



Mendelian explorations

Luigi Marengo¹ · Simona Settepanella² · Marco Valente³

Received: 28 March 2022 / Accepted: 5 January 2023 / Published online: 23 February 2023
© The Author(s) 2023

Abstract

In his most recent book, Daniel Levinthal develops a theory of strategy and organization centred around a “Mendelian executive” who consciously seeks new strategic opportunities but is constrained by path dependence to search in spaces adjacent to the current state of affairs. Building upon this intuition, we present a simple model of exploration and exploitation in complex landscapes showing that path dependence modifies the usual predictions of arm bandit models, especially when agents search in complex spaces where correlation of performance among adjacent alternative is low.

Keywords Exploration vs. exploitation · NK fitness landscapes · Search processes

Introduction

Daniel Levinthal’s book “Evolutionary Processes and Organizational Adaptation” is a milestone in scholarly research in strategy and organization. As the author explains, his book is an “*effort to construct a middle ground*” (Levinthal 2021a, p. 1) between rational choice theory and the bounded rationality approach pioneered by the Carnegie school.

Actually, as Levinthal himself acknowledges, the book is not meant to stand in the middle, as the author clearly leans towards the behavioural approach, which he prominently contributed to develop in new directions (see Gavetti et al. 2012, for a review). In particular, whereas rational choice theory considers human and organizational agency as an act of choice among given and equally accessible alternatives, the Carnegie school assumes that decision-making can be

better described as a process of costly and difficult search, where alternatives are not initially given to the decision-makers but must be discovered and constructed (March and Simon 1958; Cyert and March 1963). Levinthal’s book is a state-of-the-art account of this latter approach. An account where path-dependency, boundedly rational search, imperfect organizational and internal selection processes assume a central explanatory power. Rather than the middle of the segment between rational choice and bounded rationality, the book can be better interpreted as exploring a square where the opposite segment comprises the evolutionary and complexity approaches, as we are going to argue below.

Generalized evolutionary theory, extended beyond the biological building blocks of genes and inter-generational inheritance, complements the behavioural approach with the fundamental processes of selection and path-dependence. However, an evolutionary approach is in principle compatible with rational choice, as selection forces could drive a population of agents towards behaviours very close to optimality, even if none of them consciously optimizes. In a simplistic perspective, it may appear possible to define rational outcomes as the results produced by an extremely powerful evolutionary process removing any alternative to optimality (Friedman 1953). Daniel Levinthal exposes the fallacy of this perspective, showing that this “*as-if*” evolutionary justification of the rationality assumption does not hold if agents operate in a complex environment, made of many dimensions or components that interact non-linearly. Indeed, in this case the resulting “*landscape*” on which agents search is heavily rugged made of a multitude of peaks and trough.

✉ Luigi Marengo
lmarengo@luiss.it

Simona Settepanella
simona.settepanella@unito.it

Marco Valente
marco.valente@univaq.it

¹ Department of Business and Management, LUISS University, Rome, Italy

² Department of Economics and Statistics, Torino University, Torino, Italy

³ Department of Engineering and Economics, L’Aquila University, L’Aquila, Italy

Consequently, the power of local adaptation and selection, far from leading to the global peak, is bound to drive agents to the closest local peak, which may be very far from optimality (Levinthal 1997).

Thus, boundedly rational agents navigate these complex landscapes, with intentionality but also with limited knowledge of the all available opportunities but for the few locally accessible. The “Mendelian executive”, the main character in Levinthal’s book, is subject to these bounds and exposed to path dependence with a vanishingly small probability of ever reaching the global optimum. As a consequence, there is no interest in focusing on the global optima of a problem, since agents are constantly dealing with a search process aimed at local improvements characterized by a high level of uncertainty.

