Towards a Comparison of Evolutionary Creativity in Biological and Cultural Evolution

Andre Skusa and Mark A. Bedau*

Reed College, 3203 SE Woodstock Blvd.,Portland OR 97202
andre.skusa@reed.edu, mab@reed.edu

*To whom correspondence should be addressed.

Abstract

Bedau and Packard have defined evolutionary activity statistics that illuminate the adaptive evolutionary creativity of biological evolution. The statistics enable us to visualize adaptive evolutionary dynamics and to measure the intensity and extent of adaptive evolution, and can be used to define qualitatively different kinds of evolving systems. Here we describe how to apply evolutionary activity statistics to systems undergoing *cultural* rather than biological evolution, and we report preliminary results of implementing this method in technological evolution as reflected in patent record data. Measuring evolutionary activity in patent records provides a clear picture of the major adaptive phenomena at work in the evolution of technology. It also enables the quantitative and empirical comparison of the adaptive evolutionary dynamics of biological and cultural evolution.

Evolution of life and culture

A key question about evolving systems is to explain the source of their adaptive creativity. This question has broad applicability, concerning both artificial and natural evolving systems, and it applies to systems exhibiting either biological or cultural evolution. In a series of papers Bedau and Packard have shown how evolutionary activity statistics can be used to visualize and measure the creation of adaptations in many evolutionary systems (Bedau & Packard 1992; Bedau *et al.* 1997; Bedau, Snyder, & Packard 1998; Bedau & Brown 1999; Rechtsteiner & Bedau 1999) . These statistics are quite general and apply to data generated by both artificial and natural systems, and they apply at different levels of analysis. The study of evolutionary activity in natural and artificial biological evolution has yielded an intriguing picture of qualitatively different kinds of evolving systems (Bedau, Snyder, & Packard 1998). The biosphere as reflected in the fossil record shows an especially interesting and explosive kind of evolutionary creativity (Bedau *et al.* 1997; Bedau, Snyder, & Packard 1998), and it has been conjectured that the same kind of explosive adaptive creativity would be seen in certain kinds of cultural evolution (Bedau, Snyder, & Packard 1998). This paper takes the first step toward assessing that conjecture. We show how to apply evolutionary activity statistics to cultural evolution as reflected in patent records. This is a pilot project applied to patent data covering the past five and a half years. Our aim is to show how to create an empirical picture of the adaptive evolutionary dynamics in the evolution of patented inventions. Such pictures will enable us to compare the dynamics of patented technology with those exhibited in biological evolution. It is especially interesting whether this kind of cultural evolution is qualitatively like that exhibited in the fossil record.

The ultimate aims of this work is to illuminate the relationship between life and culture. One relationship is trivial: cultural phenomena involve the behavior or psychology of living creatures, especially humans. A much more interesting and controversial question is whether living and cultural phenomena and the mechanisms shaping them are in some way essentially the same. This is closely connected to question thirteen in the list of grand challenges in artificial life produced at Artificial Life VII (Bedau *et al.* 2000), and it is our focus here.[1] We approach this issue by showing how to compare the statistical signature of biological and cultural evolution in empirical evolutionary activity data.

There is plenty of previous work on cultural evolution and on patents, but none quite like ours. For many years cultural change has been treated as a process of the diffusion of ideas (Rogers 1995), and the scientometrics community has been investigating scientific and technological change by analysis of patent records and the like for decades (Pavitt 1985; Garfield & Welljams-Dorof 1992; Narin 1994; Albert 1998). But these approaches understand "evolution" in the sense of physics rather than biology, that is, simply as any change in time rather than just change resulting from differential imperfect replication and selection.

Sociobiology (Wilson 1978; Lumsden & Wilson 1981)

[1]The work reported in this paper is related to at least two further grand challenges in artificial life: question six about the nature of open-ended evolution and question eleven about the emergence of intelligence and mind in artificial living systems.

and its contemporary sibling evolutionary psychology (Barkow, Cosmides, & Tooby 1992) explore one kind of connection between biological and cultural evolution, specifically, the extent to which certain psychological and cultural phenomena (e.g., homosexuality and altruism) can be explained by appeal to the operation of biological evolution itself. This reduction of social science to biology is contrasted with the approach to culture illustrated by memetics (Lynch 1996; Blackmore 1999; Aunger 2000), which considers the evolution of cultural phenomena in its own right, independent from and even competing with biological evolution. The two classic quantitative treatments of cultural evolution (Cavalli-Sforza & Feldman 1981; Boyd & Richerson 1985) tend toward different answers to the question whether cultural evolution is ultimately explainable in terms of biological evolution, with Cavalli-Sforza and Feldman leaning toward explanatory dependence and Boyd and Richerson leaning toward a limited autonomy for culture. Our approach is neutral on this issue. We study cultural evolution as an evolutionary process in its own right, ignoring whether and how it might depend on biological evolution. Our goal is to provide an empirical and quantitative picture of the evolution of culture, one which allows us to compare its evolutionary dynamics with those of biological evolution. Both reductionists and antireductionists could profit from objective empirical measurement of cultural dynamics.

Although cultural evolution is not a common subject in artificial life, work on topics like the evolution of language and the evolution of economic phenomena does appear from time to time in artificial life journals and conferences; see, e.g., two recent reviews (Kirby 2002; Tesfatsian 2002). However, cultural evolution is typically construed as tangential to artificial life's central questions, connected mainly because of a similar methodology deploying evolutionary and agent-based models. Artificial life's marginalization of cultural evolution probably largely reflects the uncertainty about how life and culture are related. We aim to reduce this uncertainty and place the study of cultural evolution at the heart of artificial life.

