QUADERNI

Università degli Studi di Siena
DIPARTIMENTO DI ECONOMIA POLITICA

MAURO CAMINATI

Knowledge Growth, Complexity and the Returns to R&D

n. 386 – Giugno 2003

Abstract - This paper is concerned with the way in which R&D activity in the technological and scientific domains feeds back into the dimension, the hierarchic structure and the complexity of knowledge search spaces. The discussion sets the stage for a critical evaluation of recent contributions trying to identify foundations for the existence of laws of returns to R&D.

JEL classification: O3, D83.

I whish to thank Carlo Zappia for helpful comments. This paper is the first output of an ongoing research on structural change and technological progress carried out at Siena University under the national project on Growth and Distribution (prot. 2002131335_008) financed by Siena University and the Italian MIUR: COFIN 2002.

Mauro Caminati, Dipartimento di Economia Politica, Università degli Studi di Siena

1. Introduction

A salient feature in the recent upsurge of formal models of aggregate long-run economic growth is that technological knowledge is considered a producible good, which much like any other good, is subject to specific laws of returns to the size of the factors participating in the production thereof. Questionable as it may be, this feature responds to the need of giving synthetic formal expression to hypotheses concerning the laws of returns to knowledge production, so that their ultimate consequences are brought under analytic control[1].

In the attempt to delve more deeply into the foundations of knowledge production, and their relation with growth theory, Weitzman (1996, 1998) and Olsson (2000) provide a formal characterisation of alternative schemes which may sustain the expansion of the knowledge stock.

Weitzman's basic premise is that technological ideas proliferate through recombination[2]. New and potentially useful seed-ideas *S* available for further development grow out of the untried and successful hybridisation of existing ideas *A*. If *H(t)* is the number of new hybrids at *t*, then $S(t) = H(t)p(t)$ where *p(t)* is the hybridisation-success probability at *t*. New hybrid ideas *H(t)* arise out of the untried pair-wise combination of existing ideas[3]. Development of seed ideas *S* into new ideas *B* occurs under the constant returns technology $B = \Phi(S, J)$ where *J* is final output invested in R&D and $A(t+1) = A(t) + B(t)$. Weitzman's main result is that provided the probability *p(t)* does not decline too fast with the growth of *A(t)*, the growth of ideas will be ultimately limited only by the amount of final output to be invested in R&D and not by the availability of potentially useful ideas. It can be readily understood and is in fact shown by Weitzman how his picture of knowledge production "might serve as the core of an endogenous theory of economic growth".

Weitzman's somewhat optimistic conclusions are challenged by Olsson (2000) and Olsson and Frey (2001). Though accepting the basic premise of knowledge growth through re-combination of existing ideas, these authors confine the domain of this combinatorial process to subsets of *compatible ideas*. The technological opportunity set *Z(t)* is the set within which all potentially useful *new* combinations of ideas in the current knowledge set *A(t)* are to be found. Through a formal representation of ideas as points in a Euclidean metric space of finite dimension *n*, Olsson (2000) argues that normal science and incremental

[1] The standard replication argument refers to non-knowledge-goods production for a given size of the 'knowledge stock'. The argument implies increasing returns in this production to the joint scale of non-knowledge goods, labour and 'knowledge'. On the further simplifying premise that the contribution of non-knowledge goods to 'knowledge production' is negligible, a crucial issue is whether the latter is subject to constant or decreasing returns to the 'knowledge input'. If constant returns obtain, then steady-state growth depends on prevailing investment-output ratios; if decreasing returns obtain, steady-state per-capita output growth may be strictly positive, but is independent of the investment-output ratio; for a generalization see Eicher and Turnovsky (1999) and Caminati (2003). Caballero and Jaffe (1993), C. Jones (2002), Porter and Stern (2000), and many others, discuss some of the issues involved in the choice of different specifications for the technology of the R&D sector. Admittedly, their discussion is only tentative, reflecting the poor understanding of the grounds for bringing knowledge growth under the discipline of law-like regularities.

[2] He refers back to Poincarè (1908) and Hadamard (1949) rational reconstructions of mathematical creation, to anecdotal evidence on Edison's invention of the electric candle, to Usher (1927) and Schumpeter (1934).

[3] More formally, $H(t) = C_2(A(t)) - C_2(A(t-1))$ where $C_2(A)$ is the set of pair-wise combination of the elements in *A*.

R&D expand the knowledge set A through *convex* combinations of 'close' ideas on the boundary of A. In this representation, technological opportunities arise out of residual regions of non-convexity that may survive near this boundary. Once A is fully convexified, the technological opportunity set becomes empty. In other words, far from leading to an expansion of Z, the growth of A takes place *at the expense* of Z; thus, Weitzman's hybridisation-success probability may well decline very fast down to zero along the terminal phase of a technological trajectory. Scientific revolutions and paradigm shifts are the only means to exit states of technological stand-still; they are modelled as global, discontinuous changes in the knowledge set giving rise to new areas of non-convexity. The arrival of such revolutions is treated as a random variable and is not investigated further.

From the view-point of this paper, the main problem with Ollson (2000) and Olsson and Frey (2001) formalisation[4] is that it merely represents the notion of knowledge expansion through re-combination of compatible ideas, but, as it stands, it gives no clues on the existence of structural factors behind the notion of compatibility and its degree and to study the plausible evolution of these factors over long time spans. Put differently, their analysis is of little help in the search for 'general rules of structure and change' (Gould, 1992) which may be relevant to a theory of invention and innovation. These limits can be traced to the fact that the formalisation is not suitable to study how R&D output feeds back into the complexity of knowledge search spaces, on the one hand, and into the technology and organization of R&D on the other.

An ever larger number of contributions in the theory of innovation has come to recognise how important clues in this direction are offered by Kauffman's (1993) model of evolution on *N-K* fitness landscapes. Here the knowledge set A is not embedded in the Euclidean space $\Re^n$, but in the space of strings consisting of N binary ordered components. In this representation an idea is a string $\boldsymbol{x} \in \{0,1\}^N$. Complementarity within a given technology and between different technologies is modelled in that the contribution of a component x_i to the value (fitness) of $\boldsymbol{x}$ depends not only on the configuration of x_i, but also of K other components of $\boldsymbol{x}$ (within-complementarity) and may also depend on the configuration of C components belonging to strings other than $\boldsymbol{x}$ (between-complementarity).

In this paper, the *N-K* model and its extensions provide a unifying formalism to discuss the structural regularities affecting the returns to R&D in the very-long run. The analysis exploits results on the deformations induced by different search procedures in the structure of a search- space representation[5] to

[4] Interestingly enough, these authors try to embed in their formalisation not only the economic, but also the social and institutional constraints on the pace of technological progress along a given trajectory.

[5] In economics and organization theory the *N-K* model has been mainly used to study the relative evolutionary effectiveness of different search procedures: local versus global (Frenken, Marengo and Valente, 1999; Marengo, Dosi, Legrenzi and Pasquali, 2000); centralized versus de-centralized (Kauffman and Macready, 1995; Frenken, 2001); cognitive versus experiential (Gavetti and Levinthal, 2000). The *N-K* model has been also used to discuss the relation between complexity, modularity and vertical (dis)integration (Marengo, 2000; Frenken, 2001). A common theme coming up in this literature is the interdependence between the design of a search landscape, in particular, the number and location of the points that will count as local optima, and the design of a search procedure on the landscape (Frenken 2001, McKelvey, 1999; Levinthal and Warglien, 1999).

address the evolutionary factors affecting the complexity of the search spaces, which is normally considered as a given of the analysis[6].

The notions of modularity and redundancy will come to play an important role in what follows. Modularity may be loosely defined as the extent to which the dimension of a problem-structure can be reduced through exact or approximate decomposition into *separate* problems of a lower dimension. Redundancy in a system may be defined as the probability that a change in the system configuration is unconsequential[7]. If modularity and redundancy are taken as given, the rising dimension of a search space, resulting from the discovery of novel components, brings with it a drive towards higher inter-relatedness and complexity.

The changing organisations' boundaries and the associated processes of vertical disintegration (integration) shift the R&D complementarities from complementarities within the search space of a given organization to complementarities between the search spaces of different organizations (or vice versa)[8]. As long as redundancy, modularity and search cost (the cost of a unit-step in search space) are given, the changing organizational boundaries will not be able to avoid that the rising overall dimension of search spaces, by creating greater scope for inter-dependency, brings R&D towards a complexity catastrophe (Kauffman, 1993) (part 2).

The tendency above is counteracted by other evolutionary factors. In section 3.1 we consider the reduction of interdependency resulting from the diffusion of modularity. The incremental propagation of modularity coherent with the ruling interface standards meets an obstacle when progress demands a discontinuous, global change of these standards reflecting a paradigmatic shift in knowledge[9]. Evolution on rugged knowledge landscapes can escape the attraction-basin of a paradigm, interpreted as a local peak in the landscape, if redundancy offers the possibility of drifting away from the local peak through 'neutral mutations' (section 3.2). Neutrality still holds its relevance for knowledge evolution after the possibility of forward-looking, goal-directed procedures is admitted. Long-jumps on knowledge landscapes may be induced by experience-based search exploiting long-term memory, or by goal-directed search (section 3.3). There is also a form of horizontal loose coupling between modules which is most relevant to the transition between paradigms. What is crucial here is the *variety* of functions and applications which is produced by technological innovation and scientific discovery. I suggest that there is a rigorous and relevant sense in which the availability of specific and relatively isolated application domains provides the matrix of knowledge interactions with loosely-coupled functional and structural horizontal modules. In this way,

[6] In the lines of enquiry considered in footnote 5, an *original problem* is posed, such as the exploration of a technological landscape of exogenously given complexity. The influence of bounded rationality and computing costs on the evolution of search heuristics is then investigated. It is finally shown how different organizational choices concerning the search procedure deform the search landscape corresponding to the original problem in different directions.

[7] More accurate definitions of redundancy and modularity are given in section 2.

[8] Ceteris paribus, organizational boundaries are in turn affected by the degree of modularity of problem spaces (Langlois 2001, 2002; Frenken 2002).