A useful way to characterize this kind of search process has been offered by a seminal article written by James March in 1991 March (1991). In this paper he describes the exploration vs. exploitation dilemma as one the most fundamental strategic choices. The classical model used to formalize this dilemma is the multi-armed bandit problem (Gittins 1979; Holland 1975). According to this idealized model the decision-maker must repeatedly put a coin in one out of a finite number of slot machines, each of them delivering a random reward drawn from some machine-specific distribution, unknown to the decision-maker. The strategic problem is therefore to allocate trials among the various machines balancing the two objectives of exploiting the (apparently) best machines discovered so far or exploring opportunities for possibly better results. In particular, at any generic time, the agent must choose one among three alternatives: first, bet on the machine which is currently believed to have the highest expected payoff; second, re-sample one which has been already tried before but delivered (possibly because of bad luck) a lower payoff; third, test a new machine never tried before.

This framework has been very successful and produced a good amount of valuable research in strategy and organization studies (see Denrell and March 2001; Posen and Levinthal 2012; Laureiro-Martínez et al. 2015, just to cite a few examples). However, the traditional exploration vs. exploitation approach misses an important aspect that Levinthal’s book focuses on: “*the intentionality of the Mendelian executive allows for the conscious exploration of opportunities ...but the constraining force of path-dependence tend to restrict these moves to adjacent spaces.*” (Levinthal 2021a, p. 2). In other words, exploration is usually a cumulative and path-dependent process where the discovery of new possibilities triggers the opportunity of further discoveries. Moreover, alternatives are not all given and equally accessible like slot machines in a Las Vegas casino, but are located in some space which is only accessible by following given paths. In particular, some alternatives in a neighbourhood of

those already known can be immediately accessible, while others, further away, can only be accessed by first going through some intermediate alternatives.

Kauffman (1993) call this the “*adjacent possible*” principle, which implies that every novelty has not only a value in itself, but also in the extent to which it unveils a new space of possibilities which could not be accessed before. The standard bandit model does not capture this cumulative and path-dependent dynamics. Neither does the game-theoretic/rational choice view of strategy as a selection of the best option among a given set of equally accessible alternatives. An evolutionary and complexity approach is needed, where path-dependency and local adaptation unfolding on complex spaces are the fundamental processes driving the dynamics of the systems. Levinthal (2021b) develops a similar argument adding that the standard bandit model produces a path-dependent dynamics in beliefs, but not with regard to cumulative nature of capabilities.

The business world offers many instances where we can see this principle at work. For instance, when IBM launched the personal computer, they considered only a tiny subset of all the functionalities and uses that this product was later able to offer. Most of such additional uses were actually the product of further technological discoveries and capability developments made possible by that original act of exploration combined with subsequent acts of both exploitation and exploration (Flamm 1988; Bresnahan and Greenstein 1999).

In order to address this issue we have to model the exploration vs. exploitation dilemma in some space characterized by an accessibility, or proximity, notion. Exploration can be considered as a movement to a new location which does not only give information on the value of a new location but, also, opens up a new world of possibilities which were not accessible before. A key factor is therefore the correlation among the values of neighbouring locations: if correlation is high the value of a given location is a good estimate of the values of the other locations accessible from that one. On the contrary, if correlation is low, discarding a low value location because of better alternatives may have high opportunity costs as we discard also every option accessible from the low value one, which may be high-valued.

NK landscape models have tunable correlation of fitness values as such correlation is a function of the degree of interdependency among the dimensions/elements constituting the landscape. “Simple” landscapes, whose dimensions are relatively independent, are characterized by high correlation among fitness values of nearby locations. In “complex” landscapes, whose dimensions are highly interdependent, the correlation of fitness values of neighbouring locations is low and therefore discarding an apparently bad location may have an additional cost in forgone good locations which are accessible from it. The overall outcome is that, in a simple landscape, there are many fitness-increasing paths

which lead to optimal locations, therefore in an exploration/exploitation perspective discarding a location because it apparently delivers a low value does not have dramatic consequences since there exist many alternative paths converging to high-valued portions of the landscape. On the contrary, in a complex landscape fitness increasing paths are few and far between, and very short stopping at the closest local peak. The only possibility to continue the exploration escaping local peaks, is to make a fitness-decreasing move so as to open up new areas to exploration.

In the next sections, we introduce our simple variation on a standard NK model and present some simulation results supporting these intuitions.

