was finishing his work on *Organizations*, Simon began his collaboration with Allen Newell, a celebrated founding father of the Artificial Intelligence. The collaboration gave rise to the creation of new mathematical tools with which to model human problem solving and discovery processes. *Human Problem Solving*, published in 1972 with Allen Newell, is a bridge between computation, artificial intelligence and cognitive psychology. Here Simon went beyond the notion of "computation" as a human activity which relates means to ends, replacing it with the notion of *symbolic manipulation*. Simon’s first studies on the limits of rationality, in fact, focused on the bounded ability of individuals to construct and explore their strategies for action, of which celebrated examples are the insurmountable obstacles encountered by the players of chess and other complex games when devising winning strategies. But as Simon’s observations and field researches proceeded, he realized that beyond the limits on the human ability to “compute” a strategy in depth lay further limitations, and that these involved most of the cognitive activities connected with decision-making.

Therefore when the limits to the human capacity for mental calculation were experimentally demonstrated, it became clear to him that this ability was an aspect - an important but not unique one - of the mind’s more general capacity to *manipulate symbols* and to create *mental models* of reality.

The recognition that human decision making can be understood only if mental activity is viewed as symbolic manipulation reinforced the bridge with psychology. Simon's research thus shifted to a different version of the problem, subjecting the various mental abilities essential for explanation of human action - memorization, categorization, judgment, problem-solving, induction - to increasingly intense experimental scrutiny. In parallel with this experimental work, Simon developed computational models of intelligence designed to explain the process of discovery.1 After organizational theory and artificial intelligence, it was cognitive psychology that became Simon’s next fundamental area of inquiry as he developed his idea of the human solver and sought to clarify the shortcomings of the traditional theory of decision making within institutions.

The growth of experimental economics since the 1970s has been extraordinarily helpful in highlighting the importance of this line of research. The pioneering works by Allais on expected utility theory violations, and the subsequent discoveries by Kahneman and Tversky, have made it

---

1 The BACON and DALTON programs (Langley et al. (1987)), for example, simulate scientific discovery processes.
clear that Simon’s claims about the scant realism of economic and organizational theories were solidly based.

Once again, and now with the precision of the experiments, the systematic discrepancy between the predictions of the traditional theory of decision making and real behavior became very apparent. Indeed, an enormous number of experiments conducted in recent years have shown that many types of human errors – defined as breaches of rationality – are systematic. Consequently, the benevolence still shown by the majority of economists towards the assumptions of decision theory can only be explained by the lack of a convincing alternative. Economists thus find themselves in the embarrassing situation that they have a largely falsified theory but can only offer fragments of alternative theories in its place.

Simon’s insight that these discrepancies could be overcome by radically revising these theories, and by introducing closer concern with human activity so that systematic account could be taken of the limitations and richness of human intellectual activity, has accordingly become of crucial importance. Moreover, a large body of results from cognitive psychology and experimental economics has given further support to his view. Let us cite only one of these developments of his views.

One of the profoundest of Simon’s intuitions, namely that decision making is deeply rooted in learning activity, has been largely borne out by psychological research, and it gives us important directions for future inquiry. Since the mid-1970s until their most recent study *Choices, Values and Frames*, Tversky and Kahneman have investigated the psychological principles that govern the creation, perception and evaluation of alternatives in decision-making processes. They find that preferences vary substantially according to the way in which the choice problem is presented (‘framed’). Rather than being stable, preferences are constructed by individuals in the process itself of their elicitation; a clear demonstration of this process is provided by the well known experiments in which different representations of the same choice generate a *reversal of preferences*.

This suggests that the crucial aspect of the decision-making process is the ability to construct new representations of problems. This point was already present *in nuce* in Simon’s empirical analysis of managerial decisions conducted in the 1950s. In 1956 Cyert, Simon and Trow pointed out an apparent dualism in managerial behavior, which displays on the one hand coherent choices among alternatives, on the other a search for the knowledge necessary to define the choice context. Consequently, research into rationality shifted its focus from the coherence/incoherence of choices to the representation and editing of problems. How the mental models with which individuals and institutions frame problems are constructed is a crucial issue to be addressed by decision theory in
the years to come, an issue that will yield better understanding of human innovative activities within institutions.