[9] In fact, the emergence of a set of standards, or dominant designs, is part and parcel with the emergence of a new paradigm (Caminati, 1999).

revolutionary ideas can be brought into use to serve specific purposes for which they are already fit, without interfering with the established ideas (section 3.4).

Part 4 draws together the ways and extent in which the arguments of part 3 qualify or substantially alter the main conclusion of part 2. Part 5 concludes that the returns to R&D defy the imposition of law-like regularities extending over long spans of scientific or technological history.

2. Search on technological landscapes

2.1. The family of N-K fitness landscapes

Let us think of $\boldsymbol{x}$ as a configuration, or 'design', of idea $\boldsymbol{T}$. Information on a given design is coded in a string of binary *components*, that we may interpret as characteristics which may be present or absent within the given design[10]. More precisely, a design of $\boldsymbol{T}$ is a string of N binary elements $(x_1, x_2, ..., x_N)$, where each x_j , $j = 1, ..., N$ can assume value 0 or 1. There are then 2^N possible designs of $\boldsymbol{T}$, corresponding to the number of different states in the space $\{0, 1\}^N$. In fact, $\boldsymbol{T}$ can be thought of as the set of its 2^N possible designs. Hence, we define $\boldsymbol{T} = \{0, 1\}^N$. Let $\boldsymbol{x}$ and $\boldsymbol{x}'$ be N-strings in $\boldsymbol{T}$. The *distance* between $\boldsymbol{x}$ and $\boldsymbol{x}'$ is defined by the number of components of the former having a different value with respect to the corresponding components of the latter. The neighbourhood of $\boldsymbol{x}$ is the set of strings in $\boldsymbol{T}$ with distance from $\boldsymbol{x}$ less than or equal to 1. It consists of $\boldsymbol{x}$ and its N neighbours.

The (relative) performance of a design vis a vis the other designs potentially available for the same application domain defines its competitive strength, or *fitness*. The fitness function of $\boldsymbol{T}$ is the map V: $\boldsymbol{T} \rightarrow \Re$ associating each design of $\boldsymbol{T}$ with its fitness value (a real number). In particular, the fitness value of a string is the sum of the fitness contributions of its N components. More formally, the map V is defined as:

$$V(x_1, x_2, ..., x_N) = 1/N \left[\sum_{j=1}^{N} V_j(x_1, x_2, ..., x_N) \right]$$

where $(1/N)V_j(x_1, x_2, ..., x_N)$ is the fitness contribution of the string component j. For the present purposes V_j is best treated as a random real in the unit interval [0, 1]. The above notation attempts at formalising the notion of input interdependence, in that, the fitness contribution of j depends not only on the configuration x_j, but also on the configuration of one or more string components x_h, $h \neq j$. If x_h is a non-redundant argument of the function $V_j(x_1, x_2, ..., x_N)$, then component h is *complementary* with respect to j[11]. $K_j \leq N - 1$ is the number of string components which are complementary with respect to j. For the sake of simplicity, we assume that K_j is constant across the components of $\boldsymbol{T}$: $K_j = K$, $j = 1, ..., N$, and that, for any given K, the specific pattern of complementarity relations (epistatic interactions) between string components is fixed randomly.

[10] For the time being we abstract from the otherwise important distinction between information and knowledge and the fact that techniques, like routines, may contain elements of tacit knowledge. See however below, part 3 and 4.

[11] It may be worth stressing that our definition of a complementary input does not correspond to the more restrictive definition used in Milgrom and Roberts (1990), or Topkis (1998), which is based on the mathematical notion of super-modularity.

The *fitness landscape* of ***T*** is the graph ($V(\boldsymbol{x}), \boldsymbol{x}$): $\boldsymbol{T} \rightarrow [0, 1] \times \boldsymbol{T}$.

In the absence of complementarity (K = 0), a change $0 \rightarrow 1$, or vice versa, in the configuration of a single component, does not affect the fitness contribution of any other. This implies that the fitness landscape of ***T*** has at most one isolated local peak and a local maximum of *V( )* on ***T*** is a global maximum. When $K > 0$, there may be situations where a fitness increment can be reached only through a changed configuration of two or more string components, *simultaneously*. This amounts to the possibility of multiple isolated local peaks on the fitness landscape. When complementarity is maximal ($K = N - 1$) the fitness landscape is random, in the sense that the fitness values of neighbouring states are totally uncorrelated. The cases K = 0 and K = N − 1 lend themselves to easy formal analysis (Kauffman, 1993). Here, we stress the sample properties of the large family of correlated N-K landscapes, which spans the parameter space between the single-peaked (K = 0) and the random (K = N − 1) landscape.

Remark 1. (Kauffman, 1993, pp. 55-57): Let $V^*(N, K)$ be the expected fitness of a local peak on a *N-K* landscape, the expectation being taken across local peaks and landscapes. *(i)* For a fixed *K*, the sample average computation of $V^*(N, K)$ remains approximately constant as *N* grows to $N = 96$. (ii) For a fixed and sufficiently large *N* (Kauffman used $N \geq 8$), the sample average computation of $V^*(N, K)$ is first increasing and then decreasing in *K* (starting from $K = 0$), reaches a maximum at K^*, $2 \leq K^* \leq 4$, and dwindles towards average fitness 0.5 for *K* sufficiently large (*complexity catastrophe*).

Remark 2. (Kauffman, 1993): Let *N* be given and constant. On average, the higher *K*,
(i) the higher is the number of local optima;
(ii) the lower the correlation between the fitness values *V*( ) of neighbouring strings $\boldsymbol{x}, \boldsymbol{y}$[12].

2.2 Local versus global search

So far we have provided a formal description of (codifiable) knowledge landscapes, but nothing has been said on how exploration proceeds on such landscapes. This issue can be addressed building upon Simon's notions of decomposability and near decomposability (Simon, 1962). The problem of exploration on the *N-K* fitness landscape defined by *V()* is *decomposable* if the problem can be divided into separate sub-problems that may be solved independently.

A *decomposition* of the search space $\{0,1\}^N$ is a set of *Z* subspaces $\{0,1\}^{N_z}$, $z = 1, \ldots, Z$, such that their union recovers $\{0,1\}^N$. A decomposition is *complete* for the fitness function *V()* if x_i complementary with respect to x_j implies that x_i and x_j belong to the same sub-space in the decomposition. Complete decomposability for *V()* is a sufficient but not a necessary condition in order that a problem be decomposable. A necessary and sufficient condition for problem decomposability is stated in Frenken, Marengo and Valente (1999) and Marengo (2000).

Following Holland (1975), and Page (1966) a *schema* is a ordered sequence of 0, 1 and #. The defining bits of the schema are those that differ from #. A schema has dimension, or size, *n* if it has *n* defining bits. The projection of a string $\boldsymbol{x} \in \{0,1\}^N$ on a schema $\boldsymbol{h} \in \{0,1,\#\}^N$ is a string $\boldsymbol{y} \in \{0,1\}^N$ such that

[12] It may be worth noting that, for a fixed *K*, the correlation between the fitness values of neighbouring strings increases with *N*.

$y_i = h_i$ if i is the index of a defining bit of $\boldsymbol{h}$ and $y_i = x_i$ otherwise. A schema is *dominant* if 'by projecting on it all other strings we obtain a new string with higher (or equal) fitness. Of course, only schemata which are part of strings with maximum fitness can be … dominant' (Frenken, et al., 1999, pp. 151-2). A *cover* of a *N-K* fitness landscape is a set of dominant schemata such that their union yields a maximum-fitness string of the landscape. The *cover size* of the fitness function *V()* on $\{0, 1\}^N$ is the size of the largest schema contained in the minimum-size cover for that function, that is the cover with minimum-size schemata (sub-problems). The cover size of *V()* is therefore a measure of the complexity of the fitness landscape defined by *V()*. Cover size lower than *N* is necessary and sufficient for problem decomposability into independent sub-problems which can be solved in parallel. Cover size gives the dimension of the largest sub-problem, thus fixing an upper bound[13] on the time required to find the optimal configuration in $\{0, 1\}^N$.

When the pattern of epistatic interactions between string-components is arbitrary, the cover-size of the fitness function *V()* is likely to be *N* or close to *N*, as soon as the interaction parameter *K* rises above 2 (Frenken et al., 1999). Correspondingly, computing the fitness maximizing string $\boldsymbol{x^*}$ on the landscape is a complex task: solving the problem within a reasonable computing time is most likely to be impractical, if problem-size *N* grows sufficiently large[14]. In a selection environment where computation time matters and cover size is close to *N* it may be profitable to attack the original problem with an incorrect decomposition such that the dimension of the sub-problems, or schemata, is lower than cover size. The price to pay for faster computing time is that the algorithm may not lead to the optimal solution. Search strategies induced by decompositions with schemata smaller than cover-size qualify to be called *local*, in that only 'local' interactions are considered.

Remark 3: The ground for defining a search procedure *local* has to do with the way in which mutations are generated. A decomposition defines a search euristics in the sense that the original problem space is replaced by the union of different subspaces. On each subspace search proceeds *as if* the decomposition is correct, that is, without worrying about the fitness effects of a move in the subspace on the other subspaces. The string-components belonging to the same subspace can be mutated simultaneously. Intuitively, search is local if it can not reach a global optimum in a single search round. Conditions sufficient to qualify a search procedure 'local' are that the dimension of every subspace is in the decomposition is lower than cover size, or that it is based on an incorrect decomposition, or that there is a multiplicity of subspaces that are searched sequentially.

Remark 4: The ground for accepting or rejecting a local mutation depends on the *centralized* or *decentralized* nature of the decision. If the decision is *centralized*, a generated local mutation is accepted if and only if it is fitness-improving *globally*.

Frenken et al. (1999) and Marengo (2000) build upon Simon (1962) to suggest that an ε-satisficing solution (a solution not-more-than-ε-far from optimal fitness) to the search problem will be available if the problem structure is *nearly-decomposable*. This is the case if there exists a decomposition of the search

[13] "Cover size might over-estimate the complexity of the solution for a search algorithm because, if some schemata which form the minimum cover have bits in common, solving one of them reduces the size of those with overlapping elements. Of course, the bits which are common to two or more schemata in the cover must take the same value at the optimum, thus we can search for such values in the smallest of the overlapping schemata and reduce the dimension of the larger ones…" (Frenken et al., 1999, p. 152).