Evolutionary activity in life

Bedau and Packard developed evolutionary activity statistics in order to visualize and measure the dynamics of the process by which evolution creates significant adaptations. Detailed definitions and illustrations of the evolutionary activity method have been presented in earlier publications (Bedau & Packard 1992; Bedau *et al.* 1997; Bedau, Snyder, & Packard 1998; Bedau & Brown 1999; Rechtsteiner & Bedau 1999). This section reviews the method and summarizes results relevant to the present paper. Precise definitions of evolutionary activity statistics as we apply them to patented innovations appear in a later section.

Adaptations are components of an evolving system with properties that allow them to preferentially survive, replicate, and in general play an active role in the system's behavior. Evolutionary activity statistics are derived from bookkeeping about things like the survival, replication, and activity or use of system components. The details of this bookkeeping can vary from case to case, depending on what properties of the system in question reflect the adaptive success of components of interest. For example, if the issue is the adaptive success of genotypes in a system in which a genotype's concentration reflects its adaptive success, then the bookkeeping simply integrates a genotype's concentration over its lifetime in the system (Bedau & Brown 1999). If the issue is the adaptive success of alleles in a system in which the continual expression or use of an allele reflects its adaptive success, then the bookkeeping simply sums an allele's use over its lifetime in the lineage (Bedau & Packard 1992).

Evolutionary activity data can be processed to yield various statistics that summarize the dynamics of adaptive evolution. Two statistics are especially relevant here. One is the *intensity* of evolutionary activity, which measures the rate at which new activity is being created by the system. The other is the *extent* of evolutionary activity, which measures the total amount of activity present in the system at a given time. These notions have been defined in a variety of ways in previous publications, but the underlying inspiration behind them has remained constant.

It is sometimes ambiguous whether activity data provide evidence of adaptations, because non-adaptive (or even maladaptive) components can generate a certain amount of activity before natural selection weeds them from the system. One can filter such non-adaptive "noise" from the activity data with a *neutral* model of the system under investigation (the "target" system). A neutral model is designed to be like the target system in all relevant respects except that by construction there is no natural selection; instead, all selection is random. Thus the neutral model is a no-adaptation null hypothesis that can be used to normalize the activity data observed in the target system. The *excess* activity that remains after normalization with a neutral model can be confidently interpreted as a reflection of significant evolutionary adaptations. With appropriate scaling of the data, neutral normalization allows quantitative comparison of evolutionary activity statistics across different evolving systems (Rechtsteiner & Bedau 1999).

The experience of measuring evolutionary activity in various artificial and natural evolving systems has supported two conclusions relevant here. The first is that evolutionary activity statistics, especially when normalized with a neutral model, do highlight the significant

Class	Description	Evolutionary Activity Signature excess intensity	excess extent	Paradigm Examples
1	no activity	zero	zero	neutral model
2	uncreative activity	zero	unbounded	stable ecosystem
3	bounded activity	positive	bounded	continual adaptive succession
4	unbounded activity	positive	unbounded	Phanerozoic biosphere

Table 1: Empirical classification of evolutionary dynamics, their statistical signatures, and paradigm examples of systems in each class. This classification modifies and augments the classification presented at Artificial Life VI (Bedau, Snyder, & Packard 1998).

adaptive innovations created in the process of evolution, and their broad applicability enables adaptive evolutionary dynamics in different systems to be compared. Second, comparing data from a variety of different systems suggests that evolutionary activity statistics can be used to partition evolutionary dynamics into four qualitatively different classes. Table 1 summarizes these four classes and their statistical signatures. Class 1 consists of systems in which evolution creates no adaptations at all (e.g., all neutral models, systems in which the mutation rate is too high, and systems in which the selection pressure is too low). Systems in which evolution has created adaptations but in which no new adaptations are being created fall into class 2 (e.g., stable ecosystems). Class 3 consists of systems that continually create new adaptations but are bounded in the amount of adaptive structure they contain (e.g., if new adaptations always supplant old adaptations). If new adaptations are continually created and the total amount of adaptive structure continues to grow, then the system falls into class 4. The biosphere as reflected in the fossil record exhibits class 4 dynamics. It would be interesting to be able to ask where cultural evolution falls in this classification.

Cultural evolution in patents

Those who seek to measure cultural evolution face formidable challenges in finding appropriate empirical data. A threshold issue is forming a clear and operational idea of the units of evolution. It is especially difficult to distinguish new innovations from copies of old innovations when the subject is ideas or other mental aspects of culture. Another difficulty is finding a complete time series of rich assays of some aspect of culture in which genealogical relationships can be precisely ascertained.

One can finesse these difficulties if one studies the evolution of patented inventions. Although the evolution of inventions involves the diffusion and selection of ideas, one can operationally identify individual inventions with individual patents. To be patentable an invention must meet three criteria: novelty, usefulness, and non-obviousness. So patented inventions are certified to be new and functional. A patent's novelty is documented

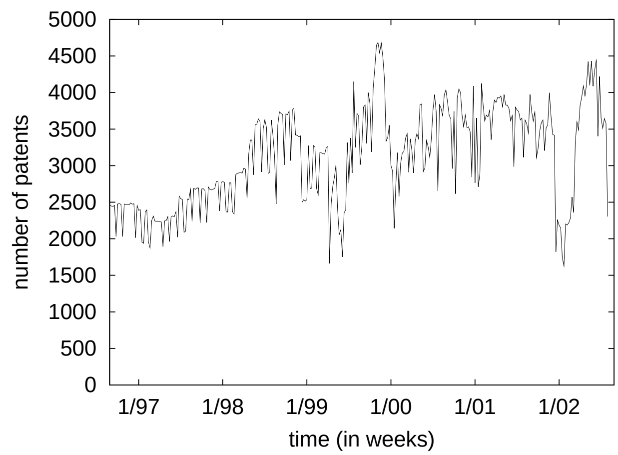

Figure 1: The number of patents issued each week. This number seems to be rising and getting more volatile.