Exploration and exploitation in a complex space

We assume that possible locations/alternatives are points in a NK fitness landscape (Kauffman 1993; Levinthal 1997). More formally, we consider N binary components and each alternative is defined as a configuration of such components: $a_i = [a_i^1, a_i^2, \dots, a_i^N]$ with $a_i^j \in \{0, 1\}$, $\forall j = 1, 2, \dots, N$, and $\forall i = 1, 2, \dots, 2^N$. Thus there exist 2^N alternatives all located in a space where the distance between two alternatives is given by the number of different components (the so-called Hamming distance) that can vary between 1 and N .

Like in standard NK models, the expected payoff of alternative a_i is the simple average of the payoff contributions of each component:

$$f(a_i) = \frac{\sum_{j=1}^N \psi(a_i^j)}{N},$$

where $\psi(a_i^j)$ is a random draw from a uniform distribution with support on the unit interval $[0, 1]$ and is conditional on the current value of $K - 1$ other components in addition to a_i^j itself.

In our model we distinguish between the true fitness of a location, constant and determined by environment, and the actual payoff that an agent receives when sampling alternative a_i , which includes a random component redrawn every time an agent assesses the configuration. This actual payoff is given by the fitness plus a random error with expected value 0:

$$\pi(a_i) = f(a_i) + \epsilon,$$

where the fitness $f(a_i)$ is the expected value of the payoff: $E\pi(a_i) = f(a_i)$.

At the outset of a simulation ($t = 0$), each agent j starts from the location with the lowest fitness a_0^j and receives its payoff (with error) $\pi(a_0^j) = f(a_0^j) + \epsilon_0$. Location a_0^j is

assigned as the initial preferred location for every agent. Subsequently, at each time $t = 1, 2, \dots$ the agent can perform one of the two following actions:

- “Exploit”, i.e. sample again the currently preferred alternative receiving a payoff $f(a_{t-1}^j) + \epsilon_t$ where the random component is redrawn. This action is randomly chosen with probability $1 - p_{\text{explore}}$ and, of course, does not modify the preferred location that remains a_{t-1}^j . The new payoff is used to update the estimation of the fitness defined as the average of all payoff’s received from the location. Formally, if Π_{t-1}^j is the current estimate of the fitness of alternative j which has been tested $T - 1$ times so far, the new estimate is $\Pi_t^j = ((T - 1)\Pi_{t-1}^j + f(a_t^j) + \epsilon_t)/(T)$.
- “Explore”, chosen with probability p_{explore} , i.e. testing an alternative adjacent to a_{t-1}^j by mutating one bit in the string representing the current location, call it a_*^j . In this case the agent receives the payoff formed by the fitness of the newly tested alternative and the random component: $\pi(a_*^j) = f(a_*^j) + \epsilon_t$. The agent decides to reject the new location, remaining on the currently preferred one, if the payoff is below a percentage τ of the estimated fitness of the old alternative, i.e. when $\pi(a_*^j) < \tau\Pi_{t-1}^j$. In case the new payoff is higher than the estimated fitness of the current alternative, $\pi(a_*^j) > \Pi_{t-1}^j$, the agent replaces the current alternative with the new one. Finally, when the payoff of the new alternative is below the estimated fitness of the estimated fitness, but by a smaller share than τ ($\tau\Pi_{t-1}^j < \pi(a_*^j) < \Pi_{t-1}^j$) the agent determines randomly whether to accept (probability p_a) or reject ($1 - p_a$) the new alternative.

To summarize, we have the following major differences with respect to the usual search model on fitness landscapes. First, the fitness/payoff value is perturbed by some random noise term. Second, agents can decide either to resample in the current location (exploit) or move to a new one in its neighbourhood (explore). Third, in a standard NK fitness landscape agents adopt a hill-climbing strategy, i.e. they move to a new preferred location only if it delivers a higher payoff; in our model, with some probability, agents can also accept to move to a location showing an inferior payoff (but within an acceptability threshold τ). Finally, we consider two performance indicators as, in addition to the usual highest achieved fitness value, we also consider, coherently with the exploration–exploitation perspective, an agent’s cumulated payoff as her/his performance indicator comprising both the payoff from chosen alternatives and from those rejected.