[14] The problem is *NP-complete* (Rivkin, 2000, p. 832-33).

space into sub-spaces with an hierarchic structure such that components belonging to the same sub-space receive *sufficiently weak* (in their fitness effects) epistatic links from components belonging to different sub-spaces at the same or at higher levels in the hierarchy[15].

For ε negligibly small, the formal conditions for near-decomposability are strong. In particular, if the dimension of the nearly-decomposed sub-spaces grows in proportion to the dimension *N* of the original problem, finding the optimum on the sub-spaces will be intractable (*NP-complete*) much like finding the optimum on the original problem space. Frenken et al. (1999) simulate a selection environment where sub-populations of agents are characterized by different decompositions of a given problem and are faced by constraints on their computing resources. They show that sub-populations with sub-optimal decompositions, tend to invade the population. This motivates the following conjecture:

Conjecture 1: Consider Kauffman's *N-K* problem and the set of possible decompositions of $\{0,1\}^N$. At $K > 2$ search time fixes an upper bound $M <$ cover size on the subspace-dimension of the *stable decompositions*. These are the decompositions that –in a selection environment where sub-populations of agents choose between different decompositions on the base of their relative success within a fixed and finite time interval- would not be invaded by any other decomposition. If we expand the scale *N* of the problem while holding *K* > 2 constant, the ratio *M/N* weakly declines.

An extreme form of local search is Kauffman's *fitter dynamics* which follows from the decomposition of the *N*-bit search space into *N* subspaces, each of size 1 bit: to generate one new configuration from a current configuration, one of the *N* subspaces is randomly selected and its configuration mutated; if the fitness value of the new design thus obtained is higher than the fitness value of the current design, the idea moves to the new configuration. The procedure is then iterated until a local maximum of *V( )* is reached. The fitter dynamics is non deterministic, but a local peak of the landscape is a stationary state of this dynamics. One may also want to consider situations in which agents choose to experiment more designs before making a move with possibly long-lasting implications. Here the polar case is that in which all the designs within the neighbourhood of the current state are tried out before making a move on the landscape[16]. Also this polar case is considered, if more briefly, in Kauffman [1993] with the label '*greedy dynamics*'. To fix our ideas, we report on simulation results based upon these particular forms of local search; the main qualitative features of the results extend to local search in general.

Remark 5. (Kauffman, 1993): Assume local search through *fitter* or *greedy* dynamics on a family of *N-K* landscapes. Let *N* be given. On average, the higher is *K*, the smaller is the basin of attraction of a local peak and the steeper the adaptive path to it. The ratio between the average number of states in the basin of a 'high' and of a 'low' local peak[17] of the landscape, increases with the correlation of the landscape, hence increases

[15] To identify the set Σ of ε-satisfacing solutions Frenken et al. (1999) introduce the notion of a ε-cover of the fitness function *V()*. This is a set of schemata which are dominant on the set of strings which do not belong to Σ and such that the union of the indexes of their defining bits recovers {1, 2, …, *N*}.

[16] Through an appropriate re-scaling of time, this generates the assumption that at any given 'date' a unit takes a 'one-step' move from the pre-determined state to the state identified by the fittest string in the given neighbourhood, thus yielding a fully deterministic dynamics on ***T***.

[17] 'High' and 'low' local peaks identify here local peaks with fitness above and below average, respectively.

with $1/K$, as K declines from $N - 1$ to zero at given N. Thus, local search converges to a 'high' local peak of the landscape with a probability that is decreasing with K, at given N.

2.3. Centralized versus decentralized search

A message of section 2.2 is that on rugged landscapes, the gap between maximum sub-space dimension in the stable decompositions (in the sense of conjecture 1) of ***T*** and cover-size is bound to increase with N. If local search is the best one can do, it is worrying that local search on rugged landscapes is likely to be trapped on poor local optima (remark 5). A way of escaping very poor local optima is offered by admitting *local* search procedures which may temporarily go in the 'wrong' direction, that is, which are allowed to move some steps *down-hill* on fitness landscapes (Kauffman and Macready, 1995, pp 35-40). A way of achieving this is to decentralize the search procedure across a multiplicity of 'agents'. Consider N as given exogenously. A *decentralization* is a disjoint decomposition of ***T*** into S sub-systems such that subsystem $\boldsymbol{T}_h$, $h = 1, \ldots, S$, consists of a sub-string of N_h binary components and $N = \sum_{h=1}^{S} N_h$. Decentralized search means that each $\boldsymbol{T}_h$ evolves on its landscape $\{0,1\}^{N_h}$ *autonomously*, that is, according to decisions that have in their view the performance of $\boldsymbol{T}_h$ and neglect the effects on the performance of other sub-strings. The decentralization $N = \sum_{h=1}^{S} N_h$ induces the corresponding *separation* $G_h = C_h + K_h$, where K_h and C_h are the average number of links that a component of sub-string h receives from components belonging to the same or to other sub-strings, respectively. G, K and C are the cross-h averages of G_h K_h and C_h. Notice that G is independent of the decentralization in use. The rest point of a decentralized search process is such that all the sub-strings are simultaneously at a local optimum on the subspaces. Hence it is unlikely that a very poor local optimum in the original (non-decentralized) problem space is a rest point of a sufficiently decentralized procedure. The other side of the coin is that co-evolutionary relations triggered by decentralized search may displace sub-strings from relatively adapted positions on their landscapes[18].

Co-evolutionary convergence: Kauffman and Macready (1995) suggests that with extreme local-search procedures (1-bit mutations) the optimal number S of subspaces, each of dimension N/S, hence the optimal determination of C and K, given N and G, is such that C is poised at a critical intermediate level; above this level the co-evolution dynamics does not find a rest point. This defines the 'best' position on the trade-off between co-evolutionary turbulence and the probability of escaping poor local optima.

Obviously enough, the same logic[19] behind Kauffman and Macready's decentralized search can be applied to local-search procedures less extreme than the 1-bit mutation procedures considered in their paper. In this respect, it can be confidently conjectured that local search defined by decompositions admitting schemata of varying but on-average-larger dimension would be 'optimally' decentralized through a lower number of

[18] It is worth stressing that the accepted mutations and the rest points of a decentralized search procedure would not generally correspond to those of a centralized procedure.

[19] It may be worth observing that Kauffman and Macready (1995) decentralization is not aimed at speeding-up search through parallel, independent search activities. However important, the speed-up side of decentralization is not considered here.

subspaces. The intuition is that a farther-reaching local search would face lower risk of being trapped on poor local optima and on this ground it would require a lower injection of co-evolutionary disorder.

2.4. Increasing the scale of problem spaces without selection on structural parameters induces a complexity catastrophe

Definition1: A radical innovation is defined by an expansion of the set of knowledge components; incremental innovations discover new, fitness-increasing configurations in the space defined by the existing set of knowledge components. On this ground, a sequence of radical innovations increases the scale N of $\boldsymbol{T}$.

At this stage, I do not yet introduce selection on the structure of epistatic links or on the ruling decentralization of search fixing the separation $G = K + C$. These will be considered at length in later sections. For the sake of the argument, it is temporarily assumed:

Assumption 1: For given N, the existence of an epistatic link from component $j \neq i$ to component i of $\boldsymbol{T}$ is a random event occurring with probability $\pi_{ij} = \pi(N) > 0$, $i, j = 1, \ldots, N$. The average number G of epistatic links received by a component of $\boldsymbol{T}$ increases with N. Recall that G has upper bound $N - 1$. The motivation behind the assumption is that we are ruling out factors, other than the scale of $\boldsymbol{T}$, which may affect the probability π. On this ground, it is assumed that $\pi(N)$ does not decline 'too' fast when N increases.

Remark 6: For given N and G, assumption 1 implies that the 'wiring' of epistatic links is random in the sense that if the component x_i of $\boldsymbol{T}$ receives epistatic links from G other components of $\boldsymbol{T}$, then each of the $N - 1$ components of $\boldsymbol{T}$ other than x_i has the same probability $G/(N - 1)$ of sending a link to x_i.

The main emphasis of section 2.4 is on the following remark:

Remark 7: *(i)* Assume that the structural relation between knowledge and performance of a system $\boldsymbol{T}$ is described by a Kauffman N-G fitness landscape, where $\boldsymbol{T}$, N and G are defined as above. *(ii)* Assume that there is a finite bound on search time and R&D is guided by local search procedures. *(iii)* Assume that the scale N of $\boldsymbol{T}$ increases indefinitely and assumption 1 holds. It is an implication of remark 5 and of the notion of optimal co-evolutionary search that, as G keeps increasing, then *no matter how organizations' boundaries move the partition* $G = C + K$, from within-organization interactions to between-organization interactions, or vice-versa, R&D will not be able to avoid the *eventual* onset of a *complexity catastrophe*.

Assumption 1 as well as assumptions *(i)* and *(ii)* in remark 7 are all, to some extent, questionable. Hypotheses aiming at a strong qualification and revision of these assumptions are presented and discussed in part 3. To this end, and by way of introduction, the following section brings in selection forces.

2.5 Introducing selection on structural parameters

Selection on the structural parameters G, K, C and S is introduced as follows.

At any given date there exist alternative knowledge systems available and a population of units explores a population of systems. While a population of units explores the landscape $\{0, 1\}^{N(T)}$ corresponding to system $\boldsymbol{T}$, another population explores the landscape $\{0, 1\}^{N(T')}$ or $\{0, 1\}^{N(T'')}$ … corresponding to system $\boldsymbol{T'}$, $\boldsymbol{T''}$ … respectively. Selection on the populations of units induces a selection on the population of systems[20] and a corresponding selection on the structural parameters *N(**T**)*, *N(**T'**)* …, *G(**T**)*, *G(**T'**)* …. Moreover, the population searching on a given landscape may be divided into sub-populations implementing alternative decentralizations of the landscape. In this way a selection is induced on decentralization modes.

