Laboratory of Economics and Management

Sant'Anna School of Advanced Studies
Piazza Martiri della Libertà, 33 - 56127 PISA (Italy)
Tel. +39-050-883-343 Fax +39-050-883-344
Email: lem@sssup.it Web Page: http://www.lem.sssup.it/

# LEM

## Working Paper Series

# Toward Formal Representations of Search Processes and Routines in Organizational Problem Solving. An Assessment of the State of the Art

Giovanni Dosi [o §]
Marco Faillo [*]
Luigi Marengo [o]
Daniele Moschella [o]

[o] LEM-Scuola Superiore Sant'Anna, Pisa, Italy
[§] Schiller University, Jena, Germany
* University of Trento, Italy

2011/04 February 2011

ISSN (online) 2284-0400

# TOWARD FORMAL REPRESENTATIONS OF SEARCH PROCESSES AND ROUTINES IN ORGANIZATIONAL PROBLEM SOLVING. AN ASSESSMENT OF THE STATE-OF-THE-ART

Giovanni Dosi[*x], Marco Faillo[*#], Luigi Marengo[*] and Daniele Moschella[*]

[*] LEM -*Sant'Anna School of Advanced Studies, Pisa*

[#] *University of Trento, Italy*

[x] *Visiting Professor Schiller University, Jena*

*giovanni.dosi@sssup.it* *marco.faillo@unitn.it* *lugi.marengo@sssup.it*
*d.moschella@sssup.it*

**Abstract**

This paper presents a critical overview of some recent attempts at building formal models of organizations as information-processing and problem-solving entities.
We distinguish between two classes of models according to two distinct objects of analysis. The first class includes models mainly addressing information processing and learning; the second class includes models focusing upon the relationship between the division of cognitive labor and search process in some problem-solving space.
The results begin to highlight important comparative properties regarding the impact on problem-solving efficiency and learning of different forms of hierarchical governance, the dangers of lock-in associated with specific forms of adaptive learning, the relative role of "online" vs. "offline" learning, the impact of the "cognitive maps" which organizations embody, the possible trade-offs between accuracy and speed of convergence associated with different "decomposition schemes", the (ambiguous) role of organizational memory in changing environments.
We argue that these are important formal tools towards the development of a comparative institutional analysis focusing on the distinct properties of different forms of organization and accumulation of knowledge.

**Keywords:** Information processing, Problem-solving, Organizational structure.
**JEL classification**: D23, D83, L22

## 1. Introduction[1]

This work is meant to offer a critical overview of the achievements and challenges ahead facing explicit formalizations of organizations as information-processing and problem-solving entities.

The importance of the information-processing arrangements is well acknowledged within both *agency* and *capability*-based theories of the firm, even if only the latter focuses on the problem-solving features of organizations.

Firms after all “do things” – whether material as a car or more “immaterial” as a software program or an airline reservation system - , try to improve over time what they do and quite often also try to innovate and find new things . “Problem-solving” is a synthetic notion covering both the current operations of an organization and its search for novel ones.

In this respect, note that most formal representations of organizations tend to offer highly *blackboxed* accounts of such activities. In that, agency models are an extreme case to the point where the whole activity of information processing is compressed in some function maximization conditional on the appropriate processing of the available information while “problem solving”, in the above sense, is almost entirely neglected. On the contrary, here we shall survey those endeavours which try to account for organizational information processing and problem-solving in terms of explicit sequences of activities and procedures nested into specific organizational arrangements prescribing "who send which signals to whom" and "who does what and in which sequence".

The appreciative theories upon which such model draw represent a small – but not negligible and growing – minority of the economic profession who place their “primitives” of the nature of economic organizations in their problem-solving features, in turn nested in ubiquitous forms of human “bounded rationality”, grossly imperfect processes of learning and diverse mechanisms of social distribution of “cognitive

[1] The work draws upon other works of the authors, in particular: Cohen *at al.* (1996), Dosi, Nelson and Winter (2000), Marengo and Dosi (2005), which the reader is referred to for further details.

labor". The roots of this approach can be found in the works of Herbert Simon, James March, Alfred Chandler and Richard Nelson and Sidney Winter[2].

The problem-solving activities of the firm can be conceived as combinations of physical and cognitive acts, within a procedure, leading to the achievement of a specific outcome. Its internal organization determines the distribution of the informational inputs across specific task units and, as such, the division of the cognitive labor. The general idea is that firms possess the specific problem-solving competencies associated with their own operational procedures and routines, in turn embedded into the patterns of intra-organizational division of labor and assignments of decision entitlements.

An illustrious antecedent of this view dates back, indeed, to Adam Smith's "Pin Factory" example in *The Wealth of Nations:*

> "One man draws out the wire, another straights it, a third cuts it, a fourth points it, a fifth grinds it at the top for receiving the head; to make the head requires two or three distinct operations; to put it on, is a peculiar business, to whiten the pins is another; it is even a trade by itself to put them into the paper; and the important business of making a pin is, in this manner, divided into about eighteen distinct operations, which, in some manufactories, are all performed by distinct hands, though in others the same man will sometimes perform two or three of them."(Smith, 1776)

How does one formalize these basic intuitions?

It is fruitful to distinguish between two (complementary) classes of models according to two distinct objects of analysis. The first class includes models mainly addressing information processing and learning. Here the focus is on the relation between organizational performance, learning patterns and the structure of information flows. Agents are adaptive learners who adjust their information processing capability (i.e. their knowledge of the environment) through local trial-and-error.

The second class includes models focusing upon the relationship between the division of cognitive labor and search process in some problem-solving space, addressing more directly the notion of organizations as repositories of problem-solving

[2] See Chandler (1977), Cyert and March (1963), March and Simon (1993), Nelson (1991, 2008), Nelson and Winter (1982), Simon (1962) and (1981).

knowledge. Here the focus is on the problem-solving procedures which the organization embodies. Indeed, managing an organization, designing and producing cars or software packages, discovering a new drug, etc. can been seen as complicated problems whose “solutions” are made of a large number of cognitive and physical acts. These kinds of activities imply the coordination of large combinatorial spaces of components.

On the output side, components which make up an artifact can take a number of alternative states: so, for example, in the case of the production of a car, one combines different characteristics of the engine, alternative designs, different materials, etc. At the same time, innovative search may be straightforwardly represented in form of combination of multiple “cognitive acts” eventually yielding the solution of the problem at hand, e.g. the discovery of a new molecule with the required characteristics, a reasonable and coherent software package, etc. Note that in both examples the existence of strong interdependencies among the components – which often are only partially understood by all agents involved - implies that the effect on the system’s performance of a change in the state of a single component depends on the values assumed by the other ones. An implication is also that in this kind of problems it is impossible to optimize the system by optimizing each single component.

By applying this view to organizational analysis one can conceive economic organizations as bundles of routines, procedures, rules characterized by strong interrelations which often are opaque to organizational members.

Notice, first, the partial “opaqueness” of the mappings between actions and outcomes is quite in tune with “garbage can” interpretation of organizational dynamics (Cohen et al. 1972).

Second, “interrelatedness” also lies behind plenty of evidence regarding the widespread difficulties in *replication* and *transfers* of incumbent organizational arrangements (Winter and Szulanski, 1998, 2002; Zander and Kogut, 1995).

Third, an obvious implication of such relatively opaque interrelatedness is also that the introduction of a new routine which has proven superior in another situation might have negative effects on the performance of the organization if other interrelated components are not appropriately co-adapted (Marengo and Dosi, 2005; Marengo *et al.*, 2000).

## 2. Information processing and structural learning

Let us start by considering those (still few) models whereby information-processing and problem-solving activities are represented by ensembles of condition-action (that is, “if…then...”) rules.

Marengo (1992) and Marengo (1996) present a model which focuses upon the modification of such information processing capabilities of individuals or subunits within the organization, i.e. a process of "structural" learning. Agents are imperfect adaptive learners, as they adjust their information processing capabilities through local trial-and-error. This adaptive learning is (at least partly) driven by the information coming from the environment and/or from other members of the organization.

Let

$$S = \{s_1, s_2, ..., s_N\}$$

be the set of the $N$ possible states of nature and

$$A = \{a_1, a_2, ..., a_k\}$$

the set of the $k$ possible actions the decision-maker can undertake. The payoff to the agent is given by a function:

$$\Pi: A \times S \rightarrow R$$

where the agent's payoff to action $a_i$ when the state of the world $s_h$ occurs will be indicated by $\pi_{ih}$.

The action the agent chooses depends obviously on the level of its knowledge about the state of the world. The agent's state of knowledge (or information processing capabilities) can be represented by a collection of subsets $P(s_i) \subseteq S$ where $P(s_i)$ is the set of states of the world which the agent considers as possible (or cannot tell apart) when the real state is $s_i$.

The basic component of this learning system is, as mentioned, a condition-action rule, where the execution of a certain action is conditional upon the agent's perception that the present state of the world falls in one of the categories it has defined in its mental model. The condition part is a category, that is a subset of the states of the world, and is activated when the last detected state of the world falls in such a subset. Practically, the condition is a string of *N* symbols (as many as the states of the world) over the alphabet {0,1} and it is satisfied whenever the last state of the world corresponds to a position where a "1" appears. All in all, the condition:

$$c_1 c_2 ... c_N \ \textit{with} \quad c_i \in \{0,1\}$$

is satisfied when, if $s_k$ is the last observed state of the world, we have $c_k = 1$. Thus, a set of conditions defines a subset of the power set of S. It is important to notice that each condition defines one subjective state (or category) of the world, *as perceived by the agent*, and defines its relationship with the objective (true) states of the world. This relationship remains anyway unknown to the decision maker, who is aware only of its subjective states.

