Creative Development: Patterns of Learning

Jonathan S. Feinstein

John G. Searle Professor of Economics and Management
Yale School of Management
PO Box 20-8200 New Haven CT 06520-8200 USA

Introduction

Creativity and innovation are primary factors driving human economic and cultural development (for example, Aghion and Howitt (1992)). However, there is a considerable gap in developing models of human creativity in relation to economic and social systems. Further, the existing field of creativity studies focuses on the psychology of creativity, whereas it is vital - especially for links with systems theory and the social sciences - to view creativity and innovation as rooted in and growing out of knowledge structures and culture. My research program lies at the intersection among these fields - economics and social systems, creativity and innovation, and knowledge representation and decision sciences.

Beginning with my book *The Nature of Creative Development* (Feinstein (2006)) I have developed a framework to model the creative development of individuals engaged in creative endeavors in fields of knowledge and practices. This framework enables me to model and visualize the rich paths of development individuals follow, often for years, in rich knowledge and cultural environments, through which they come to be able to make important creative contributions (related work is Gruber's well known study of Charles Darwin (Gruber (1974)), among many others, and a host of biographies). In turn this framework opens the way to analysis of several important topics. One is how fields develop through the successive contributions made by individuals and cohorts over time. The framework enables us to see how the contributions influence others' creative paths, sometimes leading, only after years and many other influences, thus quite indirectly, to important creative contributions. A second important topic is education for creativity and innovation - what is the optimal curriculum to spur subsequent creativity and innovation? How can creative development best be supported? In this paper I focus on the second question. I present a model in which a cohort of individuals who enter a creative field are first taught a common core currciculum, then each individual receives a signal providing intuition about creative possibilities in the field, pursues his or her own individual path of creative development, and lastly engages in a creative project that builds a new field element using elements he has learned.

Creativity is defined, foundationally, as connecting or relating two preexisting elements - ideas or material elements - that have not previously been connected or related. Thus when a new element is created in a field, it is based on connecting together two elements in the field (or potentially from outside the field) that have not been connected previously. Geometrically and mathematically this means that the field develops as a lattice-like structure; this in turn is a bridge to the important work of Wille (1992) and Ganter, Stummer and Wille (2005) on lattice structures for representation of knowledge. I have developed these ideas in a series of papers, including Feinstein (2011; 2013) and current work. I hope my approach will spur further work linking creativity with knowledge representation and more broadly complexity research.

At the core of my work is a focus on individual learning in complex knowledge environments. In broad terms, creative development is the process of exploratory learning, guided by intrinsic creative interests and intuitive signals about potentially valuable pathways to explore. Through this process, over months and years, individuals develop creative insights and ideas, spawning creative projects they engage in through which they (attempt to) generate contributions in their chosen field (Feinstein (2006)). In fact there are two important challenges in modeling this process. First, we must model the rich knowledge (and cultural) environment within which individuals are situated - I call this the Knowledge Structure (KS). This environment involves thousands if not millions of elements, organized into hierarchies of breadth and depth, and an individual can learn just a small percentage. I model this knowledge environment using knowledge respresentation, building on the important work of Sowa (1984, 2000), Helbig, et. al. (2006), and Wille and colleagues (Wille (1992); Ganter, Stummer and Wille (2005); also related are strictly semantic representations like WordNet (Princeton (2013)). Each element that is generated in the KS has a given value: in this paper I specify the a parameterized value distribution, which has the right-tail skewed form familiar from work on innovations and creativity (e.g. log-normal / Pareto form) (see Silverberg and Verspagen (2007)). The second important challenge is to model the nature of the intuition that individuals gain, that guides them in their formation of creative interests and choices of what to learn. I model this intuition as intuitive signals that individuals gain based on their prior learning. Specifically, individuals initially learn a set of elements, in my current work a "core curriculum" that is taught to them. This initial learning could also involve additional topics they learn and specific elements they encounter or experiences they have, e.g. personal events or interactions. This initial round of learning typically involves a range of elements, many of which have never been directly connected (different subfields or topics). Individuals receive intuitive signals of the potential value of such new combinations and can choose to explore them - learning in general about them and specific knowledge elements, then can pursue creative projects to attempt to create such new combinations. The signals are informative but not perfect, and enable individuals to form updated conditional distributions about particular topic areas or topic combinations, or even specific elements. Intuitive signals may be at a broader conceptual level, of the kind that states "it may be fruitful to combine topic areas A and B" or at a more specific level, including a negativity signal in which one recognizes that a previous specific attempt to combine two topic areas that failed was misguided and that in fact an alternative approach is possible involving different specific elements within these topics. In this paper I focus on a single class of signals about new topic combinations; generalizing the model to incorporate the variety of intuitions is clearly important to help understand the diversity of patterns of creative development.