Table 1 The eight-parameter combinations used in our simulations

Series label	Complexity	Accept decr.	Noise
#1	Low	No	No
#2	Low	Yes	No
#3	Low	No	Yes
#4	Low	Yes	Yes
#5	High	No	No
#6	High	Yes	No
#7	High	No	Yes
#8	High	Yes	Yes

- **noisy fitness:** we consider increasing level of random noise ϵ , starting with no noise and then introducing noise with increasing variance.

To allow comparisons among different landscapes we normalize their fitness values between 0 (lowest fitness) and 1 (highest fitness).¹

In order to simplify the description of results, we compare eight different combinations of the various parameters. Table 1 summarizes these combinations indicating the labels used for the simulation results.

We report two performance indicators. As a measure of

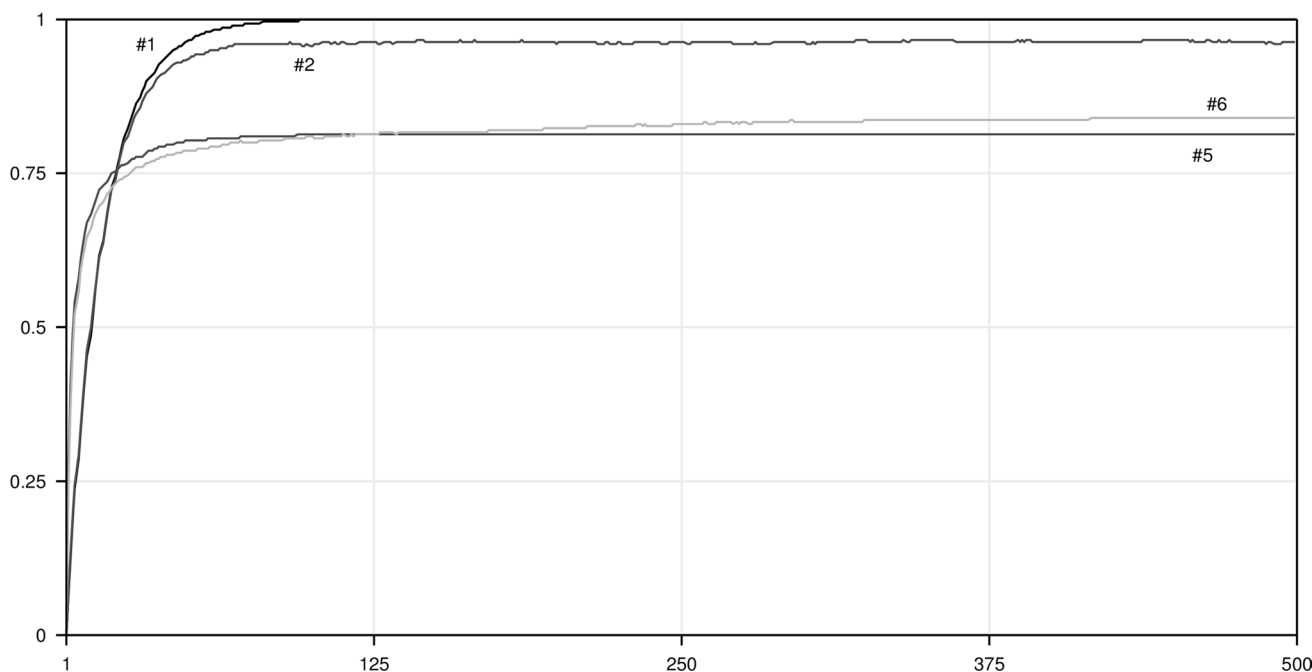


Fig. 1 Average performance without noise

In the next section we briefly report the main results we obtain by simulating this model.

