From men and machines to the organizational learning curve

Fioretti, Guido

University of Bologna

29 November 2009
1 Learning Curves as Emergence of Routines

Learning curves were first discovered in the aerospace industry, where a large number of items must be assembled with one another in order to build an airplane [41]. This is possibly not a chance, for it has been observed that the slope of organizational learning curves is generally more pronounced in assembling operations (where organizational learning has a prominent role) than in machining operations (where individual learning has a prominent role) [19] [20]. This insight suggests that organizational learning curves may stem from the coordination of large sets of men and machines.

According to this point of view, organizational learning curves reflect a distributed development of patterns of behavior leading to the emergence of routines [28] [38], meant as recurrent but flexible patterns of action [31] [32] [12]. Routines arise spontaneously in both structured environments [11] and informal communities of practice [10] [39] [40] out of repetition of successful coordination schemes. During their development the sequencing of actions is improved, which implies that the required task is accomplished earlier.

On the contrary, learning curves disappear if routines are destroyed. For instance, it is known that if production is suspended and subsequently restarted, e.g.
because of a prolonged strike, production time is generally longer than it used to be before the interruption [5] [9] [2] [37].

A prototypical example of this vision of organizational learning is Hutchins’ detailed story of the slow emergence of a manual calculation routine among the crew of a large ship suffering a breakdown that disabled an important piece of navigational equipment while entering a harbor [23] [24] [25]. By trial and error, the crew discovered increasingly faster routines until one was found that satisfied their particular needs. In the parlance of learning curves, production time decreased while the crew was exploring the possibilities of available tools, until a plateau was reached.

In this example, a set of men had to learn to use certain mechanic tools in order to read the coordinates of the ship by observing reference points on the ground. Some operations had to be carried out necessarily before others; in other cases, the sequencing of operations was a matter of convenience. Some members of the crew had unique abilities to use particular tools; other tools could be easily used by anyone. Once a pattern of action was established the crew had formed an organization, whose component units were compounds of people and the tools they were using. Routines changed with time until an optimal sequence of operations was found, that were routed onto the organization’s units depending on their endowments and expertise.

In general, one may distinguish between the two aspects in the formation of routines:

1. Given sequences of operations must be routed on a set of organizational units;

2. A set of operations must be arranged into feasible sequences at the same time they are routed on a set of organizational units.

Pure routing problems (1) arise when customers, managers, or other actors require an organization to carry out a certain sequence of operations. Sequencing problems (2) arise when customers, managers, or other actors require an organization to carry out a certain set of operations, no matter in what sequence. Sequencing involves routing, so it cannot be found in pure form unless each operation can only be carried out by one single organizational unit.

Obviously, reality is made of a mixture of routing and sequencing problems. However, it make sense to analyze the features of extreme cases in order to understand reality. We shall see that routing problems (1) are sufficient to generate
learning curves, but more interesting possibilities arise if sequencing problems (2) are considered.

In order to disentangle the impacts of routing and sequencing on learning curves, sequencing will be considered under conditions where routing problems do not arise. In particular, learning curves arising out of pure routing (case (1) above) will be analyzed by means of an agent-based simulator in § (sec:routing), whereas in § (3) a theoretical model will be used in order to understand some features of learning curves arising out of pure sequencing problems (case (2) above).

Both in the simulations and in the theoretical model an organization is conceived as a graph. Nodes are its organizational units, edges flows of semi-manufactured goods, and routines are recurring paths between units. Organizational learning means striving towards optimal paths of operations through units, which reflects into decreasing throughput time. In § (4) a possible relation between features of these organizational units and the slope of the learning curve will be discussed.

2 Learning Curves: Routing

Learning curves arising out of routing problems can be investigated by means of the *Java Enterprise Simulator* \(^1\) (henceforth *jES*), an agent-based platform for the modelization of firms where orders composed by sequences of elementary operations are routed on a set of organizational units capable of carrying out a subset of operations each.

In the applications presented henceforth the following assumptions will be made:

- Orders of given length are random sequences of operations drawn from a uniform distribution defined over the set of possible operations;
- Orders are routed on organizational units with the criterium that, if two or more units are able to carry out the required operation, the unit with the shortest waiting list is chosen;
- The outcomes of accomplished orders are stored in one end unit, which may represent an inventory of finished products.