by citing the previous patents (and sometimes published papers) that involve related ideas; these are called the patent's "prior art." The citations should identify all the important prior art on which the invention improves and, as Perko and Narin emphasize, patent examiners are charged with ensuring that no relevant prior art is missed (Perko & Narin 1997):

> These references are chosen and screened by patent examiners, who are "not called upon to cite all references that are available, but only the best." (*Manual of Patent Examining Procedures*, Section 904.02)

Such citations to prior art provide a record of a patent's evolutionary genealogy; in the aggregate the references imply a precise and complete phylogeny of all patents.[2]

We studied the evolution of culture in the complete set of utility patents granted by the United States Patent

[2]Caveat: patent applications must cite important prior art whether or not it plays a *causal* role in an invention's origin. So the evolutionary lineages derived from patent citation data can connect inventions with independent etiologies. This is a limitation of our method, for we would like to interpret evolutionary activity etiologically. We expect that most prior art cited in patents does play some causal role in the generation of those patents, but we have no strong evidence for this in hand.

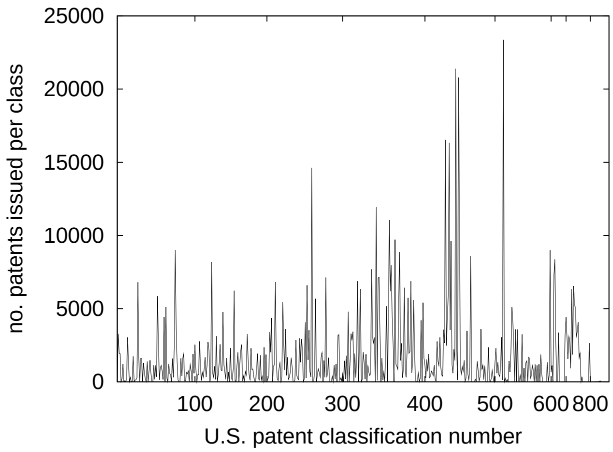

Figure 2: Number of patents issued in each class during 9/96–7/02. The number of patents in various patent classes can differ by an order of magnitude.

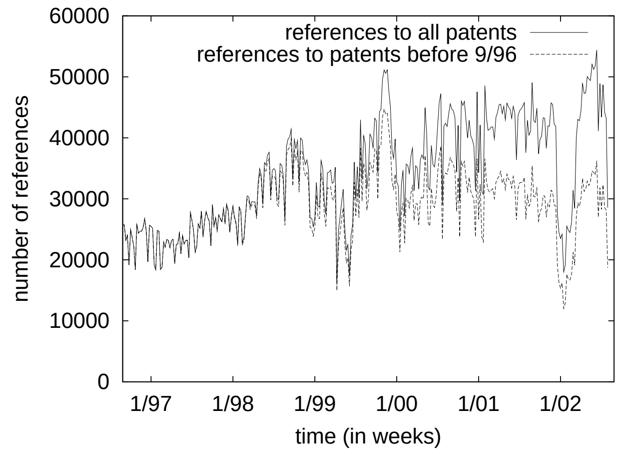

Figure 3: Time series of the number of references in the patents issued each week. The dashed line shows the references to patents issued prior to the time period shown. The gap between the solid and dashed lines shows the growing number of references to patents issued during 9/96–7/02. Note that two to three year time lag before patents issued after 9/96 start to be cited (i.e., before there is an appreciable gap between the two lines).

and Trademark Office (USPTO) during the period 9/96–7/02. (The complete record of all patent applications granted over this period is available for download at http://www.uspto.gov/web/menu/patdata.html.) Our data set contains 868,535 patents. Figure 1 shows the number of patents issued each week during this period. The USPTO classifies patents according to the category of the invention, and US patents are currently placed in one of 419 main classifications. The USPTO occasionally revises its classification scheme, dropping outmoded classes, splitting overburdened classes, and adding new classes, so the class numbering sequence has gaps. Although there are only 419 classes, some classes are now numbered over 800. Figure 2 shows the wide variability in the number of patents issued in different classes in our data.

The analogies and disanalogies between biological and cultural evolution are a matter of controversy (Hull 1988), but it is relatively straightforward to extract evolutionary activity data from patent records. The units of evolution with which we are concerned (at least in the first instance) are individual patents; these are analogous to genes (or, as memeticists might suggest, "memes"). A gene could vanish forever from an evolutionary system. By contrast, a patented invention never goes fully extinct because the invention exists forever in the patent records. We consider that a patent "reproduces" when it leads to the production of *other* patents; that is, in contrast with most biological evolution, patent reproduction necessarily involves innovation—the creation of new inventions.

Typically two or more years elapse between when a patent application is filed and the USPTO issues a patent. Thus our method of using citations to reflect when a patent reproduces introduces a time lag.[3] Figure 3 shows the total number of references contained in the patents issued each week in our data. (The shapes of Figures 1 and 3 are similar because the average number of references per patent is roughly constant.) Since a patent can cite only patents that have been issued earlier, the two or more year delay while the USPTO examines patent applications means that patents issued at a given time are typically never cited until they are two or more years old. This is evident in Figure 3 as the two to three year lag before patents in our data (i.e., those issued after 9/96) start to be cited.