Results

We tested the model with landscapes of size $N = 12$ and we vary three parameters:

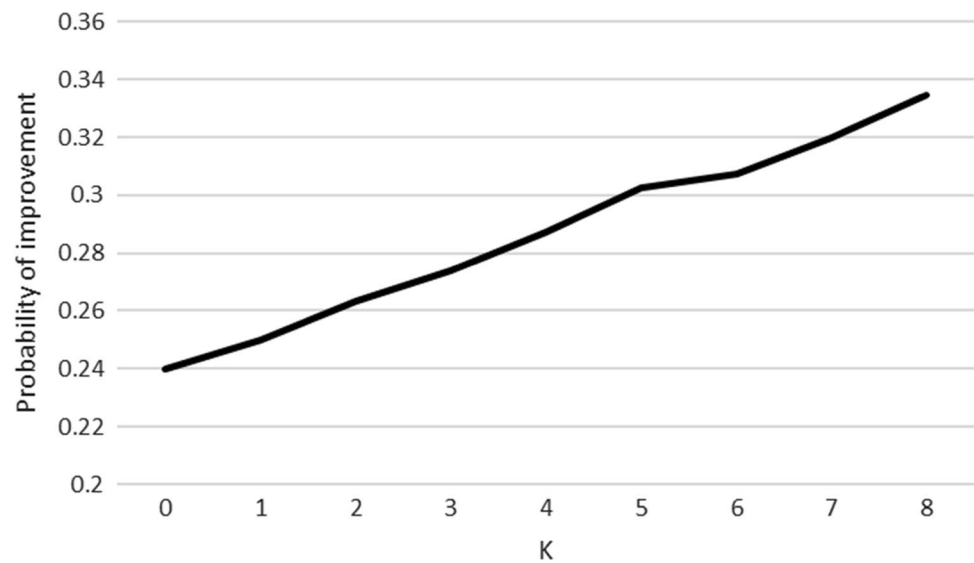
- **landscape complexity:** we consider simple landscapes with lowest complexity ($K = 1$) and highly complex landscapes with $K = 8$;
- **probability of acceptance of fitness decrements:** we compare agents who never accept fitness decrements ($p_a = 0$) with agents who accept with some positive probability $p_a > 0$.

the overall capacity to identify high-fitness points we consider the fitness of the location occupied by the agent. This index is computed on the basis of the (normalized) expected fitness $\tilde{f}(a_i)$, ignoring random noise. We also report the cumulated payoff collected by the agent up to each time step, i.e. the payoff (comprehensive of random noise) the agent has received from time 0 until the current iteration.

For each configuration we generated 1000 independent agents searching on 10 different random landscapes. The

¹ If f_{min} and f_{Max} are, respectively, the minimum and maximum fitness values of a landscape, the normalized fitness $\tilde{f}(a)$ of a generic location a whose original fitness is $f(a)$ is computed as: $\tilde{f}(a) = \frac{f(a)-f_{min}}{f_{Max}-f_{min}}$.

Fig. 2 Probability of finding a location with higher fitness following a fitness decrement across different complexity levels



results reported below are average values across these 1000 data points.

Figure 1 shows the payoffs time series for the four combinations (#1, #2, #5, and #6) without random noise.

Series #1 and #2 concern “simple” landscapes where, unsurprisingly, simple hill-climbing (series #1) converges to the unique global optimum. Series #2 reports instead the average fitness of agents who accept with a small probability ($p_a = 0.05$) a fitness decrement, and shows that in a simple landscape this generates a fitness loss comparing with the agents accepting only fitness-increasing moves. On the contrary, when the landscape is complex a small probability of accepting a fitness decrement is conducive to higher performance in the log run. Indeed, agents who do not accept any fitness decrement (series #5) are quickly stuck in a local optimum, while those who accept some fitness decrement (series #6) can move away from such local peaks and keep slowly climbing to higher portions of the landscape.

This difference of performance of the same search algorithms between simple and complex landscapes is due to their ruggedness (single-peaked vs. multi-peaked) but also, relatedly, to how informative is the fitness of a location of the fitness values of its “adjacent possible”, i.e. locations which become accessible from it. In order to highlight this aspect we have carried out the following exercise. For each possible location a_i in a landscape we list all possible one-bit mutations and, among these N neighbours of a_i , we consider only those whose fitness is lower than a_i 's. For each of these locations we compute all their $N - 1$ neighbours different from a_i , recording how many have a fitness higher than the fitness of the original a_i . Figure 2 plots the probability that a fitness higher than the initial one can be accessed after a fitness decrease. The lower such a probability, the better a fitness decrement signals that also the adjacent possible has

low fitness. The figure shows that the probability increases linearly with the complexity indicator K .