The Model

There is a cohort of individuals (two in results presented below) who are trained in a common core curriculum in period 1 then in period 2 explore individuals paths of creative development, producing a creative output / innovation. I use simulations to determine the optimal core given parameters and state of the field and identify optimal independent paths of development. The Knowledge Structure defines the field at the time the cohort enters the field. The KS includes a set of topics, denoted T, and additional helper elements denoted R that are available to help "connect" topics or paste elements together, to aid in generating new field elements (related for example to conceptual blending (Fauconnier and Turner (1998)). Topics are combined to produce new topics. The original KS includes a set of *base* topics and a set of additional topics built from combining these base topics. Topics include a list of elements that have been created under them; a topic may have no list elements, or a list up to some maximum number. List elements are created by combining two list elements under the two topics that combine to generate the given topic. Thus if T_3 is combination of T_1 and T_2 then list elements T_{1i} and T_{2j} can be combined to create a new list element under T_3. Alternatively, new elements can be created by extending an existing list element under a given topic by adding an additional R element. Each topic has associated a value τ that contributes to the value of elements generated under this topic. The τ distribution for this paper is a discrete distribution based on a log-normal (mean 1 and standard deviation 1) that is truncated at 10 with an additional probability mass at 0.01 to capture the possibility of a topic of very low value (distribution is normalized such that the cumulative probability is 1). The discrete distribution is shown in Table 1. The τ's are independent and identically distributed. When an element is created in a topic its total value depends on the τ for that topic, as well as the values of the parent elements being combined, and an idiosyncratic random element denoted ϵ that captures the inherent randomness of value of a new creative product. Denoting the topic as T_k which is the combination of topics T_i and T_j, when elements T_{ig} and T_{jh} are combined to form a new element under k, say element km, the value is: $V(T_{ig})^{\delta_p} V(T_{jh})^{\delta_p} \tau_k \epsilon_{km}$. Here δ_p is a parameter that governs the importance of a parent's value (set at zero for simplicity in the simulations). When a new element is created from an existing element, by appending an additional helper element, for from

$T_i - R_j$ building $T_i - R_j - R_k$ then only the single parent is used, with twice the power $(2 * \delta_p)$.

Learning: The model is a model of learning in two phases, core then independent, followed by creative production. There are costs to learning which limit how much individuals can learn. As a result choices about what to learn - explore - are central in the model. This captures the principle of creative development as an exploratory learning and creative process. Learning costs are best thought of as monetary costs for the core component, compensating instructors and paying for class space and so forth. For the independent learning phase costs are more likely to be time costs as it takes time to learn new material, and this limits how much can be learned in a given period; however there may also be a financial component to the independent learning. Learning costs are specified as follows. To learn a new topic costs c_0 (set at 1 in the simulations). To learn a new topic that is based on combining two parent topics requires learning the parent topics first, thus may cost as much as 3. In terms of a core curriculum this raises as an interesting issue the degree to which it is best to teach “parent” topics that may help individuals learn “child” topics tied to their signals or recognized as having high creative potential. Once a topic has been learned, each list element costs c_R (set at 0.5 in the simulations). There are thus economies of scale for learning multiple list elements under a given topic. Creative production also has limits that constrain what is possible. Specifically, in the creative production phase an individual can explore up to k_p potential new elements (set at 5 in the simulations) under a single topic. She is able to determine the value if produced of each of these elements then produces the maximal element from among this set. Overall the individual's goal is to maximize the value of what she produces and she chooses a topic accordingly.