The action part is instead a string of length *k* (the number of the agent's possible actions) over the same alphabet and with the following straightforward interpretation:

$$a_1 a_2 ... a_k \ \textit{with} \quad a_i \in \{0,1\}$$

has one and only one position which equals "1", $a_h$=1, meaning that the action "*h*" is chosen, and "0's" everywhere else.

The decision maker can be therefore represented by a set of such condition-action rules:

$$R = \{R_1, R_2, ..., R_q\}$$

where:

$$R_i : c_1, c_2 ... c_N \Rightarrow a_1 a_2 ... a_k \ \textit{with} \quad c_i, a_h \in \{0,1\}.$$

Each rule is assigned a "strength" and a "specificity" measure.

Strength basically measures the past usefulness of the rule, that is the rule's cumulated payoff. Specificity measures the strictness of the condition: the highest specificity (or lowest generality) value is given to a rule whose condition has only one symbol "1" and therefore is satisfied when and only when that particular state of the world occurs, whereas the lowest specificity (or the highest generality) is given to a rule whose condition is entirely formed by "1's" and is therefore always satisfied by the occurrence of any state of the world.

In this model, at the beginning of each simulation the decision maker is supposed to be completely ignorant about the characteristics of the environment he is going to face: all the rules initially generated have the highest generality, meaning that all their conditions are formed entirely by 1's. The action parts are instead randomly generated.

The decision maker is also assumed to have limited computational capabilities, therefore the number of rules stored in the system at each moment is kept constant and relatively small in comparison to the complexity of the problem which is being tackled.

This set of rules is processed in the following steps throughout the simulation process:

1. *Condition matching*: a message is received from the environment which informs the system (the agent or a structured collection of them) about the last state of the world. Such a message is compared to the condition of all the rules and the rules which are matched, i.e. those which apply to such a state of the world, enter the following step.

2. *Competition among matched rules*: all the rules whose condition is satisfied compete in order to designate the one which is allowed to execute its action. To enter this competition each rule makes a metaphorical "bid" based on its strength and on its specificity. In other words, the bid of each matched rule is proportional to its past usefulness (strength) and its relevance to the present situation (specificity):

$$Bid(R_i, t) = \left(k_1 + k_2 Specificity(R_i)\right) Strength\left(R_i, t\right)$$

where $k_1$ and $k_2$ are constant coefficients. The winning rule is chosen randomly, with probabilities proportional to such bids.

3. *Action and strength updating*: the winning rule executes the action indicated by its action part and has its own strength reduced by the amount of the bid and increased by the payoff that the action receives, given the occurrence of the "real" state of the world. If the $j^{th}$ rule is the winner of the competition, we have:

$$Strength(R_j, t+1) = Strength\left(R_j, t\right) + Payoff\ (t) - Bid(R_j, t)\,.$$

4. *Generation of new rules*: the system must be able not only to select the most successful rules, but also to discover new ones. This is ensured by applying *genetic operators* which, by recombining and mutating elements of the already existing and most successful rules, introduce new ones which might or might not improve the

performance of the system. Thus new rules are persistently injected into the system by recombining and/or locally modifying existing knowledge.

*Genetic operators* generate new rules which - typically but not necessarily - explore other possibilities in the vicinity of the currently most successful ones. Hence, search is not completely random but influenced by the system's past history.[3]

In Marengo (1992) and Marengo (1996) two genetic operators have been used for the condition and one for the action part. The latter is a simple local search and is just a mutation in the “vicinity”: the action prescribed by the newly generated rule is chosen (randomly) in the close proximity of the one prescribed by the parent rule. For example, a mutation in the action part may probabilistically mutate the product type prescribed by the rule into one of the neighbouring product types.

The two operators used for the condition part deserve more attention because of their role in modelling the evolution of the state of knowledge embedded into the system. They operate in opposite directions:

- *Specification*: a new condition is created which increases the specificity of the parent one. Wherever the parent condition presents a 1, this is mutated (with small probability) into a 0;
- *Generalization*: the new condition decreases the specificity of the parent one. Wherever the latter presents a 0, this is mutated (with small probability) into a 1.

Note that s*pecification* and *generalization* stand for two possible "cognitive" strategies which tend to drive the learning system towards, respectively, rules which apply to more specific states of the world and rules which instead cover a wider set of states of the world[4].

The basic model outlined so far is used to study a variety of *coordination problems possibly conditional on changing environmental states*, thus analyzing organizations

[3] In incumbent models, new rules take the place of the currently weakest ones, so that the total number of rules is kept constant.

[4] Different degrees of specification and generalizations can be simulated both by means of different combinations of these two genetic operators and by varying the coefficient $k_2$ with which specificity enters the bid equation: the higher this coefficient, the more highly specific rules will be likely to prevail over general ones. The simulations discussed below use a specificity coefficient to summarize the overall inclination of the system toward the search for specific rules, such coefficient will represent both the value $k_2$ in the bid equation and the probability of application of the genetic operator of specification every time the genetic operator’s routine is called.

which have to respond to an exogenous *and changing* environment by implementing some collective actions.

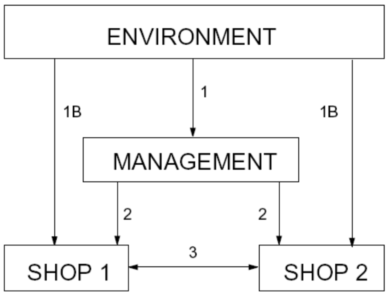
Suppose for instance that a firm can produce a certain number of product types, demanded by an exogenous market, and that the production process is divided into several parts, each of them being carried out by a different “shop”. The problem is therefore to detect correctly which product type is being demanded (the “state of the world”) and to coordinate the actions of the shops so that the correct production process is implemented.

As an illustration, suppose that there exist eight possible product types, called respectively “1”, “2”,. . . , “8”. The firm's production possibilities set is represented by sequences of operations which can be of two types (A and B). Such sequences have all the same length and map into a product type, which is conventionally designated by the number of operations of type A which are utilized in its production. For example the product of type “8” could be produced by all and only the production processes which contain eight operations of type A. Each production process is divided into two parts (of the same length) which are carried out separately by two “shops” (divisions). The problem of the firm is therefore to forecast the product type which will be demanded by the market and to implement the correct production process by coordinating the operations of the two shops. Suppose that the payoff is the following: if the firm produces the correct product type it receives a payoff of 5 units; if it does not produce the correct output it receives a negative payoff, given by the distance of the actual product type from the required one (for example, if the market demands type “7” but the firm produces type “5”, it will receive a payoff of -2).

Suppose now that the all the decision-making units which the organization is made of are represented by agents whose knowledge of the state of the world evolves exactly in the way presented above.

Marengo (1992) and Marengo (1996) simulate the behaviour of a simple but quite general organizational structure, visualized in Figure 1, composed by a "management" and two shops. The management observes the environmental message (the last state of the world), interprets it according to its, evolving, "model of the world", and sends a message to the two shops.

*Figure 1. Organizational informational flows (Marengo,1992, 1996)*




Each of the two shops can, in general, observe three kinds of signals and develop an interpretative model for each of them. These signals are, respectively, the environmental signal (last observed state of the world), the message sent by the management (based on the latter’s interpretation of the environment), and the signal sent by the other shop (in the incumbent model, its last action). The latter two messages are coordinating devices, respectively a centralized and a decentralized one, aimed precisely at fostering coordination among actions, whereas the former allows the two shops to form their own independent (from the management one) models of the world.

The weights with which these three types of messages enter the shops' decision processes define the organizational balance between differentiation and commonality of knowledge, in turn shaped by the power distribution along the organizational hierarchy.[5]

A high specificity coefficient for the condition part which classifies messages coming from the management (messages of type 2 in Figure 1) implies that shops attribute great importance to the correct interpretation of the coordinating messages which are sent by the management. A low coefficient implies instead that shops are

[5] Such weights are represented by the specificity coefficients which express the agent's search for a precise model which interprets the corresponding type of message. A high specificity coefficient for the shops' condition parts which classify messages coming from the environment (messages of type 1B in Figure 2) implies that shops are aiming at building a detailed individual model of the world. A low coefficient implies instead that shops do not pay much attention to the environment. When the coefficient is equal to zero we have an organization in which shops do not form any autonomous model of the world but rely entirely on the world's interpretation given by the management (messages of type 1 and 2).

“free” to some extent to neglect hierarchical message. When the coefficient is equal to zero we have an organization without any form of centralized coordination, in which top management has no role.

Finally, a high specificity coefficient for the condition part which classifies messages coming from the other shop (messages of type 3 in Figure 1) implies that shops are attaching high importance to mutual, decentralized coordination. When the coefficient is equal to zero we have an organization without any form of decentralized coordination, i.e. no inter-shop communication.

In general the model shows that the architecture of such information flows plays a crucial role in determining the learning patterns and the performance characteristics of the organization. In particular simulation results from Marengo (1992) and Marengo (1996) include the following.

First, in stationary environments (i.e. when the state of the world does not change) agents can in fact achieve coordination without building any model of the environment and resorting only to trial-end-error *cum* adaptive selection of rules. Interestingly, note that if instead they try to learn, i.e. to build such a model and constantly improving it, they need also to learn a model for the interpretation of coordinating messages (messages 1 and/or 1B are not sufficient, and messages 2 or 3 are also needed).

Second, if the environment undergoes predictable changes (for example of a cyclical type), high specificity coefficients on the shops' conditions which classify environmental messages (message 1B) are needed in order to exploit the environmental regularities. Shops need to have a direct access to environmental information in order to develop the necessary decentralized learning.

Third, if the environment undergoes frequent and unpredictable changes, the organization has to develop stable routines which give a “satisficing" average result in most conditions. In this case decentralized learning is detrimental, because the stability of such routine is continuously jeopardized by individual efforts to grasp unpredictable environments. Shops are better off by relying on the management's message.