Let us now move to the more general case where fitness values are perturbed by a noise factor. Figure 3 reports the highest fitness achieved by the same search strategies as in Fig. 1 on noisy landscapes. In this case, performance on simple landscapes (series #3 and #4) is generally lower than performance obtained by agents searching complex landscape with the same strategies (series #7 and #8). The reason is that noise is a source of complexity and also play a similar role as acceptance of fitness decrements. In simple landscapes, errors in fitness evaluation constantly drive agents away from the path to the global peak. The strategy to accept lower fitness values (series #4 and #8) does not produce significant changes, as noise itself has the same effect.

It is also important to notice that performance on complex landscape increases much faster than in simple ones, regardless of the strategy employed. Only in the longer run fitness achieved in simple and complex landscapes become similar, with the latter remaining higher and slowly increasing also in the long run. The reason is that, as mentioned, we let all agents start from the location of lowest fitness and in simple correlated landscape every exploration can only produce a small increase of fitness, while in complex uncorrelated landscape such an increase can be much larger.

Finally, Fig. 4 reports average total performance, i.e. the sum of total performance (including noise, when applicable) experienced by agents at every time divided by the number of iterations. The series include therefore both the fitness obtained by *exploitation* (re-testing the current location) and by *exploration* (produced by generating a new location with a mutation).

The long term trend of this indicator is much higher in the case of simple landscapes. The reason is that when attempting a mutation the fitness obtained is similar to