Figure 4 shows that the average number of citations to patents varies from class to class. Especially successful or valuable patents tend to be those that are especially heavily cited. A raft of work in scientometrics has repeatedly confirmed that number of citations is a good reflection of the technological significance and economic value of a patented invention (Pavitt 1985; Albert 1991; Narin 1994; Perko & Narin 1997; Albert 1998). This parallels the extensive evidence validating the usefulness of science citation data for measuring the importance of scientific publications (Garfield & Welljams-Dorof 1992). Once a patent has received more than ten citations, the economic value reflected by each additional citation has been estimated to be more than one million US dollars (Harhoff *et al.* 1999). For these reasons our bookkeeping of an individual patent's evolutionary activity

[3] We could remove this time lag by dating a citation back to when a patent is filed rather than issued.

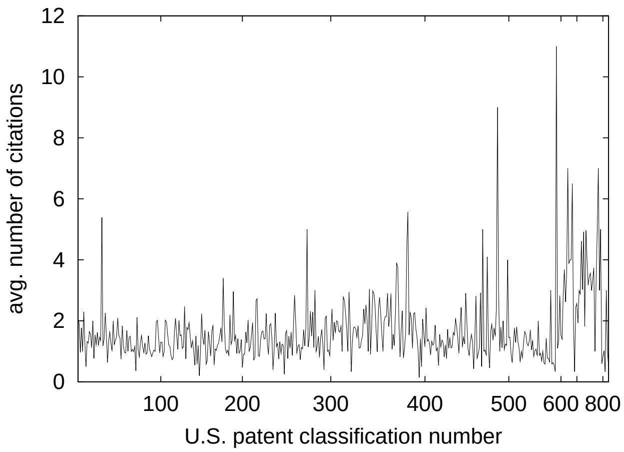

Figure 4: Average number of citations per patent in each class during 9/96–7/02. Note the wide variability between patent classes.

is based on summing the citations the patent has received. ¿From this perspective, the adaptive success of a patented innovation is measured by the extent to which it spawns subsequent patented innovations.

Definitions of patent activity

Applying evolutionary activity statistics in a given context requires settling two issues: (1) choosing which components of the evolving system will have their activity measured, and (2) choosing how to increment their activity. Our response to (1) is to examine the activity of individual patents; this allows us subsequently to examine the activity of patent classes by aggregating the activity of the patents in those classes. Our response to (2) is to increment a patent's activity at a given time by the number of citations it receives from patents issued at that time. More specifically, we let Ξ_i^t be the set of patents issued at t that cite patent i, i.e., $\Xi_i^t = \{j : j \in \Upsilon_t \;\wedge\; j \text{ cites } i\}$, where Υ_t is the set of patents issued at t. We let $\Delta_i(t)$ be the amount that i's activity increases at t and we define this as the number of new citations i receives at t, i.e., $\Delta_i(t) = \#(\Xi_i^t)$. Then, we let the counter $a_i(t)$ reflect the activity of patent i at time t and define this as the number of citations it has received up to t:

$$a_i(t) = \sum_{u=t_i}^{t} \Delta_i(u), \tag{1}$$

where t_i is the time when patent i is issued. The activity $a_c(t)$ of a patent class c at time t is then defined as the sum of the activity of the patents in it, i.e., $a_c(t) = \sum_{i \in c} a_i(t)$.

Patent neutral shadows

The mere fact that a patent has received a few citations does not prove that the invention significantly shapes the evolution of subsequent inventions. A patent might be cited by one or two subsequent patents even if patents to cite were chosen entirely at random. As with evolutionary activity measurements in other contexts, we can evaluate a patent's adaptive significance by comparing its activity with the activity observed in a neutral model of patent evolution.

Our patent neutral model mirrors a few key aspects of the real patent data. In both the same number of patents are issued each week, and they exhibit the same distribution into the various patent classes. Patent citations refer to the same number of pre-9/96 and post-9/96 patents, so real and neutral model patents both exhibit time lags before collecting many citations. Furthermore, the references to post-9/96 patents fall into the various patent classes according to the same distribution. The key distinguishing feature of the neutral model is that the patents to be cited are always chosen randomly.

More precisely, here is how the neutral model works. When a real patent is issued in a given class in a given week, a shadow patent is issued in the same class in the same week. Furthermore, when the real patent cites a post-9/96 patent, its corresponding shadow patent cites a previously issued shadow patent. Since patents in different patent classes have different expected likelihoods of being cited (recall Figure 4), we have the shadow patents mirror this bias. This gives the shadow neutral model and the real patents the same number of patents per class and the same number of citations per class. The crucial difference between the real patent system and its neutral shadow concerns the process for randomly choosing which patents to cite. Specifically, when it is time for a shadow patent to cite a previously issued shadow patent, we first randomly choose the patent class in which we will choose the patent to cite, but our random choice of patent class is *not* made with equal probability. Rather, we weight the random choice of patent class by the frequency with which real patents cite patents in different classes. This guarantees that real and shadow patent classes will have the same distribution of citations. After a patent class is randomly chosen, we randomly choose with equal probability which prior patent in that class to cite.[4]

[4] An anonymous reviewer called our attention to a disanalogy in the time lags exhibited by our neutral model and the real patents. Although there are some exceptions, a real patent usually cites only patents that were issued before the application patent was filed. The time lag before patents start to accrue citations largely reflects the gap between filing and issuing dates. But a shadow patent issued in one week is just as likely to cite a shadow patent issued in the previous week as any other patent in the same class. This disanalogy could be corrected by having the neutral model