This model is able to generate learning curves due to pure routing problems. Customers, managers or other decision-makers generate orders, that must be routed on available organizational units in order to be accomplished.

\(^1\)Freely available at http://web.econ.unito.it/terna/jes under the GNU public license.
Figure 1: The ratio of actual throughput time to minimum feasible throughput time. Random orders composed by 10 elementary operations, drawn from a set of 10 different operations, and routed over 20 organizational units capable of 1 operation each. In figure (a), orders were routed to the unit with the shortest waiting list. In figure (b), orders were routed randomly. Figures (a) and (b) were generated with the same random seed.

Even in this simple setting organizational learning can occur, and even if the single units neither overview the whole process nor remember their own decisions. In fact, although these units behave according to a rule so simple that it does not allow for individual learning, the organization learns how to route orders to decrease throughput time.

Let us first consider simulations where each organizational unit is able to carry out one single elementary operation. This may be the case, for instance, of workers operating quite simple machines.

Figure (1a) reports the ratio between actual throughput time and minimum feasible throughput time. This ratio describes a learning curve. Orders were composed by 10 elementary operations, drawn at each step from a set of 10 possible operations by means of a uniform distribution. These orders were routed over 20 units performing one operation each; thus, each operation could be carried out by 2 units. Although outcomes generally changed with the random seed, the learning curve illustrated in figure (1a) is quite typical.

Figure (1b) shows the outcome provided by the simulator when orders are routed randomly, with all parameters and the random seed as in figure (1a). Remarkably, figure (1b) gives the impression that organizational learning is still taking place, albeit to a much smaller extent than in the case orders were routed to the unit with the shortest waiting list. However, this actually occurs because once sufficiently diverse orders have accumulated, random routing is quite an efficient
strategy. It is at least questionable whether this effect may be labelled as “organizational learning”.

Figure (2) illustrates four learning curves obtained on organizations endowed with 10, 20, 30 and 40 units, respectively — figure (2b) is the same as figure (1a). All curves have been obtained with the same sequence of random orders composed by 10 elementary operations.

The clearest pattern among learning curves that originate from pure routing is that they decrease with the number of organizational units. Figure (2) makes clear that organizational learning is the greater, the more units are available. In fact, from (a) to (b) and to (c) the plateau is ever smaller.

However, figure (2) also points to the fact that beyond a threshold where many units stay idle, adding organizational units does not improve the organization’s throughput time. This can be seen by comparing cases (c) and (d), that are nearly indistinguishable from one another.

Figure (3) illustrates a similar comparison when the set of possible operations entails 20 items, i.e., orders are still composed by 10 operations, but these 10 operations are drawn from a set of 20. So in this case there is a higher variety of orders.

Figures (3a) and (3b) refer to 20 and 40 organizational units, respectively. So far it regards the number of organizational units, they correspond to figures (2b) and (2d), respectively.

A comparison between figures (2) and (3) highlights that, so far it regards learning curves arising out of pure routing of random orders, only the number of organizational units matters. In fact, figure (3a) is identical to figure (2b), both obtained with 20 units; likewise, figure (3b) is identical to figure (2d), and both have been obtained with 40 units. Variety of orders has no impact.

Since adding organizational units improves the performance of organizational learning, one may speculate that endowing organizational units with the ability of performing several operations may produce a similar effect. Figure (4) shows that this is not the case.

Figure (4) compares figure (2a), reproduced as figure (4a), with a simulation where each unit was capable of two elementary operations, all else being equal. The outcome of this simulation is illustrated in figure (4b).

Instead of aiding the task of routing, having more flexible organizational units may have made things slightly worse. In fact, in figure (4b) the learning curve never starts to descend, and the plateau is higher than in figure (4a). If the same comparison is made when 20 units are available, no appreciable difference appears when units are more flexible. Thus, it appears that increasing the flexibility
Figure 2: The ratio of actual throughput time to minimum feasible throughput time. Random orders composed by 10 elementary operations, drawn from a set of 10 different operations, routed to the unit with the shortest waiting list. Figure (a) originated with 10 units, each devoted to one operation. Figure (b) originated with 20 units, two for each operation. Figure (c) originated with 30 units, three for each operation. Figure (d) originated with 40 units, four for each operation. Figures (a), (b), (c) and (d) were generated with the same random seed.
Figure 3: The ratio of actual throughput time to minimum feasible throughput time. Random orders composed by 10 elementary operations, drawn from a set of 20 different operations, routed to the unit with the shortest waiting list. Figure (a) originated from 20 units, each for a different operation. Figure (b) originated from 40 units, two for each of the 20 operations. Figures (a) and (b) were generated with the same random seed.