With this framework in mind, we shall discuss at length in part 3 how modular structures evolve through the selection for *lower* values of *G* aimed at the sustained evolvability of the knowledge system. Before that, we stress in the following remark the existence of selection forces participating into the joint determination of *G* and of organization boundaries, but acting in a different direction. The value of *G* is increased rather than decreased. The argument introduces selection on the number of epistatic links, but not on the wiring of such links which remains random.

Remark 8: Assume that, for given *N* and *G*, the wiring of epistatic links is random in the sense of remark 6. Assume that *N* keeps increasing through time at a possibly slow pace. The conjecture stated in appendix A.0 implies a drive towards higher *C* at given *G*. Kauffman (1993, pp. 248-250) shows that *in co-evolution episodes there are forces selecting for higher K values when C increases*. Thus, decentralizations with rising number of sub-spaces induced by rising *N* (see conjecture 1) and lack of selection on the 'wiring' of epistatic links, trigger forces selecting for higher *C* and *K*, hence for higher *G*. This brings with it a positive relation between *G* and *N* which adds to the positive relation contemplated in assumption 1.

3. Evolution cum evolvability

The scenario of a looming complexity catastrophe spelled out in part 2 has been challenged by scholars of evolution of complex systems (be they natural, artificial or social). One strand of research has brought to the fore the factors which are responsible for a long-term selective suppression of epistatic links between string-components and the ensuing *increasing modularisation* of complex systems. A second strand of research has challenged the standard architecture of *N-K* fitness landscape and the ensuing evolutionary dynamics on such landscapes, in that they downplay the possibility that local optima are connected by a network of 'flat ridges' enabling a population to drift far away on the landscape.

3.1. Evolvability and the evolution of modularity

Our presentation of *N-K* fitness landscapes in part 2 exploited a number of simplifying assumptions which it is now best to remove.

The first and most relevant simplification was that every component x_j of the string $\boldsymbol{x} \in \{0,1\}^N$ exerted a distinct fitness contribution *as if* the component identifies a specific function *f* and the number of functions $F = N$. In a more general setting (Altenberg, 1994), it may be assumed that a component may affect a multiplicity of functions and $F \neq N$. Overall fitness is the product of the additively separable contributions of the *F* functions:

[20] To give a concrete example, the macro-invention of electricity took place with the introduction of two subsets of knowledge components corresponding to the direct-current and alternating-current systems, respectively.

$$V(x_1,...,x_N)=\frac{1}{F}\sum_{f=1}^{F}V_f(x_1,...,x_N)$$

The *pleiotropy* of the *j*th string component is the number of functions affected by it, that is, the number of fitness contributions $V_f(x_1,...,x_N)$ of which x_j is a non-redundant argument. Conversely, the number of string components that are non-redundant arguments of $V_f(x_1,...,x_N)$ is the *polygeny* of *f*. The structural relation between string components and functions is then described by the $N \times F$ matrix $\boldsymbol{M} = [m_{jf}]$, where $m_{jf} = 0$ or $m_{jf} = 1$ depending on whether x_j is or is not a redundant argument of $V_f(x_1,...,x_N)$. To relate the present discussion to the notion of near-decomposability, we define x_j a redundant argument of $V_f(x_1,...,x_N)$ if the fitness effect of the former on the latter is 'sufficiently weak'. How weak is sufficiently weak depends on the value of the satisficing parameter ε. Biologically inspired problem representations identify the *F* functions with *F* phenotype characters and label ***M*** *genotype-phenotype map*[21].

The main reason behind the above generalization is that the ability to achieve *evolvability*, that is, sustained variation and 'progress' over long spans of evolution time, can be referred in a meaningful way to the structure of the matrix ***M***. The modularity of the genotype-phenotype matrix corresponds to a block decomposition of ***M***, such that between-blocks pleiotropy effects on fitness are relatively weak with respect to within-blocks pleiotropy effects.

If the idea of a block decomposition of an interaction matrix has a long standing in social science (Simon and Ando, 1961), the interesting point raised by scholars of natural and artificial evolution is the *detailed specification of selection forces changing the block decomposition over time*, with the aim of preserving and increasing *evolvability*[22]. The general thrust behind such selection forces had been anticipated by Simon since his formulation of the near-decomposability concept. The implications for technology evolution and economic organization are now widely appreciated, if not deployed in all their manifold relevant dimensions[23].

---

[21] Cf. Altenberg (1995), p. 231. ***M*** has *F* polygeny column vectors and *N* pleiotropy row vectors.

[22] The grouping of the subset of string components serving the same subset of functions within a relatively isolated 'gene net' "… means that genetic change can occur in one of these gene nets without influencing the others, *thereby much increasing its chance of being viable*. The grouping leads to a limiting of pleiotropy and provides a way in which complex developing organisms can change in evolution." (Bonner, 1988, p. 175, emphasis added.) This implies that genetic change can proceed rather smoothly along the subset of dimensions participating in the regulation of a functional complex occasionally conveying adaptive advantages, while preserving the functionality of the whole. Various selection mechanisms for the evolution of modularity have been suggested (Wagner and Altenberg, 1996; Wagner, Mezey and Calabretta, in press). These range from the selective suppression of pleiotropic effects (Wagner, 1995) to the growth in the number *N* of string-components, through sub-string duplication, with subsequent selection of those strings such that the duplicated segments have relatively low and localized pleiotropic effects (Altenberg, 1994, 1995).

[23] Cf. Baldwin and Clark (2000), Langlois and Robertson (1992), Buenstorf (2002), Langlois (2002), Axelrod and Cohen (2000), Calcagno (2002), Brusoni and Prencipe (2001), Devetag and Zaninotto (2002). As suggested below in the text, a discussion of the relation between selection for modularity and paradigmatic changes in knowledge is still lacking.

Two general conclusion intersect the application domains of the modularity literature. The evolution of modularity in complex systems: *(i)* takes place through a selective pressure against the growth of diffused pleiotrpopic effects; *(ii)* enables high innovation rates along the dimensions in which evolution opportunities are favourable, while preserving stability and continuity in the dimensions which guarantee the viability of the system as a whole.

From the view point of the present discussion conclusions *(i)* and *(ii)* are most relevant in that they single out a process acting counter to the growing interconnectedness of knowledge systems which was hypothesized in part 2. Still, a careful reading of *(ii)* reveals that there are definite bounds to the scope of the counter-acting force thus identified. To see this, it is worth recalling that, according to definition 1, a radical innovation expands the dimension of the search space and generates a new fitness function *V()* on the expanded domain. A sequence of radical innovations will then preserve the pre-existing modular structure[24] of the genotype-phenotype map only if knowledge expansion occurs in restricted directions, such that the new string components, or their configuration, do not establish non-negligible interdependences between the pre-existing blocks. The restricted directions referred to above are set by the design rules (Baldwin and Clark, 2000) which confer continuity to a given lineage, for instance, a given theory or technological paradigm, by fixing a stable set of interface standards which are *peculiar to it*. There is in this respect a crucial difference between the scopes of the forces and counter-forces under discussion. The drive towards rising interconnectedness hypothesized in part 2 is specific to the knowledge domain and acts *across different lineages* of this domain, in that it is at work in those saltation events giving rise to new paradigms. The counter-acting force related to the evolution of modular architectures, as described in this section, has to do with the evolvability of a given lineage; thus, it is at work *within* the evolution episodes which mark the birth, development and maturity of a given paradigm.

The history of ideas shows that radical innovations do not always conform to the pre-existing design rules. Addition of new components to N, F or both, or selection of new string-configurations may destroy the near-independence between the *pre-existing* blocks; in this case, near-decomposability can be re-established only through a global re-design of the block structure of $\boldsymbol{M}$. On search spaces of rising dimension, global re-design is an increasingly demanding task.

We may provisionally conclude that selection for modularity favours evolution along special dimensions enabling the acquisition of favourable characteristics. The benefit is circumscribed to the evolution of individual lineages and may turn into an obstacle to the advent of macro-mutations involving global changes in the product technology and a re-definition of interface standards (for a recent re-statement, see Brusoni and Prencipe, 2000). In this sense, the modularity argument, as developed before, does not fully escape the logic inherent to the rules of local search on rugged fitness landscapes, which envisage the

[24] A block decomposition of the matrix $\boldsymbol{M}$ corresponding to the fitness function *V()* is induced by an ε-cover of *V()*. Consider the phenotype-genotype maps $\boldsymbol{M}$ and $\boldsymbol{M'}$ corresponding to the functions $V(\boldsymbol{x})$, $V'(\boldsymbol{x'})$ on $\{0, 1\}^N$ and $\{0, 1\}^{N'}$, respectively and such that $N' \geq N$. For fixed ε, let ε-$\boldsymbol{C}$, ε-$\boldsymbol{C'}$ be ε covers of $V(\boldsymbol{x})$, $V'(\boldsymbol{x'})$, respectively. ε-$\boldsymbol{C}$ and ε-$\boldsymbol{C'}$ are consistent *iff* the following conditions hold for $i, j = 1, \ldots, N$: *(a)* x_i, x_j defining elements of the same schema in ε-$\boldsymbol{C}$ implies that x_i, x_j are defining elements of the same schema in ε-$\boldsymbol{C'}$; *(b)* x_i, x_j defining elements of different schemata in ε-$\boldsymbol{C}$ implies that x_i, x_j are defining elements of different schemata in ε-$\boldsymbol{C'}$.

characteristic situation of a lock-in on a possibly poor local peak. In sections 3.2 and 3.3 we consider mechanisms conferring wider scope to evolutionary change.

3.2. Escape from local peaks: neutrality

A characteristics of Kauffman's *N-K* fitness landscapes described in part 2, including their generalization offered in section 3.1 is that every change in a string configuration has relevant fitness effects. The alternative idea behind the notion of neutrality[25] is that there is a large redundancy in the way relatively unchanging functional characteristics are coded in the genotype.