Here again an important parallelism emerges with Hayek’s theory of the mind. Hayek’s theory of institutions was based on a view of the workings of the human mind set out in *The Sensory Order* in the fifties; published a few years after the celebrated *The Organization of Behavior* by D.O. Hebb, this book surprisingly deals with very similar ideas to those expounded by the neurologist. The core of Hayek’s approach is to consider brain activity as classificatory, thus developing a description closely related to the current theory of neural nets. This approach enabled Hayek to describe the process of learning as dynamic evolution of a neural net, and in this way to provide a first profound explanation of the typical features of learning: partial unawareness, unpredictability. Evidently, these characteristics are in Hayek the conceptual pillars for explanation of why the outcome of the actions of individuals is only partially what they intended. Thus for both authors learning is a cornerstone of their theories. It would be easy to claim that Hayek can be considered as a precursor of the connectionist approach to the study of learning while Simon may be considered one of the founders of the symbolic approach. But this distinction, while useful, does not imply that the two approaches are necessarily opposite and incompatible: some approaches, and most notably John Holland’s “classifiers systems” (cf. Holland *et al.* (1986) and Holland (1992)) are indeed attempts to build a bridge between the connectionist and the symbolic approaches to learning, showing the complementarities between them. One important point made clear by the classifiers systems theory is that for both approaches – connectionist and symbolic - the representation of learning process is based on problem decomposability, as very well understood by Simon. In the following section we will briefly discuss this notion and examine some of its implications for the theory of organizations.

3 Decompositions and near-decompositions.

In the *Sciences of the Artificial* Simon puts forward the proposition that basically all viable systems, be they physical, social, biological, artificial, share the property of having a near decomposable architecture: they are organized into hierarchical layers of parts, parts of parts, parts of parts of parts and so on, in such a way that interactions among elements belonging to the same parts are much more *intense* than interactions among elements belonging to different parts. By “intense” interaction is meant that the behavior of one component depends more closely on the behavior of other components belonging to the same part than on components belonging to other parts (i.e. the cross-
derivatives are larger within a part); or that this influence happens on a shorter time scale (effects propagate faster within a part than among parts); or that the influence is more widespread (within a subunit almost all elements interact, whereas interactions among elements belonging to different subunits are more scarce).

This kind of architecture can be found in business firms, where division of labor, divisionalization, hierarchical decomposition of tasks are all elements which define a near decomposable system: individuals within a hierarchical subunit have closer, more widespread, more intense and more frequent interactions than individuals belonging to different subunits. But a very similar architecture can also be found in most complex artifacts (which are made by assembling parts and components, which in turn can be assemblies of other parts and components, and so on), in software (with the use of subroutines, and even more so in object-oriented programming).

Recently the near decomposability hypothesis has been renamed the “modularity hypothesis” and is pervading diverse disciplines, ranging from software design (where the object oriented programming paradigm is nothing but a prescription of a highly modular system) to management science (where largely the same principles are applied to the organization of firms and manufacturing systems, cf. Baldwin and Clark (2000)).

In biology, modularity has been recognized as a fundamental design principle of both the RNA structure and genotype phenotype mapping (cf. Wagner and Altenberg (1996), Callebaut and Rasskin-Gutman (forthcoming)).

In cognitive science, Fodor (1983) is the main proponent of a modular theory of cognition. He argues that certain psychological processes are self contained, or modular. This is in contrast to the “Modern Cognitivist” positions which hold that nearly all psychological processes are interconnected and freely exchange information. Strangely enough, this is perhaps the field of inquiry where the modularity hypothesis is more controversial and strong empirical evidence has been gathered against it (to the point that Fodor himself has recently made his position less clear-cut).

Why should near decomposability be such a general organizing principle characteristic of virtually all complex systems, as diverse as those just mentioned? Herbert Simon gives us a series of possible explanations, which provide food for thought for the various different disciplines connected with this hypothesis.

A first explanation refers to efficiency and is better cast in economic terms: it is basically the old Smithian argument for the efficiency of the division of labor. But we must note that per se this argument can at most explain the decomposition into modules, not the fact that such modules are organized into hierarchical systems. Moreover, it is not clear why there should be any limit to the
division of labor, i.e. why we should not observe (not only in business firms but in all viable natural and artificial systems) an endless process of decomposition into finer and finer elements. In economics the importance of these puzzles has been understood and widely investigated: the former in particular has been the focus of, for instance, the transactions costs literature (cf. for instance Williamson (1975)) which heavily relies on Simon’s bounded rationality theory. As well known, the transactions costs explanation claims that the process of decomposition of economic activities is limited by the availability of coordination mechanisms which can efficiently coordinate sub-units. In the economic realm it is usually believed that competitive markets constitute such mechanisms, but the transactions costs theory observes that under some circumstances markets cannot achieve their full efficiency.