Excess activity, its intensity and extent

Once we have a neutral model for the patents, normalizing by the neutral model yields a measure of the excess activity in the patents—that is, the amount of activity that is unambiguously attributable to the patent's adaptive significance. To achieve this effect, we define the excess activity $a_{i\in c}^{\text{excess}}$ of an individual patent i in class c at time t as the amount by which i's activity exceeds the maximum activity of *any* shadow patent in class c. More specifically, we let $a_{i\in c}^{\text{shadow}}$ be the activity of shadow patent i in class c after the neutral shadow has run its course (i.e., after it has shadowed the real patents for the entire period over which the patents are under investigation). Then $\Psi_c = \{a_{i\in c}^{\text{shadow}} : i \in c\}$ is the set of activity values of all the shadow patents in class c. Finally, we define the excess activity of patent i in class c at t as follows:

$$a_{i\in c}^{\text{excess}}(t) = \begin{cases} a_i(t) - \max(\Psi_c) & \text{if } a_i(t) > \max(\Psi_c) \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

So, patents in a given class with positive excess activity have acquired more citations than those acquired by any shadow patent in the same class. The excess activity of a patent class c at time t is defined as the sum of the excess activity at t of all the patents in class c: $a_c^{\text{excess}}(t) = \sum_{i\in c} a_{i\in c}^{\text{excess}}(t)$.

Figure 5 shows the dramatic difference between the activity accrued by shadow patents and real patents with excess activity. Another picture of this difference comes from comparing the distribution of the patents' final activity values (Figure 6). Overall, the citation levels of shadow patents are much lower than the citation levels of patents with excess activity.

Only 17,783 of the patents issued during 9/96–7/02 have positive excess activity; this represents about 2% of the patents issued during this period. Thus our neutral model is a very liberal no-adaptation null hypothesis. It might well screen out some significant innovations, i.e., it might yield some false negatives, but it should allow no false positives. This liberality is no problem in the present context for our primary concern is to focus on those patents that unequivocally play a significant role in the evolution of inventions.

In order to classify the kind of evolutionary dynamics evident in the patent record, we must define two global activity statistics—excess intensity and excess extent of activity—in a way that can be usefully applied to the evolution of patented inventions. Excess intensity of activity is intended to reflect the rate at which new significant innovations are being created. We define the excess intensity of evolutionary activity of the patented inventions at time t as the number of patents issued at t that have positive excess activity at the time when excess intensity is measured. More precisely, we define the excess intensity of activity of the patents at time t, as measured at time τ, as:

$$\text{intensity}_{\tau}^{\text{excess}}(t) = \#\{i : i \in \Upsilon_t \ \wedge \ a_{i\in c}^{\text{excess}}(\tau) > 0\}, \quad (3)$$

where c is ranges over all classes and where Υ_t is the set of patents issued at t, as before. That is, excess intensity at t as measuered at τ is the number of patents issued at t that will have positive excess activity at τ. When we report excess intensity below, our observations are made at the end of the data we examine; i.e., we set $\tau = 7/02$.

The excess intensity at t is not absolute; rather, it depends on the time τ when it is measured. Thus, the value of the excess intensity at t can increase as later and later measurements are made, i.e., $\tau - t$ increases. No moment in the evolution of patented technology can be seen to have positive excess activity contemporaneously, i.e., if $t \approx \tau$. Positive excess intensity can be observed only years later, after the patents have accumulated a sufficient number of citations.[5] A true reckoning of the eventual excess intensity can be settled only when those patents are no longer cited—a time measured in generations.

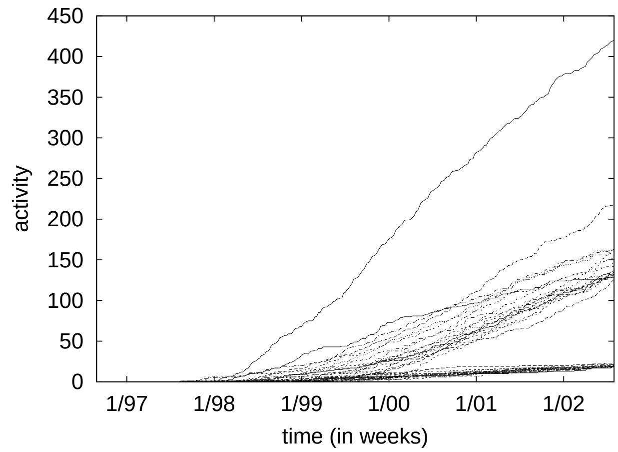

Figure 5: Comparison of the activity of real patents and shadow patents, showing the twenty most heavily cited real patents and the twenty most heavily cited shadow patents. The activity waves of the real patents all rise above 100 while those of the shadow patents remain below 25. Note that the activity accrued by significant real patents can vastly exceed that accrued by any shadow patent.

mirror the distribution of the dates of cited patents. We suspect that this correction would not substantially alter our results.

[5] It is possible to define excess intensity at t in a way that is not relative to time of observation, perhaps most naturally as the number of patents whose excess activity first become positive at t. This kind of definition has the drawback that its value at one time reflects the patents that were issued years previously. We employ a future-dependent definition to avoid precisely this problem.

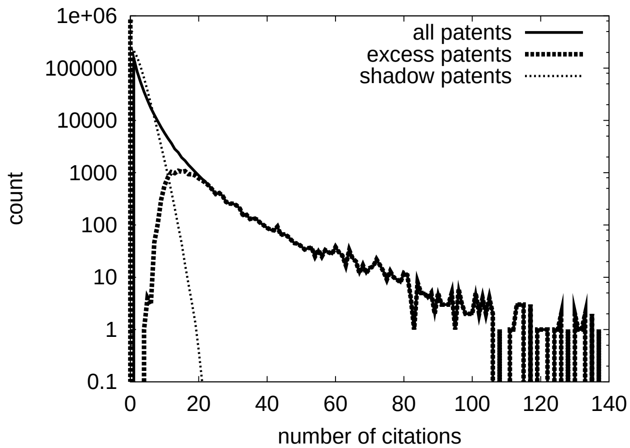

Figure 6: Log-normal plot of the distribution of final activity for all patents at 7/02, shadow patents, and just those patents with positive excess activity. Note that shadow patent activity never approaches the activity of the most heavily cited patents.