Figure 4: The ratio of actual throughput time to minimum feasible throughput time. Random orders composed by 10 elementary operations, drawn from a set of 10 different operations, routed on 10 units with the criterion of the shortest waiting list. In figure (a), each unit was devoted to a different operation. In figure (b), each unit was capable of carrying out 2 operations. Figures (a) and (b) were generated with the same random seed.
of organizational units has at best no effect, and at worst a negative impact on organizational learning.

This is possibly due to the fact that, if two or more flexible units carry out the same operation, some other operation may find no unit able to process it. This does not occur if 20 units are available, but it may become a problem if — as in figure (4) — only 10 units are there.

The shape of learning curves originating from pure routing depends very strongly on random seed, as well as on the number of elementary operations an order is composed of. The variety obtained by changing the random seed or the length of orders is very large, and no clear pattern can be distinguished by changing these two parameters. Some curves descend smoothly after an initial peak as in figure (1a), others exhibit several peaks, each followed by a descent, and a few have a very low initial peak followed by a weak descent. There are also a few instances where the descending phase never starts, but this occurs only when the peak is extremely low or inexistent.

On the whole one may conclude that with problems of pure routing organizational learning necessarily sets in if random orders generated queuing problems. The number of available units is the only relevant parameter for this kind of learning curves.

However, it is remarkable that if the simulator is fed with deterministic orders — i.e., series of operations that repeat identical to themselves — no learning curve appears. Real routing problems, and organizational learning, begin when orders are unpredictable. We shall see that this is a crucial insight in order to understand the arousal and the slope of learning curves.

3 Learning Curves: Sequencing

Let us consider a situation where the sequence of given elementary operations is not specified. An organization receives a set of elementary operations to be carried out and, similarly to the previous case, each operation can only be carried out by one organizational unit. However, contrary to the previous case, the sequence of operations is not chosen randomly and obeyed by the units, but it is chosen by the units themselves. Managers, customers or other actors require a set of operations to be carried out, but they may not care about the exact sequence. Orders are set of operations, rather than required sequences of operations. In this case, finding out optimal sequences is the task the organization has to learn.

In general, sequencing involves routing. In fact, once operations are arranged
in a sequence, they must be routed onto organizational units. However, by considering organizations where each operation can only be carried out by one single unit we can ignore routing and focus on pure sequencing. We shall do so in order to deal with a problem that is simple and opposite to the one of § (2).

However, even if in this simple setting managers leave organizational units the freedom to arrange a set of operations in any possible sequence, organizational units in general do not combine the required operations in any conceivable way. Technical, legal or administrative constraints generally force organizational units to work on subsets of all possible sequences. Thus, organizational units must classify the sequences of operations that they receive, distinguishing those on which they can carry out their operation, from those on which they cannot operate.

Henceforth, all sequences with an operation at a position such that it can processed by a particular unit will be called the set of feasible sequences for that unit. Organizational units must be endowed with categories in order classify the sequences proposed by other organizational units as feasible or unfeasible. Only feasible sequences are accepted and scheduled for processing.

Categories may be coarse or sharp, depending on the ability of a unit to perform its operation at various stages or at different stages — e.g., more flexible machines may not care whether a certain operation has been performed before they perform their own one. Coarseness of categories will be often referred to as the flexibility of an organizational unit.

Sequences of operations are strings of integers representing one operation each. Let us represent categories by means of strings made of the integers used to represent operations, plus “don’t care” characters #s. A category classifies all strings having either the same number at the same positions, or whatever number where the category has a #. Figure (5) illustrates an example.

Here are some assumptions that allow a simple analytic treatment of organi-
zational learning arising out of sequencing problems:

- Each operation can be carried out by one and only one unit. Consequently, this is a problem of pure sequencing.
- Each organizational unit has one instance of all categories represented in the organization.
- Any sequence is feasible for at least one category. Thus, each organizational unit can accept strings from any other.
- A category has the same probability to process any feasible sequence. Contrary to classifier systems, no strength is passed on.

Let $N \in \mathbb{N}$ denote the number of possible elementary operations. Let $L \in \mathbb{N}$ denote the length of sequences of operations, as well as the length of categories.