Another explanation is more cognitive in nature and springs directly from Simon’s work on problem solving and bounded rationality (see especially Newell and Simon (1972)). Problem solvers faced with problems whose complexity outweighs their bounded computational capabilities are forced to work on conjectural decompositions of the problem into sub-problems (cf. Egidi (1992)). On the one hand such decompositions are necessary heuristics for computationally limited individuals, as they reduce to a collection of sub-problem of more manageable size problems whose complexity largely defeats their computational capabilities. Moreover, very often some of these sub-problems may already be familiar to the problem solvers or at least display some analogies with known problems. However by treating sub-problems as independent or quasi-independent some existing interdependencies are almost inevitably ignored and sub-optimality, biases, systematic mistakes are almost inevitably introduced.

This a fundamental source of suboptimality due to bounded rationality and computational constraints: it is a source which relates to limitations to bounds in the representational capabilities of individuals, organizations and society and sets a clear constraints on the efficiency of the division of labor and knowledge. In a world of interdependencies the process of decomposition encounters an intrinsic (in the sense that it is not due to the quality of the coordination mechanisms) dynamic inefficiency, which arises from the inevitable separation of interdependent elements. In our view transaction costs economics does not recognize the importance of such a trade-off because it does not consider that a given partition of activities between markets and hierarchical organizations embeds a particular “decomposition” of the “economic problem” and that, as a consequence, the features of this decomposition impinge on the efficiency of the system regardless the efficiency of markets themselves (cf. Marengo et al. (forthcoming)).
Biologists have a clearer view of the issues involved in a process of decomposition and modularization, and as Simon himself stresses (cf. Simon (2000) and (forthcoming), can teach interesting lessons to economists and organization scientists.

Indeed sub-problem decomposition, complexity reduction and modular architecture also operate in the biological domain, where we observe modularity in the genotype-phenotype mapping and in the morphological properties of organisms. Thus modularity is a very important property which has evolved through selection. According to Wagner and Altenberg (1996) modularity may evolve through two opposite processes which they call, respectively, integration and parcellation in the genotype-phenotype mapping: integration means that genes acquire new pleiotropic effects phenotypic characteristics, while parcellation is the suppression of existing pleiotropic effects.

Of course this applies also to social organizations and business firms where sub-problem decomposition is indeed one aspect of the division of labor. This sub-problem decomposition theory suggests that two fundamental processes engender the dynamics of the “organizational architecture” of the economic system (and in particular the relative distribution of market and non-market coordination mechanisms): a process of deepening the sub-problem decomposition, with an endless division of problems and tasks into sub-problems and sub-tasks, and a process of recombinining sub-problems and sub-tasks into modules.

Note that the transaction cost argument is mainly restricted to consideration of the latter process, as it takes the division of labor to be given and looks for an explanation of the organizational structure of the economic system by establishing the conditions under which existing modules (organizations) cannot be efficiently coordinated by markets and must be combined hierarchically into higher order modules.

As Simon (forthcoming) observes, the latter process is likely to produce a near decomposable system by agglomeration of parts, while the kind of near decomposability we observe in the economic realm most often originates from the division of previously fully integrated parts.

That the process of integration should normally give rise to a near decomposable system has been clearly stated in Simon (1996) and even more clearly in Simon (2000), where he claims:

“If we begin with a set of simple elements that, when they meet at random, are capable sometimes of forming stable combinations, and if the combined systems [...] are themselves capable of combining into still larger systems, then the complex systems we will observe [...] will almost all be near decomposable systems” (p. 6)

In Simon (1996) this proposition was exemplified by the parable of the two watchmakers, called Tempus and Hora. Both watchmakers are so successful that they receive many phone calls from potential clients which disrupt the production process. If the latter is organized as a unitary
assembly chain then after every disruption the watchmaker will have to start all over again. If instead the production process is organized into a hierarchical system of sub-assemblies of stable components, then a phone call will disrupt only the assembly of one component. Thus systems which are nearly decomposable are much less vulnerable than systems which are not, as disturbances are more likely to remain confined to specific subcomponents: near decomposable systems limit interactions and information flows among different parts of the system and thus are better able to keep damaging events confined to sub-parts.