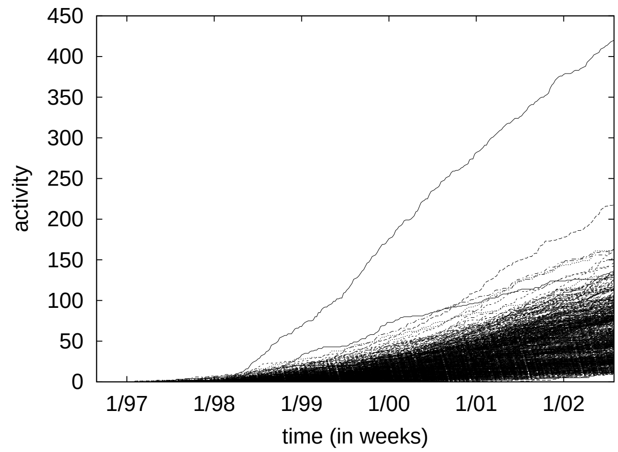

Figure 7: The activity waves of patents with positive excess activity. We graph the top two hundred excess waves and then omit redundant information by sampling the remaining excess activity waves.

Excess extent of evolutionary activity is intended to reflect the mass of excess evolutionary activity at a given time. We define the excess extent of activity of patented inventions as the sum of the activity of those patents that have positive excess activity:

$$\text{extent}^{\text{excess}}(t) = \sum_{i,c} a_{i\in c}^{\text{excess}}(t). \qquad (4)$$

Unlike the definition of excess intensity of activity, the definition of excess extent of activity is not relative to a time of observation. The value of excess intensity at t cannot change if it is measured at different times, and the excess extent of activity of the present moment can be determined at the present. Nevertheless, excess extent is an "historical" statistic, so the value of excess extent at t depends on the amount of data collected before t. Including earlier patents in our data would enable more patents to have a chance at positive excess activity.

Observations of excess activity

Figure 7 shows the activity accrued by patents with positive excess activity. Note that one patent stands far above the rest; it accrues almost twice as many citations as any other patent. This patent covers the technology that allows web browsers to display information such as advertisements while a page is being loaded a link is clicked. The second most heavily cited patent covers the technology that allows cell phones to receive email and faxes, and the third most heavily cited patent allows remote control of the receipt and delivery of wireless and wireline voice and text messages. All of the ten most heavily cited patents fall into the information technology sector (see Table 2).

Figure 8 shows the excess activity of all patent classes. Information technology covers ninety percent of these classes. The most-heavily cited patent falls in class number 709. This class has the third highest excess activity, and it has the highest average excess activity per patent. The ten patent classes with the most excess activity are listed in Table 3. Most of these concern information technology, and the second largest group concerns health. When these patent classes are ranked according to average excess activity per patent, the top three information technology patent classes have a 15–30% lead over all other patent classes.

To place the cultural evolution exhibited by patents into the classification of Table 1 we need to measure the excess intensity and extent of evolutionary activity in the patent data. Figure 9 shows the excess intensity of activity. The dramatic fall to zero over this graph is an artifact of the finite size of the data. We saw above that it takes two or more years for the USPTO to process a patent application. This means that it takes two or more years before a patent can receive many citations. Hence a patent's status of receiving excess citations can become recognized only two or more years after it has been issued. So, we would expect to continually learn that additional patents have excess activity—and hence learn that excess intensity is higher than we previous thought—two or more years after patents have been issued. (Recall that excess intensity depends on the time τ of observation.) So as we collect additional patent record data for the months and years following 7/02, the calculated excess intensity over the time period shown in Figure 9 will continue to increase. This increase will not continue forever, though, since sufficiently old patents are no longer cited. Thus, as long as some patents achieve positive

Rank	Excess Activity	Patent Number	Patent Title	Class Numb.
1	420	5572643	Web browser with dynamic display of information objects during linking	709
2	218	5608786	Unified messaging system and method	379
3	163	5742905	Personal communications internetworking	455
4	162	5708780	Internet server access control and monitoring systems	709
4	162	5557518	Trusted agents for open electronic commerce	705
6	159	5632021	Computer system with cascaded peripheral component interconnect (PCI) buses	710
7	152	5774660	World-wide-web server with delayed resource-binding for resource-based load balancing on a distributed resource multi-node network	709
8	147	5655081	System for monitoring and managing computer resources and applications across a distributed computing environment using an intelligent autonomous agent architecture	709
9	142	5623601	Apparatus and method for providing a secure gateway for communication and data exchanges between networks	713
10	137	5610910	Access to telecommunications networks in multi-service environment	370

Table 2: The ten patents with the highest excess activity. Note the total dominance by information technology.

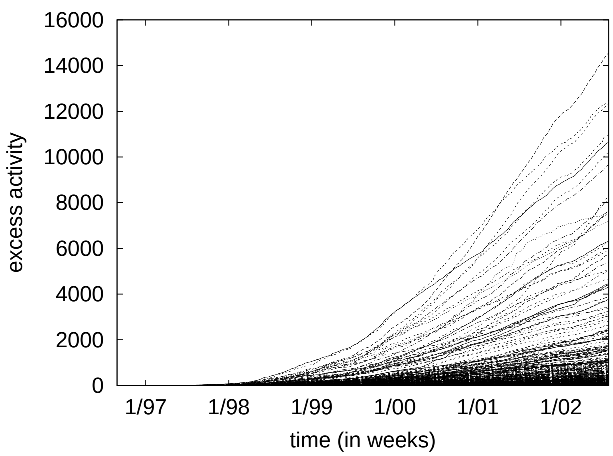

Figure 8: The excess activity of all patent classes. Note that very many patent classes show significant positive excess activity.