For the sake of simplicity, we assume that $\boldsymbol{M} = [N \times N]$ Identity matrix. Following Barnett (1998), a one-bit mutation $\boldsymbol{x'}$ of $\boldsymbol{x} \in \{0,1\}^N$ is *neutral* if and only if $V(\boldsymbol{x'}) = V(\boldsymbol{x})$; the *neutral degree* of $\boldsymbol{x}$ is the number of neutral mutations of $\boldsymbol{x}$. The possibility of neutral mutations arises on an *N-K* fitness landscape if for any $f \in [1, \ldots, N]$ there is a probability $p > 0$ that the fitness contribution $V_f(\boldsymbol{x}) = 0$. The probability p defined above introduces the class of *N-K-p* landscapes. Neutral mutations induce a partitioning of $\{0,1\}^N$ such that $\boldsymbol{x}$ and $\boldsymbol{x'}$ belong to the same equivalence class if and only if there is a sequence of neutral one-bit mutations connecting $\boldsymbol{x}$ and $\boldsymbol{x'}$. "The *neutral networks* of the fitness landscape are defined to be the equivalence classes of this partitioning" (Barnett, 1998, p. 19). A coarser partitioning is induced on $\{0,1\}^N$ by imposing that $\boldsymbol{x}$ and $\boldsymbol{x'}$ are in the same equivalence class if and only if $V(\boldsymbol{x}) = V(\boldsymbol{x'})$. An equivalence class of this coarser partitioning is a *neutral set*. The neutral networks are the connected components of the neutral sets. The reason for introducing the latter is that their properties are easier to study analytically (see appendix A.1). It is worth insisting that neutral networks, neutral sets and their properties are observed in spite of the 'ruggedness' of the landscapes, which is scarcely, if at all affected by neutrality (see appendix A.1).

A discrete evolutionary dynamics of *S* strings on an *N-K-p* landscape is induced through selection and mutation operators[26]. Barnett (1997, 1998, 2000) reports on the typical findings from the simulation of evolutionary dynamics on *N-K-p* landscapes (Appendix A.1, Remark 13).

On the premise that evolution on *N-K-p* landscapes is relevant to the knowledge domain, the main implications that are of potential interest for the fate of R&D may be summarised as follows. As *N*, and (possibly) *K* slowly increase over time, it may well take longer, on average, to find the way out a neutral network of given fitness; but the point is that *a way out exists* and will be sooner or later be found. This hints at factors which make the long term-behaviour of R&D success potentially irregular, with long intervals of gradual, relatively-slow progress or even stasis interrupted by rapid outbursts of innovations corresponding to the population transition to a new network of higher fitness. We are left with the task of showing how the above premise may hold true.

---

[25] The idea can be traced to the contribution of Kimura (1983); recent reformulations in the biological domain are corroborated by experimental studies of protein evolution (Huynen, Stadler and Fontana, 1996; Huynen, 1996).

[26] Population at the next 'generation' results from *S* selections, with replacement, from the current population. The probability that a string gets selected is proportional to its relative fitness. Every string configuration in the new population is then mutated with a fixed, low probability *m*. It may be worth noting the biological inspiration of this dynamics where, unlike the 'fitter' dynamics considered in section 2, selection operates only at the population level.

As a preliminary to this enquiry, it is worth dispensing with Barnett's restriction of neutrality to one-bit *mutations* within strings with a constant number *N* of information components. We can extend the theme of neutrality to *local changes* in information that do not exceed a given tolerance range in their fitness effects. The extended definition admits deliberate downward steps on the landscape designed to increase the rate and scope of mutations[27], multi-bit mutations and even changes in *representation* expanding the dimension of the search space (see appendix A.2). With this extension in mind, we can state that the *possibility* of neutral changes in information results from a structural rule which is quite generic in its domain of application. The rule states the *multiplicity* of the ways in which the activation of 'functional characteristics'[28] achieving a given performance value (fitness), up to the tolerance range, can be coded in a replicable string of 'information'[29]. In the technological domain the performance value is defined in terms of engineering and economic criteria; in the scientific domain it can be defined in terms of 'explanatory coherence' (Thagard, 1992): the coherence between the items of a conceptual structure and the coherence between such items and those contained in the set of the empirical evidence under consideration.

A multiplicity of coding and an associated instance of neutral local mutations is revealed by the tolerance ranges in the parameter settings of a technological design. It is also inherent to the fact that a number of simultaneous parameter changes that still qualify the implied mutation as 'local' may mutually compensate in their effects. Corresponding phenomena arise in the scientific domain when we consider alternative theories which locally agree in their explanations and predictions that apply to a restricted subset of the experimental domain (think of Newtonian mechanics and Einstein's relativity dealing with experimental speeds far from light speed).

Neutral networks may therefore exist on knowledge landscapes and populations of information strings may correspondingly drift on such networks even if the allowed evolutionary dynamics are restricted by strictly-local search heuristics.

3.3. Neutral mutations with non-local search

Drift on neutral networks being a possibility, it can be argued that the idea of neutrality is relevant to knowledge revolutions in a more general sense, which takes due account of the variety of search heuristics (local versus non-local, experienced-based versus goal-directed)[30] available on knowledge landscapes. In particular, unfrequent *long jumps* induced via goal-directed or experience-based procedures, though reaching configurations of similar fitness value (neutrality) may deeply mutate the local-search knowledge base. In what follows we focus our attention on scientific revolutions.

Connecting concepts

---

[27] On search rules designed to speed-up experimentation, see Hovhannisian (2003).
[28] The term 'functional characteristics' identifies a particular implementation of a set of functions, as determined by the 'genotype'.
[29] In terms of the genotype-phenotype matrix ***M***, two string populations with different genotype-phenotype matrixes and competing in the same application domain, can attain the same fitness.
[30] Gavetti and Levinthal (2000).

According to Thagard (1992), fundamental scientific discoveries apparently conforming to the discontinuous 'Gestalt-change' representation originating with Kuhn (1962), reveal in fact a characteristic pattern ingrained in a complex network of evolving conceptual innovations building a link between the old theory and the new. Typically, the intermediate steps of the evolving conceptual structure do not yet possess the higher 'explanatory coherence' which is a property of the new paradigm only in its developed and consolidated versions leading to its final adoption by the scientific community. In fact, during the early stages of theory development, the attention of the creative mind(s) is focused on a restricted set of new concepts and theoretical relations that are credited of being highly coherent with a specific subset of the evidence. To the extent that the links of the new concepts and relations with the overall conceptual structure are still ill defined, the explanatory coherence of the whole may be in doubt, both for the scientific community and, most importantly, for the creative mind(s). In these early stages, it is often the case that connecting concepts and relations are introduced, which deform the subjectively perceived fitness landscape by adding or deforming information schemata in a way aimed at building a network of coherence links connecting the pre-existing theory with the new concepts and relations in the focus of the creative mind(s). Such links serve the purpose of filling the gap in explanatory coherence suffered by the emerging, but still fuzzy, theory[31]. We claim that the introduction of the schemata building such links are analogous to neutral changes in information; they enable units in the population of strings to search away from the population-average location, still centred on the received theory, while reducing the risk of a premature suppression by selection. In some cases the bridging links are introduced through recombination between the items of the pre-existing conceptual structure and new concepts and relations independently formulated through a goal-oriented activity, for instance, the explanation of a novel experimental evidence. Examples of this type are some connecting links in Lavoisier conceptual structure prior to the formulation of his fully developed theory of combustion. In fact, for some time after 1772 Lavoisier still regarded his conceptual structure not necessarily alternative to the theory of phlogiston. In other cases the connecting concepts pre-exist to the formulation of a discovery and indeed serve the purpose of open gates making the discovery itself accessible. For instance, Clerk Maxwell through his notion of a 'mechanical ether' could formulate his electromagnetic equations without deep questioning his un-shaking faith in Newtonian mechanics. Such deep questioning only became ripe in the subsequent decades, partly as a result of those equations.

It may be worth adding that the development of a new theory to the stage in which it represents a fully alternative conceptual structure will most often entail the removal of the conceptual links which provided a coherent connection with the pre-existing rival.

Selective forgetting, or 'returning on one's steps'

[31] This may occur through the introduction of a concept without detectable empirical counterpart. The notion of a 'mechanical ether' enabled Maxwell to interpret his newly formulated electromagnetic wave equations in the light of his unscathed faith in classical mechanics. In a short time these equations would pave the way to a different non-Newtonian explanation of wave propagation through space, preparing the ground for special relativity and the notion of a mechanical ether could be dispensed with altogether.

Other search heuristics on knowledge landscapes exploit the possibility of 'neutral mutations' (according to our extended definition) through the combined use of short-term and long-term memory. Referring to the traditional view of great scientific discoveries originating from sadden 'illuminations' of apparently unconscious origin, Simon (1966b) argues that both the discontinuity implied by such illuminations and the revolutionary aspect of the discovery can be understood as resulting from normal problem solving activities. In fact, these activities include experienced-based search strategies that are effectively designed to avoid trapping on poor local optima. Experience-based search is guided by focalisation mechanisms (inducing local search) which give rise to the stepwise formation of an information hierarchy temporarily stored in the short term memory and only selectively transferred in the long term memory. To the extent that the frustration resulting from local-search failures makes the focalisation mechanisms temporarily dormient, the information stored in the short term memory is lost, but this is precisely what activates long-term memory and enables local search to start afresh from the exploration of a previously discarded alternative path (Simon, 1966b and 1977, pp.294-99).

3.4. Escape from local peaks: design revolution through segregated modules and knowledge spillovers

In this section we elaborate on a road of escape from local optima on technological landscapes which helps explaining how radical innovations, after seeing the light in restricted application domains, can develop into macro-innovations. The escape gate is offered by the protected niche of application of a new or still unexploited idea which so avoids being prematurely suppressed by the strong competition from fitter ideas dominating other application domains. The absence of fitter rivals within the niche enables survival and recruitment of the resources necessary for development. In the language of fitness landscapes, a population of strings identifies a set of units concerned with a specific application domain, characterized by a fitness function $V(x)$ on the search space $\{0,1\}^N$. The identification of the new fitness function associated with the protected niche enables, through birth and migration processes, the agglomeration of a corresponding population in a still unexplored region of the search space where the population can experiment with alternative configurations until the new idea is developed into a dominant design.