Under predictably changing environments the most appropriate organization is the one which, by partly decentralizing the acquisition of knowledge about the environment, can achieve higher levels of sophistication in its model of the world, provided that coordination mechanisms - which are *centralized* - are powerful enough to enable the organization to solve conflicts of representations. On the other hand, this very decentralization of the acquisition of knowledge can be a source of loss when it is more efficient for the organization to cling to a robust and stable set of routines. The explorations so far suggest that “Knightian uncertainty” requires strong coordination enforcing a set of coherent and robust routines over the entire organization. Autonomous and decentralized experimentation can only disrupt such a coherence.

In a somewhat similar modelling vein, Pentland and Reuter (1994) formalize organizational routines as a set of functionally similar patterns represented via rule-based grammar models. So a routine is a “grammar” which defines all the action patterns which are, so to speak, “legal”, having different action patterns as possible instantiations triggered by different environmental or intra-organizational signals (the “if” part).

Moreover, it is quite straightforward to represent also the *memory* of an organization (both its collective “cognitive” memory and its “operational” one) in terms of structured ensembles of “if…then…” rules (cf. the classic Walsh and Ungson, 1991). With such apparatus, Dosi, Marengo, Paraskevopoulou and Valente (2011) try to answer some questions about the relationship between memory characteristics, organizational architectures and patterns of environmental change (What are the effects of different distributions of memory elements within the organizations? How does a shock like labor turnover act upon both operational and cognitive memories?).

The bottom line is that one ought to consider the foregoing models as a template for a largely unexplored family of exercises which takes seriously on board (i) informational imperfections; and even more importantly, differences in cognitive models, (ii) "boundedly rational" information processing; (iii) adaptive learning; and (iv) inter-organizational differences in information channels and decision rules. Indeed in the foregoing types of exercises, "blackboxing" is reduced to a minimum in so far as flows of information, cognitive dynamics, and decision acts are explicitly modelled. The downside rests precisely in the high dimensionality of the space in which rules evolve and the related difficulty in identifying robust features of the mappings from rules to organizational performances, however defined.

## 3. Models of evolution in the space of "traits" and problem solving

A way of overcoming such drawbacks involves precisely some “blackboxing”, in particular concerning the relationship between organizational traits (including of course behavioural rules) and their actual expressions. Such modelling genre prominently includes a new family of evolutionary models of organizations inspired by biologist Stuart Kauffman's so-called “NK model” (Kauffman, 1993). His model of selection and adaptation in complex environments represents evolving entities characterized by non-linear interactions among their elements. In Kauffman (1993) the “NK-model” primarily deals with the evolution of populations of biological entities described by a string of "genes", but its formal structure allows for various applications in other domains. The model, indeed, has lent itself to a growing number of applications, extensions and modifications within the realm of organization studies. In this section we will present the general characteristics of the NK model and review some of its applications, well short of a comprehensive survey, with the primary purpose to flag some of the main results and incumbent challenges.

### *3.1 The NK model*

In the NK model, an entity (an organization for our purposes here) is represented as a string of (binary) traits linked together by a thread of interdependencies (referred to as “epistatic” relations in population genetics) which map into an equally stylized environment delivering performance feedbacks which, in turn, select in favor/against such configuration of traits.

More formally, an organization is described by a string of $N$ loci which refer to the set of traits ($i=1\ldots N$) that make up the organization (the system). For each element $i$, there exist $A_i$ possible states[6]. The set of all possible configurations (strings) of system’s elements $A_1\times A_2 \times\ldots A_N$ is called the possibility space of the system.

---

[6] In most applications and in all those we consider in this paper, the number of states is reduced – for the sake of simplicity – to two: $A_i \in \{0,1\}$.

Next, a fitness function F: $A_1 \times A_2 \times \ldots A_N \rightarrow [0,1]$ is defined which assigns a (normalized) real number to each possible string as a measure of its relative performance. The fitness of the string is usually defined as the mean value of the fitness values of each element ($f_i$), which are in turn randomly drawn from a uniform distribution between 0 and 1:

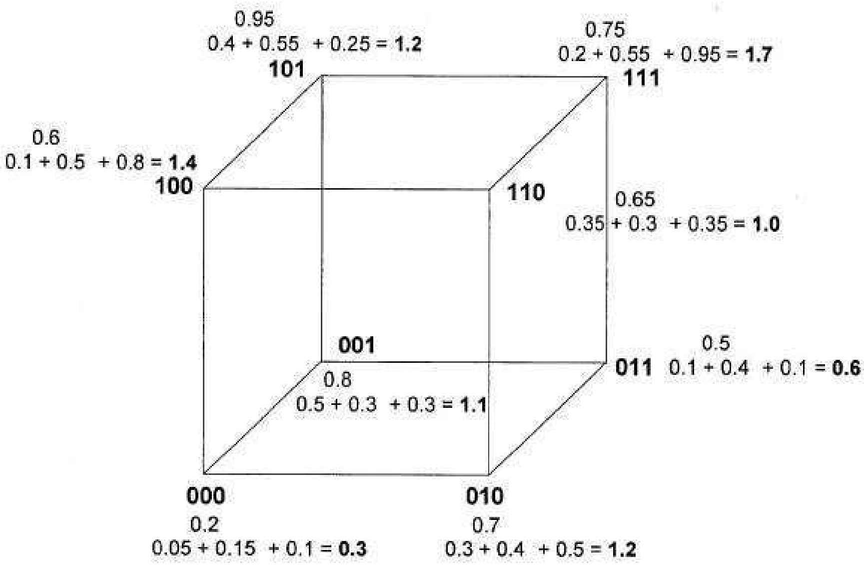
$$F = \frac{\sum_{i=1}^{N} f_i}{N}$$

The degree to which the fitness of the organization depends on the interaction effects among the traits is specified by the variable $K$, which refers to the number of "epistatic" relations among elements (in fact representing the structure of the system itself). The existence of these relations implies that the contribution of one element to the overall fitness of the system is dependent both upon its own state and upon the state of $K$ other elements. Thus, each trait can take on $2^{K+1}$ different values, depending on the value of the trait itself and the value of the $K$ other traits with which it interacts. Two limit cases of complexity can be distinguished, ranging from the *minimum complexity* when $K$=0, to the *maximum complexity* when $K=N-1$.

The distribution of fitness values to all possible configurations defines the *fitness landscape* of the system. This landscape can be explored in search for the configuration with the maximum fitness value, moving from one configuration (a point in the fitness landscape) to another, by changing the value of one element. This "adaptive walk" ends when a configuration is reached which has not immediate neighbours with better fitness.

Consider for example a system characterized by $N$=3, $A_i \in \{0,1\}$ and $K$=2. In this case, all eight ($=2^3$) possible configurations can be depicted on a cube. Each vertex of the cube represents a different configuration of the system; vertices that are connected to each other differ in only one trait. The fitness value of each configuration is, in this case, just the sum of the fitness value of each trait:

*Figure 2. Example of a performance landscape (from Siggelkow and  Levinthal 2005*




### *3.2 Organizational dynamics on complex selection landscapes*

With such a model in mind, let us build upon one of the earliest applications of the "NK" approach to organizational analysis, presented by Levinthal (1997). In Levinthal's simulations, populations of randomly generated structures (organizations) evolve on a fitness landscape, whereby the evolution is driven by variation selection and retention processes.

Variation, i.e. the generation of variety, is provided by two mechanisms:

- *local search*: one-feature mutation with retention of strings with higher fitness value.
- *Radical changes ("long jumps")*: mutation of many (possibly all) features with retention of strings with higher fitness value.

Selection is obtained by simple birth and death process: organizations die with a probability inversely proportional to their relative fitness and are replaced by newly born ones. Some of these organizations are randomly generated, owing possibly no resemblance to the existing ones, while others are replica of existing successful organizations.

Information is maintained intertemporally by means of two mechanisms:

- *retention*: successful existing organizations have a higher probability of surviving. Their features tend therefore to survive with them.
- *replication*: some of the newly born organizations, which replace bad performing ones which are selected out, are copies of the most successful existing organizations. The features of the latter tend therefore to spread in the population.

Consider a large population of randomly generated organizations which evolves according to the just mentioned mechanisms of selection and information reproduction but suppose that variation can be only local, i.e. that only one bit at a time can be mutated for every organization. Local adaptation and selection will reduce the heterogeneity of the population: bad performers will be selected out and replaced by copies of good performers. In the meantime good performers will climb with local mutations the fitness peaks on whose slopes they are located.

However, the final outcome of the evolutionary process will crucially depend on the value of $K$, i.e. the complexity of the fitness landscape. With $K$=0 local adaptation will quickly take all the organizations to the only global optimum: thus selection and adaptation will completely wipe out the initial heterogeneity of the population and yield convergence to unique optimal organizational form. For higher values of $K$ the landscape will display an increasing number of local optima on which subsets of organizations will converge according to their initial configurations. Selection and adaptation will reduce the heterogeneity but will never make it disappear.

This result, robust and general in this framework, must not be overlooked, as it provides a simple and intuitive explanation of the persistence of heterogeneity among firms, a piece of evidence widely reported by the literature but at odds with standard theories, according to which deviations from the only best practice should be only a transient property inevitably due to fade away as market selective forces operate. Note also that as $K$ increases not only does the number of local optima increases, but also the size of the basin of attraction of each of them tend to shrink. It could well be therefore that none of the organizations might be located in the basin of attraction of the global optimum and therefore no organization will ever find the globally optimal configuration.

In complex environments diversity of organizational forms can even emerge out of homogeneity. Levinthal (1997) shows that even if one starts from a population of homogeneous organizations, random local search induces mutations in different directions in the landscape. If $K>0$ such initial random mutations will take organizations in the basins of attraction of different local optima. On the other hand, selection and adaptation will only *partially* reduce such diversity.