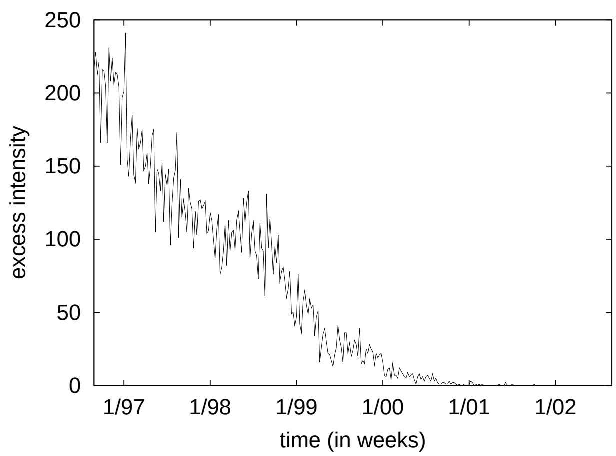

Figure 9: Time series of the excess intensity of activity of patents. The drop in excess intensity to zero is a finite-size effect. See text.

excess activity, the excess intensity of the patents will be positive over data with a sufficiently long time course (on the order of decades). So, even though our excess activity data suffers a finite size effect, we can still tell that the cultural evolution of patented inventions shows the positive excess intensity signature of classes 3 and 4.

So whether the cultural evolution of patented inventions falls into class 3 or 4 turns on whether its excess extent of activity is bounded or unbounded. Figure 10 shows excess extent over 9/96–7/02. The first thing to notice about this graph is that it also shows a finite size effect. Our data starts with the patents issued in 9/96, and no patent issued at that time will have positive excess activity in 9/96. It's excess activity can become positive only some months or (more likely) years in the future, when it has accumulated more citations than shadow patents. We could remove this finite size effect at 9/96 by collecting information about patents issued earlier. If this were done, the excess extent of activity in the the first fifth of the period shown in Figure 10 would start to rise. But this will not *remove* the finite-size effect; it will just shift it back in time. There will be an initial period of zero excess extent of activity no matter when we start collecting data.

What is more intriguing—and ambiguous—is the pattern in excess extent toward the *end* of the period shown in Figure 10. Excess activity is clearly positive; the question is whether it is bounded, that is, whether it will continue to show an overall rising trend into the future or whether it will top out. The excess extent of activ-

Rank	Total Excess Activity	Average Excess Activity	Class Number	Class Title
1	14553	30.38	438	Semiconductor device manufacturing: process
2	12483	35.36	370	Multiplex communications
3	12343	40.74	709	Electrical computers and digital processing systems: multiple computer or process coordinating
4	10973	29.11	606	Surgery
5	10640	35.12	707	Data processing: database and file management, data structures, or document processing
6	10198	26.91	257	Active solid-state devices (e.g., transistors, solid-state diodes)
7	9649	27.49	345	Computer graphics processing, operator interface processing, and selective visual display systems
8	8258	19.52	435	Chemistry: molecular biology and microbiology
9	8105	30.13	600	Surgery
10	7646	16.48	514	Drug, bio-affecting and body treating compositions

Table 3: The ten patent classes that collect the most excess activity during 9/96–7/02, with the total excess activity class and the average excess activity per patent. Note that most of the most activity classes fall into the information technology sector, including the four classes with the highest average excess activity. The USPTO descriptions of all patent classes are available at http://www.uspto.gov/web/patents/classification.

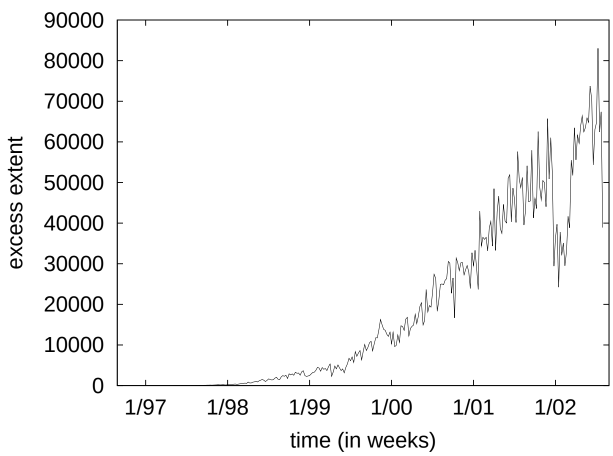

Figure 10: Time series of the excess extent of activity of patents. The initial period of zero or very low excess extent is a finite-size effect. See text.

ity shown in Figure 10 is consistent with the hypothesis of unbounded growth, but it is also consistent with a finite bound. Extrapolating long-term trends is always somewhat uncertain, and the volatility in the number of patents issued during the first half of 2002 (the tail end of our data) makes it especially risky. However, analyzing patent records for the preceding decade or two should go a long way toward revealing whether the extent of excess activity for the patents is unbounded. So, although our pilot data cannot resolve whether the evolutionary dynamics of patented inventions is in class 3 or 4, the question is now precisely formulated and quite amenable to empirical investigation.

Conclusions

Our pilot project has proved the feasibility of visualizing and quantitatively assessing the adaptive evolutionary dynamics exhibited in cultural evolution. We have applied the method to technological evolution as reflected in patent record data, but it can be applied at a variety of levels of analysis in a variety of cultural systems. The main problem is getting enough appropriate data. This problem should be easily solvable for cultural evolution reflected in such things as scientific citation data, financial data about economic markets, databases of newspaper articles, and data about the evolution of the world wide web.