The crucial point is that the improved knowledge of the search space which is so generated may be relevant to other application domains. Some of the information and understanding developed in a specific domain may turn out to be general-purpose knowledge. This leads to identify schemata yielding high fitness contributions for a large number of fitness functions and a corresponding set of application domains. In other words, well targeted long jumps on knowledge landscapes are brought in the reach of other populations. Such spillovers are characteristic of knowledge evolution and are unavailable in the biological domain.

Telling case studies which conform to this basic pattern abound in the history of technological ideas. They are offered by the history of wireless communication as synthetically reconstructed in Levinthal (1998), by the early development of the steam engine (initially conceived as a device for pumping water out of mines), of the machine-tool industry (Rosenberg, 1976) and of electricity (Hughes, 1983).

It is worth emphasising the premier position occupied by the *variety* of application domains in the mechanism described above. Variety in applications, enables the initial *specialization in use* of a novel idea which is instrumental to its introduction, survival and development up to the stage of final spreading across different uses. Most notably, if we focus our attention on the technological domain, this points to a relation between specialization and technological change which is altogether different from the relation implicated in the illustrious line of reasoning associated with Adam Smith (1776) and Allyn Young (1928) and still traceable in the more recent contributions of Romer (1990) and Arrow et al. (1998). The main point of departure is that indivisibilities play a crucial role in the latter, but not in the former. As a result of indivisibilities, Adam Smith and Allyn Young consider specialization-per-se as a *source* of increasing returns. The benefits that are so achieved are persistent only to the extent that specialization is persistent. This is not so in the mechanism described before, where specialization in use is a *temporary vehicle* for the exploration of knowledge search spaces[32].

4. A final re-assessment

The discussion in this part of the paper attempts at a final assessment of the ways (if any) in which the arguments of part 3 question the fundamental assumptions *(i)*, *(ii)* and *(iii)* behind Remark 7 of part 2, thus preparing the ground for the concluding remark on the existence of 'laws of returns' to R&D.

4.1. The structure of knowledge landscapes, variety and local search

If complementarity implies the ruggedness of fitness landscapes, the latter does not rule out the existence of narrow, but relatively smooth paths for escaping the basins of poor local optima. As long as the local-search heuristics are a fundamental, but not necessarily the unique, mode of exploration on knowledge landscapes (on this see below), the existence of such escape paths is a powerful argument for showing the long-term influence of variety on the ability to evolve. If a population of almost identical designs has access to a narrow escape gate with vanishing-small probability, the chances would be much higher for *some* member of a population of vastly differentiated designs. Under suitable conditions of knowledge transfer, the entire population may follow en masse the lucky or clever innovator.

Variety may be preserved and increased: *(i)* by making 'far' combinations reachable through exploration, or by eliciting controlled downward movements on the landscape; *(ii)* by placing constraints on selection. Redundancy and neutral mutations on a specific application domain identify a variety producing factor belonging to mode *(i)* and which is available even in the case (quite extreme on knowledge landscapes) that search heuristics are exclusively local. A factor belonging to mode *(ii)* is the existence or creation of new, relatively segregated application domains.

[32] A further point of departure of the application-niche argument developed in this section can be marked with respect to the relation between variety and technological progress which is to be found in the recent family of neo-Schumpeterian endogenous-growth models without scale effects (Aghion and Howitt, 1998, ch. 12; Howitt, 1999, Dinopoulos and Thompson, 1998, Segerstrom, 1998). Since the source of increasing returns is identified in sector-specific R&D employment, the growing variety of intermediate goods has the effect of diluting R&D employment across a growing number of sectors, with a corresponding dilution of increasing returns.

4.2. A multiplicity of search heuristics and the changing technology of R&D

Sections 3.2 and 3.3 brought to the fore the quite obvious remark that human exploration of knowledge landscapes exploits a vast repertoire of search-heuristics which correspond to an array of problem-space decompositions of different cover-size (Marengo et al., 2000). The main reason for the co-existence of different search heuristics is that there are different kinds of human knowledge, which for the ease of exposition and the purposes of the present discussion we may classify according to the twofold opposition: general versus local; tacit versus codifiable (Antonelli, 1999)[33].

Ceteris paribus (in particular, given the computing resources available), if the weight of the tacit and local knowledge components is relatively large, the exploitation versus exploration trade-off is more likely to be re-shaped and somewhat relaxed through more extensive reliance on experience-based and local-search heuristics. A higher accessibility to general and codifiable knowledge will lower the cost of searching away from current practices, thus assigning wider scope to non-local and goal-directed search (Gavetti and Levinthal, 2000).

The organization of knowledge creation (search activity) within and between organizations comes to depend upon three main factors. *(i)* The ruling representations of the problem spaces. *(ii)* The ruling decompositions of such representations as influenced inter alia by the perceived relevance of the different forms of knowledge and by computing resources. *(iii)* The parallel exploration of the sub-spaces through decentralized, autonomous search-processes, versus the sequential, centrally planned organization of search activity. Decentralization responds to a large number of incentives, some of which are non-cognitive and are discussed in the property-right, principal-agent and transaction-costs literature[34]. Drawing attention to the cognitive incentives, Dosi et al. (2001) insist that problem spaces are at best only nearly-decomposable and the ruling decompositions are at best an approximation to correct decompositions of the 'true' problem space. The pay-off expected from more parallel or more sequential search procedures comes to depend in the first place upon the strength of the perceived interactions between the sub-spaces of the approximate decompositions and in the second place upon the possibility and cost of achieving 'on-line', smooth integration of diverse knowledge items across decentralized search units[35]. It is worth stressing how the three factors *(i)*, *(ii)* and *(iii)* are strongly complementary, each one affecting the other two[36].

[33] Tacit knowledge is typically context dependent; it is transferred through direct interactions and created through learning by doing and by using. Codifiable knowledge is different from information (words of a language), in that it is information understood and ready for use. It is communicated as a sequence of information strings, but re-conversion of such information into knowledge is a process requiring a human-capital input often referred to as competence. In turn, the accumulation of competence requires codifiable and tacit knowledge[33]. General knowledge is that which maintains it relevance across a large number of application domains; local knowledge is application and context specific. Broadly speaking, scientific knowledge tends to be more codifiable and general, when compared to technological knowledge which contains a larger number of tacit and local components.

[34] Some contributions to the ongoing discussion are collected in Foss (ed.) (2000).

[35] A relevant set of explanatory variables and coordination mechanisms involved in knowledge governance within and across business firms are synthesized in Grandori (2000); guidelines on the factors affecting the organization of government-funded science are presented in David et al. (1999).

[36] For instance, the organization of search into larger or smaller 'patches' (see above, section 2.3) partly reflects the more or less widespread reliance on local decompositions of the (perceived) problem spaces, but the ruling

The emerging configurations of the three factors is also heavily conditioned by the output of previous R&D activity. In other words, the technology and organization of R&D changes dramatically as a result of scientific and technological progress. The last part of the twentieth century has been marked by a rising integration between the four types (tacit-local, tacit-general, codified-local, codified-general) of knowledge, a growing pay-off from knowledge codification (and a corresponding growth in codified knowledge[37]), an increasingly multidisciplinary nature of research (Rosenberg 1992) and a shift away from the centralized knowledge creation well exemplified by the R&D laboratory of A. Chandler's large corporation, towards more decentralized forms sustained by market and non-market interactions. It will not escape the reader's attention how the innovations in the information and communication technologies greatly contributed to these changes (Antonelli, 1999) by reducing the costs of knowledge communication and integration. The great majority of studies has considered how a design of micro incentives maps to a form of R&D organization[38]. It is here worth indicating a different and complementary line of investigation. This has in its focus the overall cognitive relevance of the emerging decentralized R&D structures as considered from a complex-system perspective in which the system behaviour is 'more than the sum of the parts' (Simon, 1962, p. 99). Slightly generalizing the idea that individual organizations perform the cognitive function of implementing decompositions of organization-specific search spaces (Marengo et al., 2000; Dosi et al., 2001), it is here suggested that collective organizational structures perform the cognitive function of decomposing collective problem spaces and of decentralizing the emergent sub-spaces to individual nodes in the structure[39].

4.3. Complexity, selection for modularity and knowledge revolutions

Concerning the long-term influence of R&D on the complexity (ruggedness) of the problem spaces faced by discovery, the main point to consider is if and to what extent the selection for low-pleiotropy and modularity can exert an effective control and prevail over the drive towards rising inter-connectedness, which seems to be otherwise a natural corollary of the rising dimensionality of the search spaces in the knowledge domain. Herbert Simon was quite optimistic on the effectiveness of hierarchic and near-decomposable structures to preserve what we have called 'evolvability', both in nature and in knowledge. In the closing paragraphs of section 3.1 I have been less optimistic, on the ground that selection for low-pleiotropy and modularity is successful at explaining the evolvability of individual lineages, in our case technological and scientific paradigms, but may loose its grip when paradigm shifts, that is, technological

decompositions are simultaneously the outcome of the prevailing forms of R&D organization. Kauffman and Macready (1995) simulation results on optimal patch size assume that search is strictly local. We may expect that searching through larger-size decompositions makes larger patches more attractive.

[37] On this point see Cowan, Foray and David (2000), Foray and Cowan (1997), David and Foray (1995) and the critical re-assessment in Nightingale (2002).

[38] A perspective of this kind is adopted in the so called 'new economics of science'; see David et al. (1999), Mirowsky and Sent (2002).

[39] In this respect, the network of market and non-market interactions between interconnected R&D units may be regarded as a connectionist network performing the task of problem decomposition. Future research will address the task of clarifying if and to what extent the idea is more than a useful metaphor.

and scientific revolutions are at stake. A close reading of Simon's presentation of the relation between short-run and long-run interactions in nearly decomposable systems offers a way of revealing a source of the divergence. In extreme synthesis, Simon argues *as if* the within-module equilibria resulting from the frequent interactions of a complex system are essentially *unique and stable*. In as far as the between-block interactions are concerned, the single-module equilibrium may be characterized in terms of its input from and its output to the other blocks. The module may discover different, more efficient ways of transforming input into output, but the idea is that such variations will leave input and output essentially unchanged. In fact, this is precisely what has been often *not* fulfilled during scientific and technological revolutions. For the very reason that human problem solving can be forward looking and goal directed, it can brake the constraints imposed by the ruling interface standards. When this occurs, the condition enabling the evolvability and rising to dominance of a new paradigm is the deconstruction of the pre-existing interaction matrix and the global re-design thereof aimed at creating the standards of a new (near) decomposition.