If organizations can perform more radical changes ("long jumps"), i.e. mutate many (possibly all) features, also in presence of large $K$ heterogeneity tends to disappear, though very slowly, as organization located on sub-optimal peaks can always perform - though with low probability - a radical mutation which allows them to jump on a higher fitness "hills", until they reach the highest one (i.e. the one whose peak is the global optimum). However, note also that if $N$ is large enough such a process may have a very low probability and be of no actual consequence for the medium term evolution of the population under consideration.[7]

Consider now the case of environmental changes, which can be modelled by re-drawing the fitness contributions of some features after the population has evolved and stabilized over previous local optima.

Suppose first that such a change concerns only one feature and $K=0$, then if the fitness contribution of only one attribute is modified, the global optimum will either remain where it was or move to a point which is at most one mutation away. Thus, if the population has already evolved and located on the global optimum, it can easily and quickly adapt and move to the new global optimum. Simulations show that all incumbent organizations survive to such an environmental change.

However, if the complexity of the landscape is high ($K>0$), even the modification of the fitness contribution of just one attribute can cause a large alteration of its shape. In high dimensional landscapes with large N local optima may well move far away. This implies that a population which has settled on the local optima of the initial landscape will find it generally very difficult to adapt to the change. Mortality of incumbents will rapidly rise as $K$ increases.

If the environment changes more radically, i.e. the fitness contributions of many (possibly all) the attributes are re-drawn, we get a different picture. As we have already mentioned, in a "simple" landscape with $K = 0$ all organizations quickly converge to the same configuration, which correspond to the unique global optimum and diversity dies out. If a dramatic environmental shock happens for which the global optimum moves far away from its initial position, the entire population will find itself in a low fitness area of the landscape and incumbent organizations are likely to be outperformed by newly created ones with random configuration.

On the contrary, with high $K$, to repeat, the population tends to remain distributed over a large number of local optima but the upside of all that is that with some probability a subset of the population might well find itself not too far from the high fitness portion of the new post-shock landscape: diversity helps the population adapt to dramatic environmental changes.

Levinthal's analysis has been expanded and broadened by a few works which have further studied the relationship between organizational design and environmental complexity and turbulence. Rivkin and Siggelkow (2002) (cf. also Siggelkow and Rivkin 2006) tackle the issue of multilevel organizational search by introducing an explicit representation of organizational structures in NK-type models. Decisions over the $N$ policies (bits of the string) are allocated among different departments and a superordinate CEO has the function of coordinating departmental decisions.

More in detail, each department controls a given number of policies and is engaged in increasing the fitness contribution of such policies (climbing the departmental "subscape", i.e. the landscape generated by only those policies). As – in general – any policy change in one department changes also the other departments' fitness values, each department may also attach some weight to fitness changes of other departments.

[7] There is a much more general point here related to the time scale of evolution. In many dynamics there might well be an asymptotic state which however does not have any interpretative relevance for empirical phenomena as the time of convergence is extremely long.

This weight, ranging from 0 to 1, is a model parameter which stands for the degree of "horizontal" inter-department coordination.

Finally, the organization has a CEO endowed with the power of taking the final decisions by selecting departments' proposals. For this purpose, the CEO asks each department *i* for their most preferred alternatives and selects those combination of departments' proposals which deliver the highest organizational fitness.[8]

The interplay between departments and CEO creates what the authors call a set of "sticking points", i.e. organizational configurations to which no alternative exists which can go through the approval of all subjects involved. Sticking points do not necessarily correspond to organizational local optima: first, cross-vetoes of departments and CEO can prevent also improvements which would increase the fitness of the organization and, on the other side, a department can, in some circumstances, implement a change which is beneficial for itself but not for the entire organization and therefore unlock the organization from a local optimum, if it happened to be in one.

Divergence between the set of local optima and the set of sticking points is larger when the following conditions are met:

1. decisions are allocated among a larger number of departments;
2. interdependencies among policies allocated to different departments are stronger;
3. the weight that each department attributes to other departments' fitness is lower;
4. the number of proposals the CEO receives from departments is larger and the latter give higher weight to others' fitness.

Sticking points are *competency traps* that organizations might want to escape from. One way to accomplish this is by changing the organizational structure. Siggelkow and Levinthal (2003) and Siggelkow and Levinthal (2005) analyze the performance consequences of changes in organizational structures, say from a centralized to a

[8] Some parameters $d_i$ measure the degree of CEO discretion: at one extreme, if $d_i$ is equal to one for all departments, then the CEO can automatically approve each department's most preferred alternative, without any de facto selection power. At the other extreme, if $d_i$ is equal to the number of all envisageable alternatives for all departments, then the CEO has a de facto full discretionary control over all policies.

decentralized one, or vice versa. Quite a few, albeit not all, changes tend to be beneficial to performance, not so much due to the intrinsic superior fitness of the “new” organizational form but rather to the very fact that the switch has a de-locking effect upon past “sticking points”.

The set of goals a successful organization should pursue is not limited to the broad search on the performance landscape. Once a good set of decisions has been found, stability should also be among the priorities of organizational design (cf. Rivkin and Siggelkow 2003). Moreover, firms must take into account not only the performance level they can reach, but also the speed at which they can improve on it, especially in environments that change frequently and in no predictable ways (cf. Siggelkow and Rivkin 2005).

A comprehensive analysis of differences in performance between organizational structures in terms of stability, convergence time and solution quality is offered by Mihm *et al.* (2010). In the analytical model they present, an organization is engaged in a purely decentralized search to solve a complex problem with many interdependent sub-problems.

The first crucial parameter is $b_{i,j}$, which measures how much each agent discounts the importance of the other sub-problems compared with her own one. In one extreme case, all employees take fully into account the overall performance of the firm, $b_{i,j}=1$; in the opposite case, $b_{i,j}=0$ when $i \neq j$, each employee acts myopically. The second parameter is the rate at which agents update their information about the others’ decisions, the update time being an independent Poisson process for each decision maker.

The decentralized search performs well, provided that all employees act fully holistically ($b_{i,j}=1$), and updating is immediate. What if $b_{i,j}<1$? Mimh *et al*. show analytically that in this case the probability that the search is unstable approaches 1 as $N$ grows. Intuitively, potential loops of mutual influence between agents grow up as interdependences are not taken fully into account into the decision process.

Introducing hierarchy can change the problem-solving dynamics in two different ways. First, it can give managers a veto power over the others’ decisions. When this happens, the analytical result shows that the solution quality converges monotonically to a final level even at the cost of an inferior quality.

Second, the hierarchy can change communication and influences patterns among employees. By creating departments, workers in one group may consider the performance of other groups less important, or they can be less affected by the others' decisions. If this second effect is allowed to operate, then it can be shown that there is always a choice of departments and of interdependence importance such that search progress toward a solution becomes fast and stable, again at the cost of solution quality. Basically, cycling behaviours are avoided by weakening interdepartmental interdependencies.

In the simulated model, Mihm *et al*. consider a structure in which 48 workers are grouped into six departments, which in turn are structured into two areas of three departments each; the area managers report to the CEO. The front-line workers are assumed to act holistically while differences in "myopia" are at the managers level. Moreover, updating among subgroups is delayed.

Three dimensions of decision making are analyzed. The first is the order of problem solving, which can be parallel or sequential. The second is the *locus* of decision making. The third is the structure of the hierarchy.

Simulations results can be summarised in the following way. In the first dimension, sequential search performs better in most cases while parallel search is desirable only when speed is much more important than solution quality. In the second dimension, the key result is that decisions should be delegated to the lowest level that has the information necessary to make the decision. Centralization at the lowest management level provides the same effect as full centralization, whether or not the managers act holistically. In the third dimension, the main driver of the search performance is the size and the number of departments; how the structure is organized at the intermediate levels is irrelevant for the firm's performance.

A final comment about the main building-block in NK models is in order here. In most analysis, interactions among decisions are assumed to be randomly generated. However, organizations tend to show highly *patterned* interdependences between decisions. Rivkin and Siggelkow (2007) address this issue by emphasising the implications of different *interaction patterns* in terms of the long-run value of exploration along the landscape. Different interaction patterns are modelled as *influence* *matrices* that differ with respect to the actual *arrangement* of interdependences among decisions.

Their simulation shows that even if the total number of interactions among decisions is held constant, performance landscapes can differ markedly both in the number and in the average height of the peaks they contain. The key variables to understand what drives these differences are the number of “uninfluential” decisions, that is decisions that do not affect any other decision, and the number of “uninfluenced” decisions, that is decisions that are not affected by any other decision. The presence of “uninfluential” decisions creates large smooth subspaces on each performance landscape that limit the number of local peaks; on the other hand, when many decisions are uninfluenced, it is more likely to have a handful of decisions that are very sensitive to many other choices. This creates the potential for many conflicting constraints and lots of internally consistent configurations of choices.

As local peaks proliferate, it becomes more unlikely for a searching firm to climb a high peak. This result is quite common in NK models, but Rivkin and Siggelkow show that the proliferation of local peaks comes also from the very *pattern of interactions*, and not only from the actual degrees of epistatic correlations amongst decisions. Together, the performance difference between firms engaged in *high exploration* (that is, firms that in each period try to change many decisions) and firms engaged in *low exploration* (that is, firms that in each period try to change only one decision) is increasing in the number of local peaks. That is, the more rugged the performance landscape is, the more valuable is to be engaged in radical changes.

### *3.3 Cognitive and experiential search*

Gavetti and Levinthal (2000) add another perspective to the analysis of search processes and look at the relations between forward-looking and backward-looking search and their effects on performances. The roots of the distinction between the two search processes go back to Simon (1955): the former involves cognition-ridden, forward-looking choices based on off-line evaluation of alternatives, even very distant from current behavior; the latter entails experiential choice based on on-line evaluation of a limited set of alternatives which are close to current behaviors.

In Gavetti and Levithal's model, the organization chooses a policy on the basis of a simplified and incomplete "cognitive model" of its environment, entailing "templates" which cannot directly prescribe actions. In this context, existing practices function as defaults for elements not specified by the cognitive representation and allow the identification of a specific course of action. Thus, it may happen that actors with the same cognitive template may engage in different behaviors.