Our preliminary analysis of technological evolution underscores the vast importance of information technology, and especially the Internet, over the past five years. This conclusion is not news, of course; it just corroborates what we already know. But it does confirm the aptness and probity of evolutionary activity analysis of cultural evolution, and it opens the door to myriad subsequent empirical investigations of cultural evolution.

The most intriguing issue highlighted by our pilot results is the question whether the technological evolution exhibits class 3 or 4 dynamics. This question can be answered once we have data with a significantly long time horizon. Whatever the answer, cultural evolution will be shown to have a precise quantitative similarity with one or another kind of biological evolution. Resolving this issue will shed new light on the nature and scope of class 3 or 4 evolutionary dynamics, and it might help us to formulate deeper and more refined classifications of evolutionary creativity. So, evolutionary activity analysis of patented technology opens a constructive path toward understanding whether life and culture are two manifestations of one fundamental kind of process.

Acknowledgments

Thanks to Norman Packard for helpful discussion and extensive collaboration on evolutionary activity statistics. For constructive suggestions, thanks to Steen Rasmussen, Mike Raven, the Artificial Life VIII reviewers, and the audiences at Lake Arrowhead, California (May 2002), and at the Fraunhofer Gesellschaft in Sankt Augustin, Germany (July 2002), where MAB presented these results. Thanks to Howard Gutowitz for the suggestion at Artificial Life VI of applying evolutionary activity statistics to patent record data. AS thanks the Studienstiftung des deutschen Volkes (German National Merit Foundation), Bonn, Germany, for financial support.

References

Albert, M. B. 1998. *The new innovators: Global patenting trends in five sectors.* Washington D.C.: U.S. Department of Commerce.

Albert, M.; Avery, D; Narin, F.; and McAllister, P. 1991. Direct validation of citation counts as indicators of industrially important patents. *Research Policy* 20: 251–259.

Aunger, R., ed. 2000. *Darwinizing culture: The status of memetics as a science.* New York: Oxford University Press.

Barkow, J. H.; Cosmides, L.; and Tooby, J., eds. 1992. *The adapted mind: Evolutionary psychology and the generation of culture.* New York: Oxford University Press.

Bedau, M. A.; and Brown, C. T. 1999. Visualizing evolutionary activity of genotypes. *Artificial Life* 5: 17-35.

Bedau, M. A.; McCaskill, J. S.; Packard, N. H.; Rasmussen, S.; Adami, C.; Green, D. G.; Ikegami, T.; Kaneko, K.; Ray, T. S. 2000. Open problems in artificial life. *Artificial Life* 6: 363-376.

Bedau, M. A.; Snyder, E.; Brown, C. T.; and Packard, N. H. 1997. A comparison of evolutionary activity in artificial evolving systems and in the biosphere. In *Proceedings of the Fourth European Conference on Artificial Life*, P. Husbands and I. Harvey, eds., pp. 125-134. Cambridge: MIT Press.

Bedau, M. A.; Snyder, E.; and Packard, N. H. 1998. A classification of long-term evolutionary dynamics. In *Artificial Life VI*, C. Adami, R. Belew, H. Kitano, and C. Taylor, eds., pp. 228-237. Cambridge: MIT Press.

Bedau, M. A.; and Packard, N. H. 1992. Measurement of evolutionary activity, teleology, and life. In *Artificial Life II*, C. Langton, C. Taylor, D. Farmer, S. Rasmussen, eds., pp. 431-461. Redwood City, CA: Addison-Wesley.

Blackmore, S. 1999. *The meme machine.* New York: Oxford University Press.

Boyd, R; and Richerson, P. J. 1985. *Culture and the evolutionary process.* Chicago: University of Chicago Press.

Cavalli-Sforza, L. L.; and Feldman, M. W. 1981. *Cultural transmission and evolution: A quantitative approach.* Princeton: Princeton University Press.

Garfield, E.; and Welljams-Dorof, A. 1992. Citation data: Their use as quantitative indicators for science and technology evaluation and policy making. *Science and Public Policy* 19: 321–327.

Harhoff, D.; Narin, F.; Scherer, F. M.; and Vopel, K. 1999. Citation frequency and the value of patented innovation. *Research Policy* 81: 511–515.

Hull, D. L. 1988. *Science as process: An evolutionary account of the social and conceptual development of science.* Chicago, London: The University of Chicago Press.

Kirby, S. 2002. The evolution of language. *Artificial Life* 8: 185–215.

Lumsden, C. J.; and Wilson, E. O. 1981. *Genes, mind, and culture.* Cambridge: Harvard University Press.

Lynch, A. 1996. *Thought contagion: How belief spreads through society.* New York: Basic Books.

Narin, F. 1994. Patent bibliometrics. *Scientometrics* 30: 147–155.

Pavitt, K. 1985. Patent statistics as indicators of innovative activities: Possibilities and problems. *Scientometrics* 7: 77–99.

Perko, J. S.; and Narin, F. 1997. The transfer of public science to patented technology: A case study in agricultural science. *Journal of Technology Transfer* 22: 65–72.

Rechtsteiner, A.; and Bedau, M. A. 1999. A generic model for quantitative comparison of genotypic evolutionary activity. In *Advances in Artificial Life*, D. Floreano, J.-D. Nicoud, F. Mondada, eds., pp. 109-118. Heidelberg: Springer-Verlag.

Rogers, E. M. 1995. *Diffusion of innovations.* Fourth edition. New York: Free Press.

Tesfatsian, L. 2002. Agent-based computational economics: growing economies from the bottom up. *Artificial Life* 8: 55-82.

Wilson, E. O. 1978. *On human nature.* Cambridge: Harvard University Press.