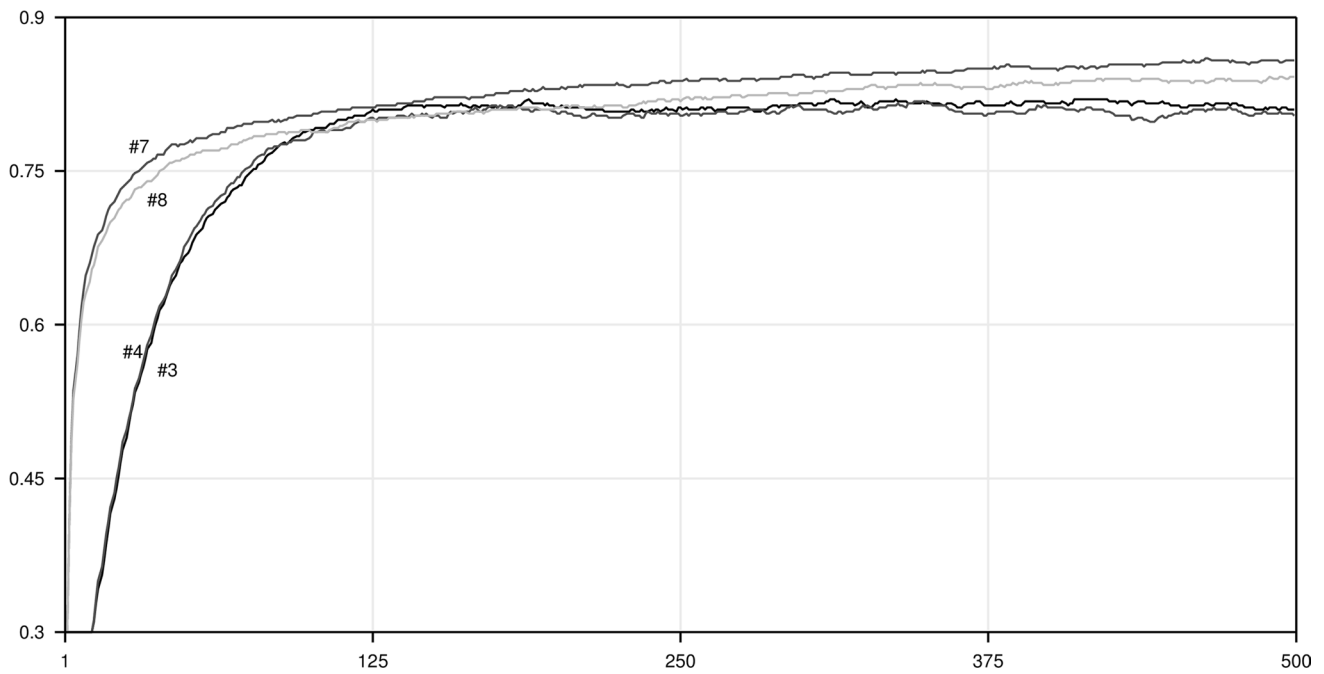


Fig. 3 Average performance with noisy fitness

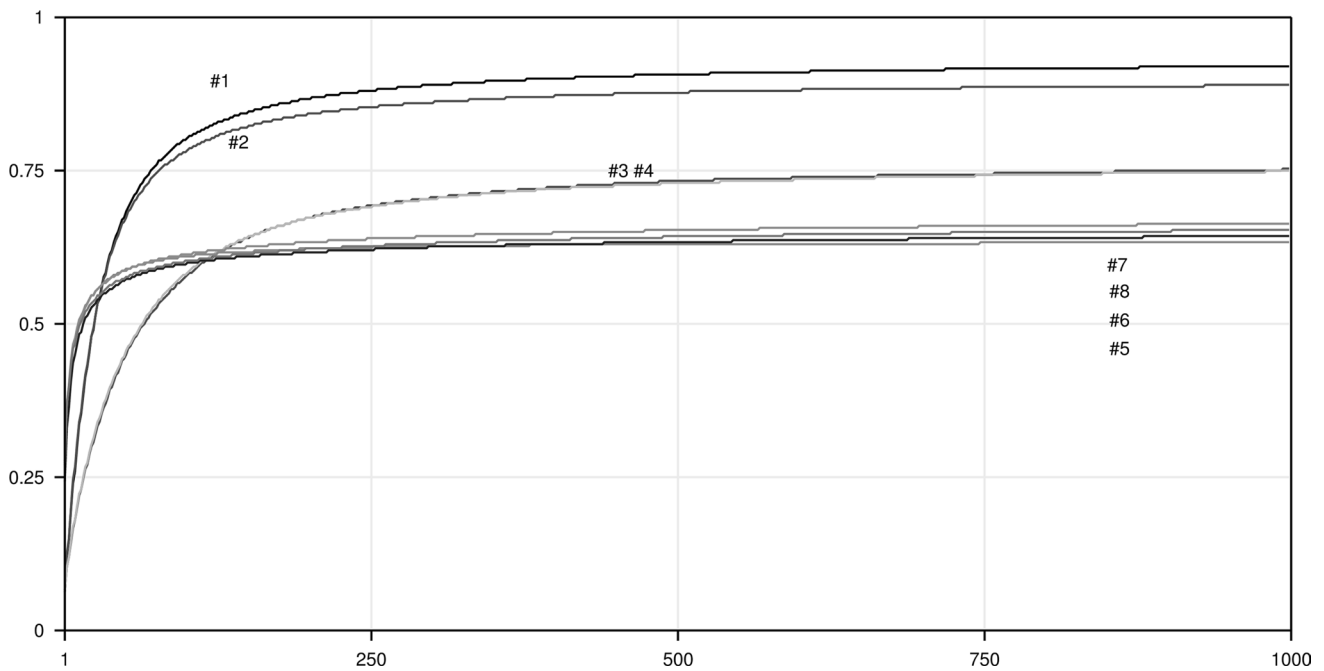


Fig. 4 Cumulated performance for all the series

that of nearby locations. Therefore, the opportunity cost of exploration are very low with respect to complex landscape, where local exploration can be very costly in terms of lower performance, especially when agents are either on or close to a peak and therefore every exploration inevitably implies

testing a location with lower expected fitness. As already explained, the opposite instead is true at the beginning of the simulations, when agents start from the worst location and uncorrelated complex landscape offer the opportunity to increase performance faster by local explorations.