The model unfolds as follows. The first period all individuals in the cohort are taught the same core. The core is chosen by the administrator and is chosen to be optimal given the core budget B_c and the strategies it is predicted individuals will follow in the second period. Results are presented varying the budget level. The core is efficient in that a single cost is incurred to teach everyone. After learning the core each individual then chooses a topic on which to focus - either an existing topic or a new topic (based on combining two existing topics) for which they will create the first list element. Individuals gain intuitive signals of the potential value of topics they might pursue, which guide them in their creative development. In the version of the model for which I present results here signals are generated based on the core that is taught; in extensions individuals may also received signals based on personal learning or experiences they have. Thus for the model here each pair of topics taught in the core can generate a signal, and this defines a pool of potential signals. Each individual receives one signal from this pool. The signal can be either High (H) or Low (L). Signals are independent across individuals, thus it is possible for two or more individuals to receive signals about the same $T_i - T_j$ combination. In the model presented here a signal refers to a partition of the τ distribution: a L signal indicates the true τ value lies in the interval $[.01, \sigma_{b-1}]$ and a H signal indicates the true τ value lies in the interval $[\sigma_b, 10]$; σ_b is set at coordinate 4 in the simulations. Thus signals are correct, but not fully revealing of the true value of τ. Individuals receive their signals at the end of the first period. At the beginning of the second period each individual chooses which topic to focus on and is given the oppotunity to learn additional elements, choosing those which will be those most valuable for producing a creative work under this topic. There is a budget B_l for this independent learning period. This budget constraint is best thought of as a time constraint, but may also involve a monetary component.

A key component of the model is the inference procedure evaluating the probability distribution for each τ given what is know about that topic. Initially the τ distribution is the prespecified distribution described above and displayed in Table 1. There are two ways information is gained leading to an update of this baseline distribution. One way is when list elements are produced under this topic. Recall from the formula above that the value of a list element is the product of (i) parent values (one or two), (ii) τ, and (iii) ϵ. Given that the parent values are known inference about the τ for this topic is based on the observed value decomposed into the product of the τ value and the ϵ for this element. In particular, given the ϵ distribution, which is uniform, standard Bayesian updating is used to update the τ distribution. If there is more than one list element updating can be done sequentially over each list element in turn. The other way information is gained about a τ is when an individual receieves a signal about the value of this τ. As defined above signals are partitions of the τ distribution. Thus a signal confines τ to either the lower (L) or upper (H) partition, and the resulting τ distribution is then the original distribution only within this interval with probabilities over the support of this interval normalized to add to 1. Since list elements are created prior to period 1 the signaling update is based on the distribution for this τ that has been generated taking into account any updates based on list elements.