These hypotheses are translated into a NK-based model in which the organization's limited cognition corresponds to a simplified representation of the fitness landscape which is assumed to be of lower dimensionality than the actual landscape ($N1<N$), even if grounded in it. This is captured by the assumption that for each point of the cognitive representation (of the perceived landscape) there are $2^{N-N1}$ points in the actual fitness landscape that are consistent with this point. The fitness value assigned to each point of the cognitive representation corresponds to the average fitness values of these $2^{N-N1}$ points.

An organization which chooses according to its cognitive representation explores *regions*, and not single points, of the landscape, while the width of these regions depends on the crudeness of the representation. When both cognitive and experiential search are at work, organization identifies a pick in its perceived N1-dimensional landscape (by cognitive or off-line search) and then explores the remaining N-N1 alternatives through a local (or on-line) search based on one bit-mutations. The role of experiential search becomes more and more important as the crudeness of the cognitive representation increases. It is important to notice here the role of the initial cognitive search in identifying the superior, on average, basins of attractions. Initial off-line search then helps in finding a good position from which the local search can start.

Gavetti and Levinthal show that in a context of competitive ecologies in which low performance organizations are selected out, organizations which adopt a joint cognitive and experiential search dominated the population. This becomes particularly evident under rugged landscapes, in which organizations which use purely experiential search are trapped into local optima.

In this framework what are the effects of adaptation through changes in the cognitive representation? Gavetti and Levinthal consider these effects both in the case of purely cognitive search and in that of joint cognitive and experiential search, also with changes in the actual fitness landscape. In the case of pure cognitive representation the organization chooses an alternative on the basis of its understanding of the payoffs as characterized by a set of N1 attributes. In this case the effects of changes in the representation depend on the complexity of the landscape (the value of *K*). If *K* is high these changes may produce good performances, as they can compensate for a poor representation of the landscape. However, if one considers organizations which use joint off-line and on-line search, the shift to a new representation may also destroy the accumulated (on-line) experience.

Changes in the representation can enhance organization's performance when the landscape itself changes as the new representation may more effectively identify new (superior) basins of attraction, and this can compensate for the loss of experiential wisdom.

However, the *locus* in which the change in the representation is decided can be crucial for the effectiveness of the change itself. Gavetti (2005) explores this linkage by situating the cognitive and the experiential search within hierarchical structures characterized by different allocations of "cognitive rights". In his NK model, each organization creates a new division engaged in a new line of business after an initial period of activity in a single line. Managers first explore the new landscape by way of local search, then decide which representation to use by comparing the representations they have in their cognitive memory with the actual payoff of the local search. The organizational hierarchy determines exactly at which level (firm or divisional) and in which way this choice is made. The performance outcome of the various organizational structures is analyzed in four contexts that differ along two dimensions. The first is the degree of economies of scope between the two divisions; the second is the heterogeneity between the problems they face.

Simulation results show that the crucial mechanism driving the difference in performances is the way in which information is processed in the exploratory phase. In particular, two properties seem to emerge. First, matching the outcome of the local search with an appropriate cognitive representation becomes more difficult as the manager in charge of decisions is higher in the hierarchy. This effect is stronger when divisions face heterogeneous problems as top managers, unlike divisional mangers, have to assess potentially contrasting action-outcome relationships. Second, a systematic bias favors signals originating in the old division: managers tend to select representations that fit the original business instead of choosing representations that capture the new domain. This effect is again increasing in the heterogeneity of the divisions' businesses.

These models shed light on the role of cognitive search both in conditioning experiential learning by constraining the local search to the most promising regions of the landscapes and in shaping organizational search under different hierarchical structure. The analysis of the interplay between the two logics of action in different contexts represents indeed a significant progress vis-à-vis representations of organizational search processes just via "one-bit mutation" search or totally random "big jumps".

Knudsen and Levinthal (2007) look at another cognitive dimension of the search process, that is the capacity of evaluating alternatives. In fact, all the models considered so far take for granted that agents are always good at comparing the outcomes of the local search, but in many task environments this might not be true. Simulation results show indeed that an imperfect evaluation of alternatives can be beneficial for organizations in that it avoids the rapid identification of the local peak within the initial basin of attraction.

A further step in the direction of opening up the "organizational problem solving black box" entails an explicit representation of organizational problem solving procedures, their emergence and their dynamics.

### *3.4 Problem solving organization and the division of labor*

Following Simon (1981), Marengo and Dosi (2005)[9] focus on strategies for the reduction of problem complexity through the division of problem solving labor, that results in the decomposition of large and complex problems into smaller sub-problems which can be solved independently. In fact, process of division of labour is a major and

[9] See also Marengo *et al.* 2000.

long neglected driving force in explaining the inner features and boundaries of economic organization. In particular, traditional organizational economics has concentrated upon the governance of transaction and contractual relations between given "technologically separable" units, but does not tackle the analysis of where such technologically separable units come from or, even more importantly, of whether organizational structures have some at all.

The issue bears a fundamental importance because, first, most processes of division of labour take place *within* organizations and, second, it empirically happens that most of the times technologies are born in a highly integrated fashion, and possibly undergo subsequent vertical disintegration both within and among firms. In other words, one could say that "in the origin there were organizations" and then markets develop along the lines defined by the processes of division of labour, rather than the other way round as postulated by transaction costs economics.

In Marengo and Dosi (2005) different organizational structures (with varying degrees of vertical integration) are compared in terms of their dynamic problem-solving properties determined by their patterns of division of labour and problem decomposition. The basic assumption is that solving a given problem requires the coordination of $N$ atomic "elements" or "actions" or "pieces of knowledge", which we can generically call components, each of which can assume some number of alternative states. The one-bit mutation algorithm at the basis of the NK model can be conceived as a particular case in which the problem is fully decomposed and the search process is fully decentralized: each sub-problem consist of a single component (bit). As showed by Kaufmann (1993), this algorithm is very quick, but it can converge only to the local optimum whose basin of attraction contain the initial configuration. On the opposite extreme, there is the case of no decomposition at all, corresponding to a strategy in which all the components (bits) are simultaneously mutated. In this case the global optimum can be reached by exploring all the possible configurations. In between there are all the other possible divisions of labor strategies.

Note that the effectiveness of the decomposition, in terms of system performances, is strongly affected by the existence of interdependences among the components of the problem: so, for example, separating interdependent components and then solving each sub-problem independently will prevent the very possibility of overall optimization. Note also that, as pointed out by Simon, because of the opaqueness of the interrelations between components, an optimal decomposition – a division of labor that separates into sub-problems only the components that are independent from each other - cannot be generally achieved by bounded rational agents, who normally are bound to aim at *near-decompositions*, that is decompositions that try to put together within the same sub-problem only those components whose interdependences are "more important" for the performance of the system.

Finally note that the search space is not given exogenously, but is constructed by agents that possess subjective representations of the structure of the problem. In that the distance between the real structure of the problem (its real decomposition) and the subjective representation that agents have of it has a dramatic effect on problem solving outcomes.

More formally, one can characterize a problem by the following elements:

The set of components: $C=\{c_1, c_2, \ldots, c_N\}$, where each component can take one out of a finite number of states. Normally, a binary set of components is assumed for simplicity: $c_i \in \{0,1\}\ \forall i$.

A configuration, that is a possible solution to the problem: $x^i = c_1^i c_2^i \ldots c_N^i$.

The set of configurations: $X = \{x^1, x^2, \ldots, x^{2^N}\}$.

An ordering over the possible configurations: $x^i \geq x^j$ (*or* $x^i > x^j$) holds whenever $x^i$ is weakly (or strictly) preferred to $x^j$.

A problem is fully defined by the pair $(X, \geq)$.

As the size of the set of configurations is exponential in the number of components, whenever the latter is large, the state space of the problem becomes much too vast to be extensively searched by agents with bounded computational capabilities. One way of reducing its size is to decompose[10] it into sub-spaces.

[10] A decomposition can be considered as a particular case of search heuristics: search heuristics are, in fact, ways of reducing the number of configurations to be considered in a search process.

Let $I=\{1,2,\ldots,N\}$ be a set of indexes and let a block[11] $d_i \subseteq I$ be a non-empty subset of it; the size of block $d_i$ is its cardinality $|d_i|$. A decomposition of the problem $(X, \geq)$ is defined as a set of blocks:

$$D=\{d_1,d_2, \ldots ,d_k\} \text{ such that } \bigcup_{i=1}^{k} d_i = I .$$

Note that a decomposition does not necessarily have to be a partition (that is the intersection between two decompositions need not be the empty set).

Given a configuration $x^i$ and a block $d_j$, the block-configuration $x^i(d_j)$ is the substring of length $|d_j|$ containing the components of configuration $x^i$ belonging to block $d_j$:

$$x^i(d_j) = x^i_{j_1} x^i_{j_2} \ldots x^i_{j_{|d_j|}} \qquad \forall j_h \in d_j .$$

The notation $x^i(d_{-j})$ is used to indicate the substring of length $N$-$|d_j|$ containing the components of configuration $x^i$ not belonging to block $d_j$.

Two block-configurations can be joined into a larger block-configuration by means of the $\vee$ operator so defined: $x(d_j) \vee y(d_h) = z(d_j \cup d_h)$ where $z_k = x_k$ if $k \in d_j$ and $z_k = y_k$ if $k \in d_h$.

The size of a decomposition is defined as the size of its largest defining block:

$$|D|= \max\{|d_1|,|d_2|, \ldots, |d_k|\}.$$

Coordination among blocks in a decomposition may either take place through market-like mechanisms or via other organizational arrangements (e.g. hierarchies). Dynamically, when a new configuration appears, it is tested against the existing one according to its relative performance. The two configurations are compared in terms of their ranks and the superior one is selected, while the other one is discarded.