Note that this argument is not limited to damaging events *stricto sensu*, but, perhaps more importantly, to all disturbances to the system which may feed adaptation and change. This point is very important: successful change and adaptation requires a great deal of trial and error. Every adaptive system must be able to bear a high rate of mistakes without losing its functionality: variability is the driving engine of adaptation, but of course most of the variation produced amounts to error.

The point has recently attracted the interest of computer scientists (mainly those working in the theory and application of evolutionary computation) and biologists, but it has enormous implications for the theory of economic organization and once again one of the earliest illustrations of the point can be found in Simon (1965). Adaptation and evolution can work effectively only if variations are not too often very disruptive and not too rarely favorable. In computer science it is well understood that this depends crucially on how the task environment is coded (the so-called representation problem). By acting on the representation we can make simple adaptive mechanisms such as mutation and recombination either very effective or totally ineffective.

In biology a similar problem concerns the “evolvability” of organisms: if, as biologists claim, evolution is the result of the interplay between mutation and selection, then we must ask where these mutations come from, and in particular whether an evolvable genome is itself a product of evolution. A representation problem also arises in biology, because the crucial factor affecting “evolvability” is the genotype-phenotype map, which determines how variations at the genotypic level affect the phenotypic characteristics and thus the fitness of organisms. As mutations in the biological world are believed to be totally random, this randomness must be able to produce some improvement.

Again the likelihood of producing ordered and stable structures out of totally random events depends crucially on the representation: modularity and near decomposability can be a way to solve the representation problem (cf. Wagner and Altenberg (1996)). We can better understand this by asking whether there is some chance that a group of monkeys with their random actions are able to produce a verse of Shakespeare: if monkeys are simply given a piece of paper and a pencil they will
probably draw signs which are no way similar to anything understandable, but if we give them a typewriter and let them play with its keyboard they will at least produce strings of letters, and the possibility that one of these strings is an English word (possibly also found in Shakespeare) is probably not so remote. If we give them tiles with English words and let them arrange them in sequences we might even from time to time get a short meaningful sentence. The likelihood of obtaining a verse of Shakespeare is probably still very low, but is certainly orders of magnitude higher than in the case in which we gave the monkeys only paper and pencil.

Note that what we are doing here is simply endowing monkeys with different representational systems; that is, different modules which they can recombine randomly. More structured representations involve higher level modules which on the one hand greatly limit the space of possible variation (with pencil and paper monkeys could produce not only a verse of Shakespeare but also a drawing of Leonardo, while with a typewriter the latter is no longer possible), but impose constraints upon variability which safeguard some kind of coherent structure.

We believe that it could be very fruitful to pursue a line of research in organizational economics which starts from the acknowledgement that the way in which economic activities are organized, and in particular their distribution between markets and hierarchies not only determines the static efficiency of the economic system but also, and more importantly and with inevitable trade-offs, shapes the possible dynamic “morphological” paths which the system can follow, its space of possibilities (for an interesting formal characterization of the space of possibilities in biological evolution see Stadler et al. (2001)).

4 Final remarks

Besides re-establishing close connection among economics, psychology and the cognitive sciences, Simon’s research suggests that the actions of decision-makers in the real economic world should be studied, not merely in terms of rationality, but in the light of the capacity of the human mind to frame problems, and to represent reality in innovative ways, in an endeavor to reduce their uncertainty and ignorance. The “Sciences of the Artificial”, namely organization theory, cognitive psychology and artificial intelligence, are the new disciplines now investigating the classic yet still unresolved question of human creativity and learning, and their relationship with the evolution of institutions.

In these notes we have briefly elaborated on learning process, focusing the attention on near decomposability, a notion which plays a very important role in Simon’s theory of learning, organization and, more in general, complex systems. Simon ideas with regard to near
decomposability are far from being exploited (and sometimes even from being understood) by current research in economics, and therefore developing on them (as biology is doing to a much greater extent) we could open new perspectives on key questions concerning the organization of economic activities and their distribution between market and non market coordination modes.

References