Solution: The model is solved using standard decision analysis and game theoretic methods. Computer simulation is used to solve for a Nash equilibrium over strategies for the individuals in the cohort. A strategy specifies, conditional on the signal the individual receives, what topic he chooses to focus on, the set of elements he selects to learn given the learning budget constraint in the second period and his topic choice, and the set of projects he will explore (k_p or fewer if there are fewer feasible projects given his topic and learning choices). In fact we focus on a particular class of strategies for the simulations. In this class a strategy divides into two parts. If the individual receives a H signal for a topic for which the expected value of what he will earn pursuing this topic, given the τ distribution updated based on his signal, he pursues that topic. Otherwise, he follows a mixed strategy randomizing over a set m of topics that are the best topics to pursue given the strategies of others, the current τ distributions, and available elements to learn. My focus is thus more restrictive than the full class of possible strategies in two ways. First, I assume that all individuals who follow the mixed strategy portion of the strategy follow the same mixed strategy, so that this is a symmetric mixed strategy equilibrium. Second, I assume pure strategies for the H signal component of the strategy - an individual either pursues the topic to which the signal refers or not based on whether the expected value exceeds a threshold we compute. It is possible in some cases for an equilibrium to require a mixed strategy over whether or not individuals pursue a H signal topic, since two or more individuals may receive a H signal about the same topic. However, I am able to compute the more intuitive strategies with a pure strategy for the H signal in all cases considered in this paper. I compute an equilibrium for the class of strategies under consideration by identifying a pool of m "top" topics, a probability p_i of playing topic i in this pool, with $\sum p_i = 1$, and a threshold t_{crit} for H signals such that for every possible signal the individual receiving that signal will find it optimal to follow the defined strategy, in particular pursuing the topic to which the signal refers if the signal is H and the expected value of pursuing that topic, given the strategies played by all others (including the possibility of one or more receiving the same H signal) exceeds t_{crit}. The mixture probabilities are chosen such that the expected value of each topic in the pool is equal, taking into account that some individuals may choose a given topic because they receive a H signal for it, and that more than one individual may choose the same topic, whether because of signals or simply due to how the mixture choices work out. Finally, each topic not in the pool must be shown to provide lower expected value if a single player were to deviate and pursue it.

The model extends to a framework for modeling how a field develops over time - cohorts enter over time, each going through the process of creative development. In Feinstein (2013) I have developed such a model and analyzed the model via an extensive set of simulations. The field begins from an initial states and grows as individuals enter the field and make new contributions; its basic structure resembles a lattice. Individuals working in the field follow a defined process of creative development similar to that outlined above, and new elements that are created are added to the field, so that the field grows over time. The simulation analysis reveals a set of key features that characterize the development of fields through this process. Here are two key results. (1) There is a rich diversity of possible paths of development; this diversity is generated especially by the intuitive signals individuals receive, which lead them to attempt to make elements they might otherwise not pursue, thus shaping the development of the field in important ways. (2) The results also reveal a high degree of path dependence, generated as individuals build on the work of their predecessors, and interesting temporal patterns for how output in one period is linked with what occurred in the previous period.

Results

Table 2 presents results for a set of simulations of the model. For these simulations the number of base topics is 5 and the number of additional topics in the KS is 17, thus there are 22 topics in total when this cohort enters the field. Depth of topics is two. For this case there are 91 topic areas that can be focused on representing the sum of all current topics except base topics (assumed not to be a focus for further creative work) plus all new combinations of existing topics. Most of these topics are thus new, have not been defined previously and have no list elements. Results are shown for the case of a cohort of 2 individuals. The simulations vary the core budget as well as the independent learning budget. In particular, results are presented for a core budget of 6 and 12 and for independent learning budgets of 0 (no independent learning), 1, 2 and infinite (no constraints - individuals can learn what is required to focus on any feasible topic). Recall that the cost to learn a topic is 1 plus 1 for each parent topic that has not been learned in the core, thus up to 3, and the cost to learn individual elements under a topic is 0.5. Thus a budget of 1 is very small allowing at best one new topic or two elements under an existing topic; 2 allows more learning, a topic plus two list elements or four list elements under previously learned topics.

Each cell in Table 2 shows the optimal core for that cell as well as the characteristics of the equilibrium identified for that cell, including the expected value received by each individual, which is equivalent here to the per capita expected social welfare; the size m of the pool of topics over which individuals play a mixed strategy, and the probability an individual chooses a topic not in the pool due to receiving a H signal leading to an expected value for the signal topic that is at least as great as the expected value from playing the mixed strategy over the pool, together with the number of such topics for which a H signal triggers pursuing that topic. It is noteworthy that expected values are higher for the larger core, though by far less than a factor of two. In addition, the optimal core values with the independent learning budget for the lower budget core but not for the higher budget core. Finally, for most budget levels there is a significant probability of choosing to pursue a topic for which one has received a H signal. At the IMCIC I will present additional results.

References

Aghion, P. and P. Howitt (1992): "A model of growth through creative destruction," *Econometrica*, 60, 2, pp. 323-51.