More precisely, let us assume that the current configuration is $x^i$ and take block $d_h$ with its current block-configuration $x^i(d_h)$. Let us now consider a new configuration $x^j(d_h)$ for the same block, if:

$$x^j(d_h) \vee x^i(d_{-h}) \geq x^i(d_h) \vee x^i(d_{-h})$$

[11] Blocks in our model can be considered as a formalization of the notion of modules used by the flourishing literature on modularity in technologies and organizations (Baldwin and Clark, 2000) and decomposition schemes are a formalization of the notion of system architecture which defines the set of modules in which a technological system or an organization are decomposed. We will come back to modularity literature later on.

then $x^j(d_h)$ is selected and the new configuration $x^j(d_h) \vee x^i(d_{-h})$ is kept in place of $x^i$, otherwise $x^j(d_h)$ is discarded and $x^i$ is kept.

It might help to think in terms of a given division of labor structure (the decomposition scheme) within firms, whereby individual workers and organizational sub-units specialize in various segments of the production process (a single block). Decompositions, however, sometimes determine also the boundaries across independent organizations specialized in different segments of the whole production sequence.

Note that, dynamically, different inter-organizational decompositions entail different degrees of decentralization of the search process. The finer the inter-organizational decompositions, the smaller the portion of the search space which is being explored by local variational mechanisms and tested (however indirectly) by market selection. Thus there is inevitably a trade-off: finer decompositions and more decentralization make search and adaptation faster (if the decomposition is the finest, search time is linear in N), but on the other hand, the process explores smaller and smaller portions of the search space, thus decreasing the likelihood that optimal (or even "good") solutions are ever generated and tested.

Decompositions are sorts of templates ("categorizations" in the "mental models" perspective) which determine how new configurations are generated and can be tested afterward by the selection mechanism. In large search spaces in which only a very small subset of all possible configurations can be generated and undergo testing, the procedure employed to generate such new configurations plays a key role in defining the set of attainable final configurations.

Marengo and Dosi assume that boundedly rational agents can only search locally in directions which are given by the decomposition: new configurations are generated and tested in the neighborhood of the given one, where neighbors are new configurations obtained by changing some (possibly all) components within a given block.

Given a decomposition $D=\{d_1,d_2, \ldots ,d_k\}$, a configuration $x^i$ is a preferred neighbor or simply a neighbor of configuration $x^j$ with respect to a block $d_h \in D$ if the following three conditions hold:

1. $x^i \geq x^j$
2. $x_k^i = x_k^j \quad \forall k \notin d_h$

3. $x^i \neq x^j$.

Conditions 2 and 3 require that the two configurations differ only by components which belong to block $d_h$. According to the definition, a neighbor can be reached from a given configuration through the operation of a single decentralized coordination mechanism.

The set of neighbors of a configuration $x$ for block $d_i$ is called $H_i(x,d_i)$.

The set of best neighbors $B_i(x,d_i) \subseteq H_i(x,d_i)$ of a configuration $x$ for block $d_i$ is the set of the most preferred configurations in the set of neighbors:

$$B_i(x,d_i)=\{y \in H_i(x,d_i) \text{ such that } y \geq z \quad \forall z \in H_i(x,d_i)\}$$

By extension from single blocks to entire decompositions, the definition of the set of neighbors for a decomposition is:

$$H(x,D) = \bigcup_{i=1}^{k} H_i(x,d_i).$$

Here the configuration is a local optimum for the decomposition $D$ if there does not exist a configuration $y$ such that $y \in H(x,D)$ and $y > x$.

A search path or, for short, a path $P(x^i,D)$ from a configuration $x^i$ and for a decomposition $D$ is a sequence, starting from $x^i$, of neighbors:

$$P(x^i,D)=x^i,x^{i+1},x^{i+2},\ldots \text{ with } x^{i+m+1} \in H(x^{i+m},D).$$

A configuration $x^j$ is reachable from another configuration $x^i$ and for decomposition $D$ if there exists a path $P(x^i,D)$ such that $x^j \in P(x^i,D)$.

Suppose configuration $x^j$ is a local optimum for decomposition $D$: the basin of attraction of $x^j$ for decomposition $D$ is the set of all configurations from which $x^j$ is reachable:

$$\Psi(x^j,D)=\{y, \text{ such that } \exists\, P(y,D) \text{ with } x^j \in P(y,D)\}.$$

Now let $x^0$ be the global optimum and let $Z \subseteq X$ with $x^0 \in Z$. We say that the problem $(X,\geq)$ is locally decomposable in $Z$ by decomposition $D$ if $Z \subseteq \Psi(x^0,D)$. If $Z=X$, we say that the problem is globally decomposable by decomposition $D$.

The perfect decomposability requirement can be softened into one of near-decomposability: the problem is no longer required to be decomposed into completely separated sub-problems, i.e. sub-problems which fully contain all interdependencies, but it can be sufficient to find sub-problems which contain the most relevant interdependencies, while less relevant ones can persist across sub-problems. In this way, optimizing each sub-problem independently will not necessarily lead to the global optimum, but to a "good" solution. In other words, near-decompositions give a precise measure of the trade-off between decentralization and optimality: higher degrees of decentralization, while generally displaying a higher adaptation speed, are likely to be obtained at the expense of the asymptotic optimality of the solutions which can be reached.

As a consequence, Marengo and Dosi arrange all the configurations in $X$ by descending rank $X=\{x^0, x^1, x^2, \ldots\}$ where $x^i \geq x^{i+1}$, and $X_\mu = \{x^0, x^1, \ldots, x^{\mu-1}\}$ is the ordered set of the best μ configurations. $X_\mu$ is said to be reachable from a configuration $y \notin X_\mu$ and for decomposition $D$ if there exists a configuration $x^i \in X_\mu$ such that $x^i \in P(y,D)$.

The basin of attraction $\Psi(X_\mu, D)$ of $X_\mu$ for decomposition $D$ is the set of all configurations from which $X_\mu$ is reachable. If $\Psi(X_\mu, D) = X$, $D$ is a μ-decomposition for the problem. μ-decompositions of minimum size can be found with an algorithm which computes minimum size optimal decompositions.

It is straightforward to show that as μ increases one can generally find finer near-decompositions. This shows that the organizational structure sets a balance in the trade-off between search and adaptation speed and optimality. It is easy to argue that in complex problem environments, characterized by strong and diffused interdependencies, such a trade-off will tend to produce organizational structures which are more decomposed and decentralized than what would be optimal given the interdependencies of the problem space.

Different organizational forms implement different decomposition heuristics and might be characterized by different representations of the problem and therefore present different properties in terms of the effectiveness and efficiency of the derived search processes (cf. Marengo, Pasquali and Valente (2005) for a theoretical discussion of the topic). In particular a trade-off exists between complexity and optimality: a finer decomposition makes search faster, but the exploration of smaller portion of the search space reduces the likelihood to generate and then select an optimal solution. The application of these ideas to organizational design leads to the comparison, in terms of relative performance, between not decomposed tasks (organization-embodied) and decomposed tasks (coordinated via market-like mechanism or via simple organizations structured as sets of perfectly independent tasks). One of the main conclusions is that the advantages of decentralization (faster adaptation) usually imply a cost in terms of sub-optimality (impossibility to reach global optima). This casts strong doubts on the efficacy of market selection processes as substitutes for individual optimization: selection is not able to select out sub-optimal features nor to select for optimal ones if both are somehow complementary to each-other in actual organizations and technologies.

### *3.5 Modelling the coupling mechanisms between capabilities and governance*

Marengo and Dosi (2005), as well as most of contributions of this genre, while concentrating on the problem-solving features of organizational dynamics, censor any incentive compatibility issue. An attitude that, as noted above, is quite typical within the capability-based framework.

There is nothing, however, preventing this type of analysis to go beyond the exclusive focus on firms as *loci* of coordination and as *loci* of creation, implementation, storage and diffusion of productive knowledge[12] and explicitly take on board the issues of incentive governance and control discussed qualitatively in Coriat and Dosi (1998). Attempts in this direction are formal analyses by Dosi, Levinthal and Marengo (2002; 2003) which incorporate issues of conflict of interests, power and control over agents' decisions within the analytical framework of Marengo and Dosi (2005) and Marengo *et al.* (2000) and discuss the interaction between problem representation and incentive mechanisms. In particular, the double role of problem representation is stressed: on the one hand it defines the "cognitive" structure of the problem and the consequent decomposition which is adopted (definition of teams as subsets or blocks of components); on the other hand, it has important consequences for a reward mechanism

[12] A more complete "co-evolutionary" picture is discussed by Dosi (1995). Organizations are assumed to be characterized by six correlated dimensions: the distribution of formal authority; the distribution of power; the incentive structure; the structure of information flows; the distribution of knowledge and competence. In this context organization dynamics can be conceived as a process of adaptation and selection according to multiple, and possibly conflicting, objectives.

based on the distinction between organization's (system) and team's (block) performance as it defines what organization conceives as a team.

The analysis starts by considering the conflicts of interest among problem solving teams generated by the adoption of team-level incentive mechanisms. While under a global reward an alternative (a particular configuration of sub-problem's components) is selected if it improves the overall organization's performance, with a team-level reward mechanism a would-be alternative is accepted if it enhances the performance of the unit even if it degrades the overall organization's performance. It can be shown that if the organization's representation of the problem is not correct (it does not correspond to the right structure of the problem in terms of interrelations among components) the adoption of a global reward allows the organization to reach a global optimum. But what is more interesting is that, even if the representation of the problem is not correct, the adoption of a team-level reward structure tends, in the long run, to produce performances that are similar to the global- reward one. Thus, goal conflicts prevent the organization to remain absorbed in local optima and act as substitute for a correct representation of the problem (Dosi, Levinthal and Marengo, 2002).