Let $H \in \mathbb{N}$ denote the number of different categories available in the organization. It is $H \leq (N + 1)^L$, the number of dispositions with repetition of $N + 1$ elements (the $N$ operations plus the #) of class $L$.

Let $K \in \mathbb{N}$ denote the number of different sequences produced in the organization. It is $K \leq N^L$, the number of dispositions with repetition of $N$ elements of class $L$.

Since categories exist in order to make classifications, it must be $H < K$. Thus, $H$ can never reach its upper bound.

Let $p$ denote the probability that a link exists between any two organizational units. This parameter captures the range of possibilities for changing the production routines. In fact, the more possibilities for connecting the stages of production, the more possibilities there are for creating new routines.

Links are not established without a purpose. Links represent flows of semi-finished products through organizational units until an end unit is reached, where the good is finished. For instance, in the simulations of § (2) there was one single end unit, representing an inventory of finished products. Thus, searching for better routines means finding shorter paths to the end unit.

This search is generally not random. In general, it reflects an organization’s ability to adopt better procedures, discarding the inefficient ones. Let $r$ denote the probability of eliminating the unproductive links departing from an organizational unit. Thus, $r = 1$ corresponds to a perfect decision procedure whereas $r = 0$ corresponds to a random walk.
Huberman et al. derived a learning curve as a function of these parameters [36] [22]. In its turn, the following model links $p$ and $r$ to observable magnitudes such as $H$ and $K$ [18] [17].

However, a warning is in order in this respect. Organizational units are compounds of men and machines. Whereas $H$ and $K$ are easily observable so far they reflect machine features, objective measurement may not be available so far these magnitudes reflect human features.

Let us consider one single attempt to connect two organizational units. Let us assume that the probability that a category accepts a sequence is $1/(K - 1)$. This is a rough approximation, because (a) there are many more sequences than elementary operations or units, and (b) a category does not classify all sequences, but only the feasible ones. However, (a) and (b) push in opposite directions so one may hope that they nearly cancel one another.

Since each unit owns all categories, the probability to establish at least one link between any two organizational units is:

$$p = \frac{H}{K - 1}$$

with $K - H \geq 1$.

Parameter $r$ represents the probability that the search for better arrangements is effective. In the limit of infinite attempts to establish connections to other units, sooner or later the end unit is reached. On the contrary, if novel connections are no longer tried, the end unit may never be reached because the same connection to one and the same unit is endlessly repeated.

Let us calculate the probability to get stuck in connecting to a particular unit again and again. Parameter $r$ will be its complement to one.

Equation (1) expresses the probability that at least one link is established between any two units during a procedure where $H$ trials are made. The probability that this happens twice if the whole procedure is carried out twice is $(H/(K - 1))^2$. And so on. If these probabilities are summed it is safe to divide by a coefficient $K - 1$ in order to ensure that the sum will be less than unity. In the end, the probability to repeat a connection endlessly is the sum of $H/(K - 1)^2 \sum_{i=0}^{\infty} (H/(K - 1))^i$, which amounts to $H/(K - 1)(K - H - 1)$.

Consequently, the probability to choose the right path is $1 - H/(K - 1)(K - H - 1)$, which can be written as:

$$r = \frac{(K - 1)^2 - KH}{(K - 1)(K - H - 1)}$$

11
with $K - H > 1$.

Note that eq. (2) is quite a good estimate of $r$ when routines are under construction, but it is definitely wrong once good routines have been established. In fact, it captures the extent to which organizational units try novel paths, which is good at the beginning but is bad once a good routine has already been found. It makes sense in a theoretical model that estimates possible learning curves, but it would not fit a simulation logic as in § (2).

The implications of equations (1) and (2) become clear if one plots them for various values of $H$ and $K$. For obvious reasons, interesting values appear when the difference $K - H$ is not too small with respect to the absolute values of $H$ and $K$.

Equation (1) is defined for $K - H \geq 1$ but it is trivial if $K - H = 1$. Equation (2) is defined for $K - H > 1$. Thus, the smallest possible value of $K - H$ is 2. Correspondingly, the range of values of $H$ and $K$ should be close to the origin. Figure (6) illustrates $p$ and $r$ for $H \in [1, 10]$ and $K \in [3, 12]$.

The higher the parameter $p$, the more attempts are made at improving on the current arrangement of production. Equivalently, the higher $p$, the steeper the learning curve.