.

5. Concluding remark: are there laws of returns to R&D?

According to Olsson (2000), Olsson and Frey (2001) the distinction between incremental and radical, paradigm-shift producing changes of the knowledge stock is isomorphic to the distinction between recombinant and non-recombinant knowledge growth. This view is challenged by the foregoing discussion. Recombination of ideas is a way of generating novelty which can produce both incremental and revolutionary discoveries. Symmetrically, paradigm-shift producing novelties may or may not be the mere outcome of knowledge recombination. In fact, they conform to one of the following characteristics: *(a)* They introduce new knowledge components affecting the function characteristics of a high number of pre-existing information strings; this mode of creating novelty is non combinatorial. *(b*) They reach new combinations of pre-existing components and such combinations send fitness effects to one or more function characteristics of a large number of strings; the latter mode of discovery is indeed combinatorial.

For the above reasons, and because the addition of radical *viable* ideas to the knowledge stock requires the periodic deconstruction-reconstruction of the near block-decomposition of knowledge interactions, knowledge growth need not follow a combinatorial growth process, contrary to the view expressed in Weitzman (1996).

The technology and organization of R&D is bound to change over time in ways that are related to not only to the incentive structure provided by institutions, but also to the particular, contingent nature of the R&D output and its influence on the costs of computation and knowledge integration.

More than the slow drive towards a well defined regime of returns to R&D activity, the prediction is that of an alternation between periods of progress conforming to a given set of design rules for knowledge interaction and periods in which innovations break with the pre-existing rules until a set of standards finally emerge that fix the rules of a new modular structure. These periods will not easily map into the ups and downs of the statistical records of R&D inputs and outputs.

Appendix A.0: Scale properties of sub-space dimension

Assume that for given N and G the 'wiring' of epistatic links is random in the sense of remark 6. With this assumption in mind, it is worth considering the scale properties of the average sub-space dimension in a 'optimal' decentralization of $\{0,1\}^N$, as we let N increase holding G constant.

Remark 9: With random wiring, the ratio K/C is equal the cross-h average of the ratio N_h / $(N - N_h)$. Thus, K/C is constant if N_h^* / $(N - N_h^*)$, where $N_h^* = \frac{1}{N}\sum_{h=1}^{S} N_h = \frac{N}{S}$. The conclusion is that, with random wiring, C/K constant is supported by decentralizations such that S is constant and the average dimension N_h^* is a positive linear function of N. With sub-space dimension being a fixed proportion of N, finding the optimum configuration of the sub-spaces is intractable (is *NP-complete* like finding the optimum on the original problem space[40]).

Remark 10: As we let N increase through time, there must be eventually at some date in the future a gap between M and the cover-size in the sequence of decentralizations supporting a fixed ratio C/K. From that date onwards, M will increasingly fall short of cover-size.

Remarks 9 and 10 support the following conjecture:

Conjecture: Assume initial values of N and G, $G \leq N - 1$ and sufficiently large. Assume random wiring of epistatic links. As we let N increase, while holding G constant, the sample average S^* of the number of sub-spaces of a 'optimal' decentralization of $\{0,1\}^N$ is a non-decreasing function of N. $S^*(N_1) > S^*(N_2)$ if $N_1 - N_2$ is sufficiently large.

Appendix A.1: The structure of N-K-p fitness landscapes

The structure of N-K-p landscapes is more easily understood in the light of the following.

*Remark 11 (*Barnett, 1998, p. 21): Let $\varepsilon(\boldsymbol{x}, V, f)$ the configuration of $\boldsymbol{x}$ at the components which are non redundant arguments of $V_f(\boldsymbol{x})$. Then almost surely $V(\boldsymbol{x}) = V(\boldsymbol{x}')$ if and only if for all f such that $V_f(\boldsymbol{x}) \neq 0$ it holds true that $\varepsilon(\boldsymbol{x}, V, f) = \varepsilon(\boldsymbol{x}', V, f)$.

Neutrality is not spread uniformly on an *N-K-p* landscape; The landscape can be decomposed into subsets $Z_n(V)$, each containing only strings $\boldsymbol{x}$ such that the number of fitness contributions with the property $V_f(\boldsymbol{x}) = 0$ is exactly n. Obviously enough, the higher n, the lower the expected fitness of a string $\boldsymbol{x}$ belonging to $Z_n(V)$.

Remark 12 (Barnett, 1998): The class of *N-K-p* fitness landscapes has the following properties.

- *(a)* For large N, the probability of a neutral mutation is roughly independent of N and drops off exponentially with K.
- *(b)* The estimated mean-size of the neutral sets contained in $Z_n(V)$ and the expected neutral degree of a string contained in $Z_n(V)$ scale roughly exponentially with n; this means that the neutral degree and expected size of neutral networks fall as fitness increases. "The 'higher up' the landscape we go, the less neutrality we can expect to encounter" (p. 21).
- *(c)* The correlation structure[41] of an *N-K-p* landscape is nearly invariant with respect to p. Deviations from invariance become less significant for large N.

[40] Cf. Rivkin (2000).

[41] The usual measure of 'ruggedness' of a fitness landscape is the auto-correlation function. Weinberger (1990) and Kauffman (1993) define this function in terms of the fitness difference at successive steps along a random walk. Barnett

Remark 13 (Barnett, 1998): At sufficiently high p the evolutionary dynamics on *N-K-p* landscapes has the following properties (the parameter settings are: $S = 200$, $N = 60$, $K = 12$, $m = 0.001$, length of simulation runs = 3000 generations):

(a) Except at specific saltation episodes, the population is mostly confined to a single neutral network on which it drifts at characteristic rate which depends positively on the neutral degree of the network, hence it is inversely related with fitness..

(b) Mutation generates new strings that explore neighbouring networks. If a higher-fitness network is so encountered, the selection pressure being strong relative to mutation, there is a positive probability of a fast transfer of the whole population to the newly found network.

(c) The average distance travelled by the population after a fixed time-lag tends to be inversely related with fitness. The probability that local search finds a way out the basin of a local peak in a *N-K-p* landscape declines as the fitness of the peak increases. If we let the number of knowledge components N increase over time, then, if this occurs with the ratio $K/N \approx$ constant, the probability of finding the way out the basin of a local peak *of given fitness* declines with the rise in K and N. Indeed, as N increases, the condition for a higher probability of escape from a local peak of given fitness is very restrictive, for it requires that K/N falls sufficiently faster than $1/N$.

(d) The jumps in mean fitness described above are still observed after generation 1500 and the eventual mean-fitness level which is so attained exceeds 0.8. In simulations with identical parameter settings, except $p = 0$, mean fitness stops showing an upward trend after generation 200 and the eventual mean-fitness level does not exceed 0.6.

Appendix A.2: Neutral changes in information

Let $\{0,1\}^{N'} \supseteq \{0,1\}^{N}$, $N' \geq N$, $\boldsymbol{x} \in \{0,1\}^{N}$, $\boldsymbol{x'} \in \{0,1\}^{N'}$. The string $(\boldsymbol{x'}|\boldsymbol{x}) \in \{0,1\}^{N'}$ is such that $x'_i = x_i$, $i = 1, \ldots, N$. The fitness function $V'(\boldsymbol{x'})$ is the extension of $V(\boldsymbol{x})$ if for every $\boldsymbol{x} \in \{0,1\}^{N}$ there is $(\boldsymbol{x'}|\boldsymbol{x}) \in \{0,1\}^{N'}$ such that $V'((\boldsymbol{x'}|\boldsymbol{x})) = V(\boldsymbol{x})$. A change from $\boldsymbol{x} \in \{0,1\}^{N}$ to $\boldsymbol{x'} \in \{0,1\}^{N'}$, is local if $\boldsymbol{x'}$ is generated through a local-search procedure on the fitness landscape corresponding to $V'(\boldsymbol{x'})$ starting from some $(\boldsymbol{x'}|\boldsymbol{x})$. A *local change* from $\boldsymbol{x}$ to $\boldsymbol{x'}$ is *ε-neutral* if and only if for the given $\varepsilon \geq 0$ there exists $h \in [-1, 1]$ such that $V'(\boldsymbol{x'}) = V(\boldsymbol{x}) + h\varepsilon$. We may notice that if $V'()$ is an extension of $V()$, then there exist neutral changes from from $\boldsymbol{x} \in \{0,1\}^{N}$ to $\boldsymbol{x'} \in \{0,1\}^{N'}$.

References

Aghion P.and Howitt P. (1998): *Endogenous Growth Theory*, MIT Press, Cambridge, Massachusetts.

Altenberg L. (1995): "Genome Growth and the Evolution of the Genotype-Phenotype Map", in Banzhaf W. and Eeckman F. H. (eds.): *Evolution and Biocomputation: Computational Models of Evolution*, Berlin, Springer Verlag; revised version (1997) of the paper available at http://dynamics.org/~altenber/.

Altenberg L. (1994): "The Evolution of Evolvability in Genetic Programming", in Kinnear K. E. Jr. (ed.), *Advances in Genetic Programming*, Cambridge MA, MIT Press, pp. 47-74.

Arrow K. J., Y-K. Ng, and X. Yang (eds.) (1998): *Increasing Returns and Economic Analysis*, London, Macmillan.

Baldwin C. Y. and Clark K.B. (2000): *Design Rules. Volume 1: The Power of Modularity*, Cambridge MA, MIT Press.

Barnett L. (1998): "Ruggedness and Neutrality – The NKp Family of Fitness Landscapes"in Adami C., Belew R. K., Kitano H. and Taylor C. (eds.), *ALIFE VI, Proceedings of the Sixth International Conference on Artificial Life*, Cambridge MA, MIT Press, pp. 18-27.