Power is introduced by allowing one team (a block in the decomposition) to stop the mutation of any other blocks that decreases its own performance (veto power). The evidence suggests that, under specific conditions, the adoption of such a mechanism lead to good solutions. In particular, a team reward scheme with veto power is superior to the global reward structure when the organizational representation of the problem is based on a finer decomposition than the real one and the latter is not too complex. This is due to the fact that veto power interrupts the cycling among possible solutions generated by a team-based reward structure preserving the advantages in terms of greater search effort which are typical of this reward mechanism.

A principal-agent-like model of interaction is reproduced considering the case of control over the decisions of other organizational members by a principal, the residual claimant of the total payoff, who can "order" others to keep performing a given action or to switch to a different one. This activity is considered to have a cost which depends on the span of control, i.e. the dimension of each sub-unit, and it is higher when the principal wants to induce a change in agent's action than when he wants to elicit the same behaviour (the principal's profit is defined as the total output of the organization minus the "elicitation cost"). When actions are interdependent, the control function, as any other problem-solving activity, cannot be entirely decomposed. Thus, the interaction between a cognitive dimension and a control dimension has to be considered. The effects on total performance and the principal's profit are analyzed considering four different cases: right, almost right, wrong and minimal (one-component units) perceived decomposition by agents, with reference to different decompositions of the underlying problem and the "correctness" of the decomposition itself.

Obviously if the organizational decomposition is the "true" one, perfectly knowledgeable agents not facing any incentive compatibility problem would make costly control redundant. However, interestingly, when the organization has a wrong representation of the problem space (and in particular underestimates the span of interdependencies), agents subject to costly control may generate a better performance than the one produced by perfectly 'cooperative' agents.

Finally Dosi, Levinthal and Marengo (2002) analyze more explicitly the double role of problem representation. The work examines, in particular, by means of a simulation model, the relations between cognitive decompositions and operational decompositions. The former establish search heuristics and targets, whereas the latter implement search processes driven by those targets. The exercise shows that if cognitive decompositions are correct then it is efficient to have maximum division of labor at the operational level, as this increases speed and accuracy of adaptation to targets. On the contrary, if cognitive decompositions do not correspond to the "true" ones, coarser division of labor at the operational level ensures less accurate but prompt adaptations to the imperfectly set target.

### *3.6 Modularity and organizational architecture*

The existence of different organizational decompositions and hierarchies poses a problem: to what extent can boundedly rational agents identify the true structure? Ethiraj and Levinthal (2004a) address this question within a NK model in which organizations are characterized along the two dimensions of "decomposability" and “hierarchy”. In a loosely (tightly) coupled organization, there are few (many) interdependencies between departments; in the second dimension, the organization is said to be “hierarchical” if the structure of interdependencies between departments is unidirectional (e.g., the decisions of the first department influence the decisions of the second department, but not the other way around).

In the model, there are five relevant aspects that need to be specified. The first is the generative structure that can be of four different types, depending on the characteristics along the two dimensions. The second is the boundedly rational second-order adaptation that concerns the organization design. Managers are assumed to have control over the number of departments and the assignment of functions to them. In particular, they can split an existing department into two or more new departments, can combine two or more into one or reallocate functions among them.

The third aspect is the first-order adaptation which corresponds to the usual one-bit mutation implemented simultaneously in each of the departments. The fourth is the environmental change which is modelled as a random change in the generative structure that affects the coupling of decision both within and between departments. The final modelling specification concerns the selection mechanism. Here, the probability that an organization will be selected is proportional to its performance.

Simulation results show that hierarchical structures are always able to converge to the generative structure when managers are engaged in second order adaptation. On the contrary, non-hierarchical structures never manage to reach a stable state. Non-hierarchical and loosely coupled structures continue to exist in six different forms at the end of the experiment; non-hierarchical and tightly coupled structures preserve the initial heterogeneity throughout the experiment, suggesting that second-order adaptation is relatively ineffective.

When an environmental change occurs, the effectiveness of second-order adaptation is considerably reduced, but the ranking between organizations is the same as before. Non-hierarchical and tightly coupled organizations are the bad performers also in this case.

Ethiraj and Levinthal address also the question of complementarity of first-order and second order adaptation with results that are similar to the ones presented above. When there is a simultaneous process of first-order and second-order adaptation, all organizational structures perform better with respect to the setting in which only first-order adaptation is on the stage. In the ranking, the bad performer is, once again, the non-hierarchical and tightly coupled organization.

Ethiraj and Levinthal (2004b) use a similar setup in order to answer the following question: given a "true" structure and supposing that boundedly rational agents are unable to uncover it, is it better to "over-modularize" or to "under-modularize"? In particular, how does over- or under-modularization affect local search and module recombination performance over time?

Simulations show that when firms are engaged only in local search, the effectiveness of innovation is lower the greater the deviation of the design structure from the true underlying structure. More interesting, Ethiraj and Levinthal find that, in the long run, erring on the side of greater integration poses lower performance penalties than erring on the side of greater modularity.

When module recombination is allowed to operate, things get more complicated. With recombination but no local search, over-modularization gives more benefits than under-modularization. Recombination helps firms to avoid local peaks, but as the size of each module gets larger, a greater number of decision choices will be replaced and the probability of incorrect changes becomes higher. When recombination and local search are allowed to interact, under-modularizing is again a better strategy than over-modularizing. With few modules, local search compensates for the poor performance of recombination; with many modules, local search appears to counteract the effect of recombination.

## 4. Conclusions

Parallel to the qualitative analyses of organizations as structured bundles of problem-solving capabilities (for a critical review of the literature cf. Dosi, Faillo and Marengo, 2008), a growing number of contributions have begun to offer formal accounts of such organizational properties and their dynamics. The formal instruments are diverse: they include NK models representing organizations as ensembles of interrelated "traits" mapping into some overall environmental fitness of the firm; classifiers system representations of the problem-solving procedures triggered by diverse internal or environmental states; decomposition schemes of Simonian ascendancy allowing the analysis of the performance properties of different “representations” in the problem-solving space and different patterns of division of cognitive and operational labour.

The formal modelling of organizations as problem-solving entities bears important consequences also in terms of the theory of production and technology. In fact, the problem-solving activity conceived of as combinations of physical and cognitive acts, within a procedure, leading to the achievement of a specific outcome is quite near to a representation of *technology in action* conceived as a *recipe* or a *procedure*. However, the properties of such representation into the (lower dimensional) space of input/output relations is till underanalyzed, an exception being Auerswald *et al*. (2000) who offer a promising example of the use of the apparatus of NK models to study the microeconomic theory of technological evolution (cf. also the discussion in Dosi and Grazzi 2006 and Dosi and Nelson 2010).

The results begin to highlight important comparative properties regarding, among other, the impact on problem-solving efficiency and learning of different forms of hierarchical governance, the dangers of lock-in associated with specific forms of adaptive learning, the relative role of “online” vs. “offline” learning, the impact of the “cognitive maps” which organizations embody, the possible trade-offs between accuracy and speed of convergence associated with different “decomposition schemes”, the (ambiguous) role of organizational memory in changing environments.

In a nutshell, one has finally begun to develop formal instruments allowing exercises of comparative institutional analysis (cf. Aoki, 2001), focusing on the distinct properties of different forms of organization and accumulation of knowledge. It is a work which is only at its exciting start.

## References

Aoki, M. (2001), *Toward a Comparative Institutional Analysis*. Cambridge, Ma: MIT Press.

Auerswald, P., S. Kauffman, J. Lobo and K. Shell (2000), “The Production Recipes Approach to Modelling Technological Innovation: an Application to Learning by Doing”, *Journal of Economic Dynamics & Control*, 24, 389-450.

Augier, M. and J. G. March (eds.) (2002), *The Economics of Choice, Change and Organization: Essays in Memory of Richard M. Cyert.* Cheltenham, UK and Northampton, MA: Edward Elgar.

Baldwin, C. Y. and K. B. Clark (2000), *Design Rules*. Cambridge, MA: MIT Press.

Becker, M. (2004), “Organizational Routines: a Review of the Literature”, *Industrial and Corporate Change*, 13, 643-678.

Chandler, A., P. Hagström and Ö. Sölvell (eds.) (1998), *The Dynamic Firm*. Oxford and New York: Oxford University Press.

Chandler, A. (1977), *The Visible Hand: the Managerial Revolution in American Business*. Cambridge, MA: Harvard University Press.

Cohen, M. D., J. G. March and J. P. Olsen (1972), “A Garbage Can Model of Organizational Choice”, *Administrative Science Quarterly*, 17, 1-25.

Cohen, M. D., R. Burkhart, G. Dosi, M. Egidi, L. Marengo, M. Warglien, and W. Sydney (1996), “Routines and Other Recurring Action Patterns of Organizations: Contemporary Research Issues”, *Industrial and Corporate Change*, 5, 653-698.

Coriat, B. and G. Dosi (1998), "Learning How to Govern and Learning How to Solve Problems: On the Co-Evolution of Competences, Conflicts and Organizational Routines", in Chandler *et al*.

Cyert, R. M. and J. G. March (1963), *A Behavioral Theory of the Firm*. Cambridge: Blackwell.

Dosi, G. (1995), "Hierarchies, Markets and Power: Some Foundational Issues on the Nature of Contemporary Economic Organization", *Industrial and Corporate Change*, 4, 1-19.

Dosi, G., M. Faillo and L. Marengo (2008), "Problem-Solving and Governance in the Capability-based view of the Firm: The Roles and Theoretical Representations of Organizational Routines", in M. Becker (ed.), *Handbook of Organizational Routines*. Cheltenham, UK and Northampton, MA: Edward Elgar.

Dosi, G. and M. Grazzi (2006), "Technologies as Problem-Solving Procedures and Technologies as Input-Output Relations: Some Perspectives on the Theory of Production", *Industrial and Corporate Change*, 15, 173-202.