Thus, figure (6) shows that the greater the number of operation sequences and categories, the more possibilities for improvement. In short: the more is there to learn, the more can be learned.
However, learning may not proceed if the search for better arrangements of production is stuck in vicious circles. The parameter $r$ captures the likelihood of this possibility.

Figure (6) shows that the greater the number of sequences and categories, the more likely that no improvement will take place at all. To be concise: the more is there to learn, the more likely that nothing will be learned at all.

Thus, figure (6) illustrates a trade-off between the possibility to improve the arrangement of an organization and the danger to get lost in endless search. In fact, the more the possibilities for improvement, the more difficult it is to realize them.

Let us consider an organization with fewer categories. Fewer categories often means more generic categories. This means that workers have a wider knowledge so they can do more diverse jobs, or that machines are more flexible so they can process a wider range of semi-finished goods, or both.

Let us choose $K - H = 3$. Figures (7) and (8) show the ensuing effect on $p$ and $r$, respectively.

Even with so small a change in the number of categories, differences are impressive. The possibilities for improvement — captured by the parameter $p$ — have slightly decreased. On the contrary, the likelihood that better arrangements are found — captured by the parameter $r$ — have increased dramatically. Furthermore, the greater $H$, the more pronounced are these effects.

Figures (7) and (8) suggest that, by employing a few general categories, a large gain in effectiveness can be attained at the expense of a small loss on the possibilities for improvement. An organization of open-minded generalists and flexible machines may loose a fraction of the learning possibilities afforded by specialization, but will not get stuck in meaningless routines leading nowhere.

4 Final Discussion

Learning curves would be a valuable tool for business planning, if they were predictable. The trouble is that this is generally not the case. The slope of the learning curve is something of a guess, and it may even happen that no decreasing pattern curve sets in. Given that there is always a small but positive probability that the learning curve will not set in, it is hard for managers to rely on it in the evaluations of future costs.

It is obvious that it is necessary to understand the reasons why learning curves arise in order to be able to predict whether they will not. This chapter moved
Figure 7: Parameter $p$ when $K - H = 2$ (thin line) and $K - H = 3$ (thick line).

Figure 8: Parameter $r$ when $K - H = 2$ (thin line) and $K - H = 3$ (thick line).
from the idea that organizational learning is grounded on the formation of routines and attempted to highlight some features on which the shape of learning curves depends.

In § (2) we found that orders must be produced randomly in order for learning curves to exist, and that organizational units must be sufficiently many for learning curves to be effective. In § (3) we found that there must be sufficiently many things to do for learning curves to set in (large $H$ and large $K$), and that organizational units must be sufficiently flexible to enable the formation of routines (large difference between $K$ and $H$). Possibly, these findings point to a common pair of principles for organizational learning to take place, namely, that (i) there are sufficiently many novel possibilities for conceiving novel routines, and (ii) organizational units are sufficiently many and sufficiently flexible to implement novel routines.

Point (i) is exemplified by the case of technical innovations, that quite often take place at the same time organizational units are striving to develop better routines [4]. A few cases where innovation was totally absent [6] [13] [33] highlighted that without continuous stimulation and injection of novelties the learning curve reaches a plateau [21] [29] [7] [1]. On the contrary, a changing and stimulating environment is beneficial to both production time [35] [30] [34] and qualitative improvements [27].

Point (ii) is exemplified by the fact that, among industrial plants, learning curves are most pronounced where assembling operations are involved [19] [20]. Assembling operations require a large number of units that must be flexible enough to interact with one another in multiple configurations, a circumstance that facilitates the emergence and modification of routines. On the contrary, plants based on conveyor belts are not the typical settings where organizational learning curves arise.

More detailed simulations are in order. It is necessary to integrate all factors giving rise to organizational learning curves and to investigate their consequences beyond the level allowed by mathematical models, which is only possible by means of numerical simulations.

The application of concepts derived from numerical simulations to real cases poses still another kind of problems, for organizational units in general are not just machines, but compounds of men and machines. The features of machines can be easily measured, those of human beings often can not. Human beings exert a large influence on learning curves, as testified by the fact that the slope of the learning curve may differ across identical plants of the same firm [8] [42] [13] [14], or even across shifts in the same plant [3] [15] [2]. These episodes suggest
that there are some limits to the extent to which learning curves can be managed and predicted.

References