Fauconnier, G. and M. Turner (1998): "Conceptual integration networks," *Cognitive Science*, 22, 2, pp. 133-87.

Feinstein, J.S. (2006): *The nature of creative development.* Stanford, CA: Stanford University Press.

Feinstein, J.S. (2011): "Optimal learning patterns for creativity generation in a field," *American Economic Review Papers and Proceedings*, 101, 3, pp. 227-32.

Feinstein, J.S. (2013): "The creative development of fields: learning, creativity, paths, implications." Working paper, available on SSRN or www.jonathanfeinstein.com.

Ganter, B., G. Stummer and R. Wille (2005): *Formal concept analysis: foundations and applications.* Berlin: Springer.

Gruber, H. (1974): *Darwin on man: A psychological study of scientific creativity.* New York: E.P. Dutton.

Helbig,H., (2006): *Knowledge representation and the semantics of natural language*. Berlin, Heidelberg : Springer-Verlag.

Princeton University (2013): *WordNet.* Address: wordnet.princeton.edu .

Silverberg, G. and B. Verspagen (2007): "The size distribution of innovations revisited: an application of extreme value statistics to citation and value measures of patent significance," *Journal of Econometrics*, 139, pp. 318-39.

Sowa, J.F. (1984): *Conceptual structures: information processing in mind and machine.* Reading, MA: Addison-Wesley.

Sowa, J.F. (2000): *Knowledge representation: logical, philosophical, and computational foundations.* Pacific Grove, CA: Brooks/Cole.

Weitzman, M.L. (1998): "Recombinant growth," *Quarterly Journal of Economics*, 113, 2, pp. 331-60.

Wille, R. (1992): "Concept lattices and conceptual knowledge systems," *Computers and Mathematics with Applications*, 23, pp. 493-515.

Table 1: Tau Prior Distribution

Value	0.01	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0
Probability	0.364	0.176	0.138	.0962	.0673	.0482	.0353	.0265	.0203	.0157	.0124

Table 2: RESULTS

Independent Learning Budget	**Core Budget**	
	6.0	**12.0**
0	**Optimal Core:** Base Topics 2 & 4 Per Capita EValue/ Social Value: **1.63** *m* pool: 1 Prob (choose outside pool): .61 Number of Topics in choice set outside pool: 1	**Optimal Core:** Base Topics 1,2,3 & 4 and non-base Topic 11 Per Capita EValue/ Social Value: **2.29** *m* pool: 1 Prob (choose outside pool): .42 Number of Topics in choice set outside pool: 11
1	**Optimal Core:** Base Topics 2 & 3 Per Capita EValue/ Social Value: **2.75** *m* pool: 1 Prob (choose outside pool): 0 Number of Topics in choice set outside pool: 0	**Optimal Core:** Base Topics 1,2,3 & 4 and non-base Topic 11 Per Capita EValue/ Social Value: 3**.10** *m* pool: 1 Prob (choose outside pool): .39 Number of Topics in choice set outside pool: 10
2	**Optimal Core:** Base Topics 2 & 3 Per Capita EValue/ Social Value: **2.86** *m* pool: 1 Prob (choose outside pool): 0 Number of Topics in choice set outside pool: 0	**Optimal Core:** Base Topics 1,2,3 & 4 and non-base Topic 11 Per Capita EValue/ Social Value: 3**.16** *m* pool: 1 Prob (choose outside pool): .39 Number of Topics in choice set outside pool: 10
Infinite – all learning sets feasible.	**Optimal Core:** Base Topics 1 & 2 and non-base Topic 9 Per Capita EValue/ Social Value: **2.95** *m* pool: 2 Prob (choose outside pool): .16 Number of Topics in choice set outside pool: 2	**Optimal Core:** Base Topics 1,2,3 & 4 and non-base Topic 11 Per Capita EValue/ Social Value: 3**.16** *m* pool: 1 Prob (choose outside pool): .39 Number of Topics in choice set outside pool: 10