Dosi, G., D. Levinthal and L. Marengo (2002), *The Uneasy Organizational Matching Between Distribution of Knowledge, Division of Labor and Incentive Governance*, LEM Working Papers, Sant'Anna School of Advanced Studies, Pisa.

Dosi, G., D. Levinthal and L. Marengo (2003), "Bridging Contested Terrain. Linking Incentive-Based and Learning Perspectives on Organizational Evolution", *Industrial and Corporate Change*, 12, 413-436.

Dosi, G. and L. Marengo (1995), “Toward a Theory of Organizational Competencies”, in R. W. England (ed.), *Evolutionary Concepts in Contemporary Economic*s. Ann Arbor: Michigan University Press.

Dosi, G. and L. Marengo (2000), “On the Tangled Discourse between Transaction Cost Economics and Competence-Based Views of the Firm”, in N. Foss and V. Mahnke (eds.), *Competence, Governance and Entrepeneurship*. Oxford: Oxford University Press.

Dosi, G., L. Marengo, E. Paraskevopoulou and M. Valente (2011), *The value and Dangers of Remembrance in Changing Worlds: a Model of Cognitive and Operational Memory of Organizations*, in progress, LEM, Sant’Anna School of Advanced Studies, Pisa.

Dosi, G. and R. Nelson (2010), “Technical Change and Industrial Dynamics as Evolutionary Processes”, in B. Hall and N. Rosenberg (eds.), *Handbook of Innovation*.

Dosi, G., R. Nelson and S. G. Winter (eds.) (2000), *The Nature and Dynamic of Organizational Capabilities*. Oxford and New York: Oxford University Press.

Dosi, G. and S. G. Winter (2002), “Interpreting Economic Change: Evolution, Structures and Games”, in M. Augier and J. G. March (eds.).

Ethiraj, S. and D. Levinthal (2004a), “Bounded Rationality and the Search for Organizational Architecture: An Evolutionary Perspective on the Design of Organizations and Their Evolvability”, *Administrative Science Quarterly*, 49, 404-437.

Ethiraj, S. and D. Levinthal (2004b), “Modularity and Innovation in Complex Systems”, *Management Science*, 50, 159-173.

Ethiraj, S., D. Levinthal and R. Roy (2008), “The Dual Role of Modularity: Innovation and Imitation”, *Management Science*, 54, 939-955.

Gavetti, G. and D. Levinthal (2000), “Looking Forward and Looking Backward: Cognitive and Experimental Search”, *Administrative Science Quarterly*, 45, 113–137.

Gavetti, G. (2005), “Cognition and Hierarchy: Rethinking the Microfoundations of Capabilities’ Development”, *Organization Science*, 16, 599-617.

Kauffman, S. A. (1993), *The Origins of Order*. Oxford: Oxford University Press.

Knudsen, T. and D. A. Levinthal (2007), “Two Faces of Search: Alternative Generation and Alternative Evaluation”, *Organization Science*, 18, 39-45.

Kreps, D. M. (1990), “Corporate Culture and Economic Theory” in J. E. Alt and K. A. Sheplsle (eds.), *Perspectives on Positive Political Economy*. Cambridge: Cambridge University Press.

Lazaric, N. and B. Denis (2005), “Routinization and Memorization of Tasks inside a Worskshop: An Illustration through a Case Study”, *Industrial and Corporate Change*, 14, 873-896.

Levinthal, D. (1997), “Adaptation on Rugged Landscapes”, *Management Science*, 43, 934-950.

Levinthal, D. (1998), "The Slow Pace of Rapid Technological Change: Gradualism and Punctuation in Technological Change", *Industrial and Corporate Change*, 7, 217-247.

March, J. G. and H. A. Simon (1993), *Organizations*, 2nd edn. New York: Wiley.

Marengo, L. (1992), "Coordination and Organizational Learning in the Firm", *Journal of Evolutionary Economics*, 2, 313-326.

Marengo, L. (1996), "Structure, Competence and Learning in an Adaptive Model of the Firm", in G. Dosi and F. Malerba (eds.), *Organization and Strategy in the Evolution of the Enterprise*. London: MacMillan.

Marengo, L., G. Dosi, P. Legrenzi and C. Pasquali (2000), "The Structure of Problem-Solving Knowledge and the Structure of Organizations", *Industrial and Corporate Chang*e, 9, 757-788.

Marengo, L. and G. Dosi (2005), "Division of Labor, Organizational Coordination and Market Mechanisms in Collective Problem-solving", *Journal of Economic Behavior and Organization*, 58, 303-326.

Marengo, L., C. Pasquali and M. Valente (2005), "Decomposability and Modularity of Economic Interactions", in W. Callebaut and D. Rasskin-Gutman (eds.), *Modularity: Understanding the development and evolution of complex natural systems*. Cambridge, MA: MIT Press.

Mihm, J., C. H. Loch, D. Wilkinson and B. A. Huberman (2010), "Hierarchical Structure and Search in Complex Organizations", *Management Science*, 56, 831-848.

Narduzzo, A., E. Rocco and M. Warglien (2000), "Talking about Routines in the Field", in G. Dosi, R. Nelson and S. G. Winter (eds.), *The Nature and Dynamics of Organizational Capabilities*. Oxford: Oxford University Press.

Nelson, R. (1991), "Why Do Firms Differ and How Does it Matter?", *Strategic Management Journal*, 12, 61-74.

Nelson, R. (1994), "The Role of Firm Differences in an Evolutionary Theory of Technical Advance", in Magnusson (ed.), *Evolutionary and Neo-Schumpeterian Approaches to Economics*. Boston: Kluwer.

Nelson, R. (2008), "Why Do Firms Differ and How Does It Matter? A Revisitation", *Seoul Journal of Economics*, 21, 607-619.

Nelson, R. and S. G. Winter (1982), *An Evolutionary Theory of Economic Change*. Cambridge, MA: The Belknap Press of Harvard University Press.

Oswald, J. and M. Craven (2001), "Beyond the Routine: Innovation Management and the Teaching Company Scheme", *Technovation*, 21, 267-279.

Page, S. E. (1996), "Two Measures of Difficulty", *Economic Theory*, 8, 321–346.

Pavitt, K. (2002), "Innovating Routines in the Business Firm: What Corporate Task Should They Be Accomplishing?", *Industrial and Corporate Change*, 11, 117-133.

Pentland, B. T. and H. H. Reuter (1994), "Organizational Routines as Grammars of Action", *Administrative Science Quarterly*, 39, 484-510.

Rivkin, J. W. and N. Siggelkow (2002), "Organizational Sticking Points on NK Landscapes", *Complexity*, 7, 31-43.

Rivkin, J. W. and N. Siggelkow (2003), “Balancing Search and Stability: Interdependencies Among Elements of Organizational Design”, *Management Science*, 49, 290-311.

Rivkin, J. W. and N. Siggelkow (2007), “Patterned Interactions in Complex Systems: Implication for Exploration”, *Management Science*, 53, 1068-1085.

Siggelkow, N. and D. A. Levinthal (2005), “Escaping Real (Non-Benign) Competency Traps: Linking the Dynamics of Organizational Structure to the Dynamics of Search”, *Strategic Organization*, 3, 85-115.

Siggelkow, N. and D. A. Levinthal (2003), “Temporarily Divide to Conquer: Centralized, Decentralized, and Reintegrated Organizational Approaches to Exploration and Adaptation”, *Organization Science*, 14, 650-669.

Siggelkow, N. and J. W. Rivkin (2005), “Speed and Search: Designing Organizations for Turbulence and Complexity” *Organization Science*, 16, 101-122.

Siggelkow, N. and J. W. Rivkin (2006), “When Exploration Backfires: Unintended Consequences of Multilevel Organizational Search”, *Academy of Management Journal*, 49, 779-795.

Simon, H. A (1962), “The Architecture of Complexity”, *Proceedings of the American Philosophical Society*, 106, 467-482.

Simon, H. A. (1981), *The Science of Artificial*, 2nd Ed.. Cambridge, MA: MIT Press.

Schumpeter, J. A. (1950), *Capitalism, Socialism and Democracy*, New York: Harper.

Smith, A. (1776), *An Inquire into the Nature and Causes of the Wealth of Nations*. London: W. Stahan and T. Cadell.

Teece, D. J., G. Pisano and A. Shuen, (1997), "Dynamic Capabilities and Strategic Management", *Strategic Management Journal*, 18, 509-33; a revised version is published in G. Dosi, R. Nelson and S. G. Winter (eds.) (2000).

Walsh, J. P. and G. R. Ungson (1991), "Organizational Memory", *The Academy of Management Review*, 16, 57-91.

Winter S. (1987), "Knowledge and Competence as Strategic Assets", in D. J. Teece (ed.), *The Competitive Challenge: Strategies for Industrial Innovation and Renewal*. Cambridge, M.A.: Ballinger.

Winter, S. G. and G. Szulanski (1998), "Replication as Strategy", *Organizational Science*, 12, 730-743.

Winter, S. G. and G. Szulanski (2002), "Replication of Organizational Routines: Conceptualizing the Exploitation of Knowledge Assets", in N. Bontis and C. W. Choo (eds.) *The Strategic Management of Intellectual Capital and Organizational Knowledge: A Collection of Readings*. New York: Oxford University Press.

Winter, S. G. (2005), "Toward and Evolutionary Theory of Production", in K. Dopfer (ed.), *The Evolutionary Foundations of Economics*. Cambridge: Cambridge University Press.

Zander, U. and B. Kogut (1995), "Knowledge and the Speed of Transfer and Imitation of Organizational Capabilities: an Empirical Test", *Organization Science*, 6, 76-92.

Zollo, M. and S. G. Winter (2002), “Deliberate Learning and the Evolution of Dynamic Capabilities”, *Organization Science*, 13, 339-351.