Brusoni S. and Prencipe A. (2001): "Unpacking the Black Box of Modularity: Technologies, Products and Organizations", *Industrial and Corporate Change*, vol. 10, n. 1, pp. 179-205.

(1998, p. 20) defines this function in terms of the fitness co-variance across subsets of strings that are Hamming distance d apart.

Buenstorf G. (2002): “Sequential Production, modular techniques and technological change”, Max Planck Institute for Research into Economic Systems, Working Paper # 0201.
Caballero R. J. and Jaffe A. B. (1993): “How High Are the Giant’s Shoulders: An Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth”, *NBER. Macroeconomics Annual*, pp. 15-74.
Caminati M. (1999): “Complementarity, Opportunity and the Co-evolution of Techniques”, *Quaderni del Dipartimento di Economia Politica*, n. 260, University of Siena; revised edition available at www.econ-pol.unisi.it/pubblicazioni.
Caminati M. (2003): “R&D Models of Economic Growth and the Long-Term Behaviour of Productivity and Innovation”, in Salvadori N. (ed.), *The Theory of Growth. A Classical Perspective*, Aldershot, Edward Elgar, chapter 12, pp. 235-58, in press.
Campbell D. T. (1960): “Blind Variation and Selective Retention in Creative Thought as in other Knowledge Processes”, reprinted in Radnitzky G. and Bartley W. W. III (eds.) (1987): *Evolutionary Epistemology, Rationality, and the Sociology of Knowledge*, La Salle, Ill., Open Court, pp. 91-114.
David P., Forey D., Steinmuller E. W. (1999): “The Research Network and the New Economics of Science”, in Gambardella A. and Malerba F., *The Organization of Economic Innovation in Europe*, Cambridge, U. K, Cambridge University Press.
Dinopoulos E. and Thompson P. (1998): “Schumpeterian Growth Without Scale Effects”, *Journal of Economic Growth*, vol. 3, pp. 313-37.
Dosi G. Levinthal D, Marengo L. (2001): “Bridging Contested Terrain: Linking Incentive-Based and Learning Perspectives on Organizational Evolution”, Mimeo.
Eicher T. S. and Turnovsky S. J. (1999): “Non-Scale Models of Economic Growth”, *The Economic Journal*, vol. 109, pp. 394-415.
Foss N. J. (ed.)(2000):*The Theory of the Firm: Critical Perspectives on Business and Management*, London, Routledge, 4 volumes.
Frenken K. (2001): “The Organization of Search Activity in Random Fitness Landscapes”, Working Paper available at the homepage: *http://econ.geog.uu.nl/frenken.html*.
Frenken K. (2002): “A fitness landscape approach to complexity, modularity and vertical disintegration”, Working Paper available at the homepage: *http://econ.geog.uu.nl/frenken.html*.
Frenken K., Marengo L. and Valente M. (1999): “Interdependencies, Nearly-Decomposability and Adaptation”, in Brenner T. (ed.), *Computational Techniques for Modelling Learning in Economics*, Boston, Kluwer, pp. 145-65.
Gavetti G. and Levinthal D. (2000): “Looking Forward and Looking Backward: Cognitive and Experiential Search”, *Administrative Science Quarterly*, vol. 45, pp. 113-37.
Gould S. J.(1992): “The Panda’s Thumb of Technology”, in S. J. Gould, *Bully for Brontosaurus*, Harmondsworth, Penguin Books.
Grandori A. (2000): “Knowledge Governance Mechanisms and the Theory of the Firm”, paper presented at the Isdie-SPRU conference ‘Knowledge and the Firm’, Nice, December.
Hadamard J. (1949): *The Psychology of Invention in the Mathematical Field*, Princeton, Princeton University Press.
Helpman E. (ed.) (1998): *General Purpose Technologies and Economic Growth*, Cambridge, MIT Press.
Holland J., Holyoak K., Nisbett R., and Thagard P. (1986): *Induction: Processes of Inference, Learning and Discovery*, Cambridge MA, MIT Press.
Holland J. H. (1975): *Adaptation in Natural and Artificial Systems*, Ann Arbor, Univ. of Michigan Press.
Hovhannisian K. (2003): “’Imperfect’ Local Search Strategies on Technology Landscapes: Satisficing, Deliberate Experimentation and Memory Dependence”, University of Siena, Mimeo.
Hughes T. P. (1983): *Networks of Power. Electrification in Western Society*, Baltimore, John Hopkins Univ. Press.
Jones C. I. (2002) : “Sources of U.S. Economic Growth in a World of Ideas”, *American Economic Review*, vol. 92, pp. 220-239.
Kauffman S. (1993): *The Origins of Order*, New York, Oxford Univ. Press.
Kauffman S. and Macready W. (1995): “Technological Evolution and Adaptive Organizations”, *Complexity*, pp. 26-43.
Kealey T. and Al-Ubaydli O. (2001): “The Economics of Science and R&D Based Models of Endogenous Growth: A Critique”, *Knowledge, Technology and Policy*, vol. 13, n° 4, pp. 37-48.

Langlois R. N. (2001) : "The Vanishing Hand : The Changing Dynamics of Industrial Capitalism", Paper presented at the DRUID Nelson and Winter Conference, June 12-17, 2001, Aalborg, Denemark.
Langlois R. N. (2002) : "Modularity in Technology, Organization and Society", *Journal of Economic Behaviour and Organization*, vol. 49, pp. 19-37.
Levinthal D. (1997): "Adaptation on Rugged Landscapes", *Management Science*, vol. 43, pp.934-50.
Levinthal D. (1998): "The Slow Pace of Rapid Technological Change: Gradualism and Punctuation in Technological Change", *Industrial and Corporate Change*, vol. 7, n. 2, pp. 217-247.
Levinthal D. and Warglien M. (1999): "Landscape Design: Designing for Local action in a Complex World", *Organization Science*, vol 10, n. 3, pp. 342-357.
Marengo L. (2000): "Decentralization and Market Mechanisms in collective problem Solving", Mimeo.
Marengo L., Dosi G., Legrenzi P., Pasquali C. (2000): "The Structure of Problem Solving Knowledge and the Structure of Organizations", *Industrial and Corporate Change*, vol. 9, pp. 757-788.
McKelvey B. (1999): "Avoiding Complexity Catastrophe in Coevolutionary Pockets: Strategies for Rugged Landscapes", *Organization Science*, vol.10, pp. 294-321.
Mokyr J. (1990): *The Lever of Riches*, New York, Oxford Univ. Press.
Mokyr J. (1991): "Evolutionary Biology, Technological Change and Economic History", *Bulletin of Economic Research*, vol.43, pp. 127-49.
Olsson O. (2000): "Knowledge as a Set in Idea Space: An Epistemological View on Growth", *Journal of Economic Growth*, vol. 5, pp. 253-275.
Olsson O. and Frey B. S. (2001): "Entrepreneurship as Recombinant Growth", Göteborg University, Department of Economics, Working Papers in Economics n. 51.
Page S. E. (1996) : "Two Measures of Difficulty", *Economic Theory*, vol. 8, pp. 321-346.
Poincaré H (1908): "Mathematical Creation", reprinted in Ghiselin B. (ed.) (1952): *The Creative Process*, Berkeley, University of California Press, pp 23-33.
Porter M. E. and Stern S. (2000): "Measuring the 'Ideas' Production Function: Evidence from International Patent Output", NBER Working Paper N° 7891.
Rivkin J. W. (2000) : "Imitation of Complex Strategies", *Managment Science*, vol. 46, n. 6, pp. 824-844.
Romer P. M. (1990): "Endogenous Technological Change", *Journal of Political Economy*, vol. 98, pp. 71-102.
Rosenberg N. (1976): *Perspectives on Technology, Cambridge*, Cambridge University Press.
Schumpeter J. A. (1934): *The Theory of Economic Development*, Cambridge, Harvard University Press.
Segerstrom P. S. (1998): "Endogenous Growth Without Scale Effects", American Economic Review, vol. 88, pp. 1290-1310.
Simon H. A. (1962): "The Architecture of Complexity", *Proceedings of the American Philosophical Society*, vol. 106, n. 6, pp. 467-482; page number references refer to the reprint in Litterer J. A (ed.) (1969): *Organizations: Systems, Control and Adaptation, Second Edition, Volume II*, New York, John Wiley & Sons, pp. 98-114.
Simon H. A. (1973): "The Organization of Complex Systems", in Patee H. (ed.): *Hierarchy Theory*, New York, Braziller.
Simon H. A. (1977): *Models of Discovery*, Dordrecht, D. Reidel Publishing Company.
Simon H. A., Langley P. and Bradshaw G. L. (1981): "Scientific Discovery as Problem Solving", *Synthese*, vol. 47, pp. 1-27.
Thagard P. (1992): *Conceptual Revolutions*, Princeton, Princeton Univ. Press.
Topkis, D. M. (1998): *Supermodularity and Complementarity*, Princeton, Princeton University Press.
Usher A. P. (1929): *A History of Mechanical Inventions*, New York, McGraw-Hill.
Wagner G. P. and Altenberg L. (1996): "Complex Adaptations and the Evolution of Evolvability", *Evolution*, vol. 51, n. 2, pp. 329-347.
Wagner G. P., Mezey J. and Calabretta R. (2001): "Natural Selection and the Origin of Modules", Mimeo.
Westhoff H., Yarbrough B. and Yarbrough R. [1996]: "Complexity, Organisation, and Stuart Kauffman's *The Origins of Order*", *Journal of Economic Behaviour and Organisation*, vol. 29, pp. 1-25.
Weitzman M. L. (1996): "Hybridizing Growth Theory", *American Economic Review*, vol. 86, pp. 207-212.
Weitzman M. L. (1998): "Recombinant Growth", *Quarterly Journal of Economics*, vol. 113, pp. 331-360.
Wright S. (1931): "Evolution in Mendelian Populations", *Genetics*, 1, pp. 356-66.
Young A. (1928): "Increasing Returns and Economic Progress", *Economic Journal*, vol. 38, pp. 527-42.