Finally, differences among among different search combinations on complex landscapes are very small. However, it is interesting to notice that the ranking among these series is persistent. The best performance is generated by agents search landscapes where fitness can be observed without errors but accept fitness decrements (series #7). Then we have series #8, i.e. agents moving on noisy landscapes and accepting performance decrements. In the third position is series #6, i.e. agents searching on a noisy landscape but accepting only movements to location with higher observed performance. The worst series is #5 where strict hill-climbing in an errorless landscape determines early lock-in.

Conclusion

Daniel Levinthal's "Mendelian executive" is an intelligent and purposeful decision-maker who explores a complex world with locally constrained path-dependent actions. In this paper, we have sketched a tentative model of exploration and exploitation which, in our view, captures some basic elements of the perspective developed by Levinthal. As he argues (Levinthal 2021b), standard arm bandit models of exploration and exploitation allow path dependence only in the formation of beliefs on the payoffs of different arms, path dependence in competence development and in the accessibility of opportunities is not considered.

In this paper, we have partly addressed the latter issue by developing a simple model combining exploration vs. exploitation dynamics with a spatial structure² constraining the accessibility of opportunities and characterized by variable degrees of performance correlation among adjacent points. We have shown that in complex environments the resulting uncorrelated structure of the values of adjacent locations modifies quite substantially the traditional perspective on the exploitation vs. exploration trade-off as discarding a low performance location may prevent access to high performance ones.

Our model is only meant to be a preliminary attempt in the direction of grounding the analysis of exploitation vs. exploration trade-offs on behavioural search models where alternatives are not given but have to be discovered in a path-dependent process. We hope more investigations will follow.

Acknowledgements We thank two anonymous reviewers for their very insightful comments on a previous draft.

Author contributions All authors read and approved the final manuscript.

Funding Open access funding provided by Luiss University within the CRUI-CARE Agreement.

² Levinthal (2021b) proposes a tree structure (similar to phylogenetic trees in biology) for the space of opportunities, ours is instead a network structure which allows for multiple paths and reversibility.

Declarations

Competing interests The authors declare that they have no competing interests

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bresnahan TF, Greenstein S (1999) Technological competition and the structure of the computer industry. *J Ind Econ* 47:1–40
- Cyert R, March JG (1963) *A behavioral theory of the firm*. Prentice Hall, Englewood Cliffs
- Denrell J, March JG (2001) Adaptation as information restriction: the hot stove effect. *Organ Sci* 12:523–538
- Flamm K (1988) *Creating the computer*. The Brookings Institution, Washington
- Friedman M (1953) *Essays in positive economics*. Chicago University Press, Chicago
- Gavetti G, Greve H, Levinthal DA, Ocasio W (2012) The behavioral theory of the firm: assessment and prospects. *Acad Manag Ann* 6:1–40
- Gittins JC (1979) "Bandit processes and dynamic allocation indices". *J R Stat Soc Series B Stat Methodol* 42:148–177
- Holland JH (1975) *Adaptation in natural and artificial systems: an introductory analysis with applications to biology*. University of Michigan Press, Control and Artificial Intelligence, Ann Arbor
- Kauffman SA (1993) *The origins of order*. Oxford University Press, Oxford
- Laureiro-Martínez D, Brusoni S, Canessa N, Zollo M (2015) Understanding the exploration-exploitation dilemma: an fMRI study of attention control and decision-making performance. *Strateg Manag J* 36:319–338
- Levinthal DA (1997) Adaptation on rugged landscapes. *Manag Sci* 43:934–950
- Levinthal DA (2021a) Evolutionary processes and organizational adaptation: a mendelian perspective on strategic management. Oxford University Press, Oxford
- Levinthal DA (2021b) From arms to trees: opportunity costs and path dependence and the exploration-exploitation tradeoff. *Strateg Sci* 6:331–337
- March JG (1991) Exploration and exploitation in organizational learning. *Organ Sci* 2:71–87
- March JG, Simon HA (1958) *Organizations*. John Wiley and Sons, New York
- Posen HE, Levinthal D (2012) Chasing a moving target: exploitation and exploration in dynamic environments. *Manag Sci* 58:587–601

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.