Eigensystems and Attractors: Stable Points in Self-Organizing Information Spaces

Miles Efron
School of Information and Library Science
UNC, Chapel Hill
efrom@ils.unc.edu

14th January 2002

1 Introduction

Recent work in the field of complexity studies has shed light on a broad class of systems that engage in self-organizing behavior. Stuart Kauffman identifies persistent patterns of action in systems as diverse as idealized Boolean networks and multi-cellular organisms. Such systems are "self-organizing" insofar as their elements tend to evolve emergent order despite varied initial conditions and their lack of any guiding agent. Thus randomly configured Boolean networks develop patterns of activiation despite their stochastic initialization and lack of intentional design. Likewise, biochemical systems known as autocatylitic sets tend to stablize into orderly chemical reactions, a stability which Kauffman argues is critical for the emergence of living organisms.

It has been argued that self-organization is evident also in informative systems, such as human language and electronic networks. Balasubrahmanyan and Naranan [1] argue that natural language constitutes a complex system similar to those described by Kauffman. Lending support to this argument, Cancho and Sole have identified the so-called "small world" phenomenon in natural language corpora. Cancho and Sole argue that patterns of word co-occurrence in human discourse define a graph structure characterized by densely clustered nodes with short inter-nodal paths[2, 3]. This type of graph configuration, they argue, implies that human language has evolved into coherence despite its lack of design or purposive direction. Likewise, Bernardo Huberman and his colleagues have identified small-world structure in the adjacency matrices described by the World Wide Web. Huberman cites "patterns [that] turn out to be so ubiquitous that they can be considered as manifestations of underlying laws governing the growth of the Web and the way humans interact with it" [6, p. 17]. According to Huberman, the Web has ordered itself spontaneously, organizing its structure into statistically compelling patterns despite its chaotic development.

Central to the theory of self-organization is the notion of attractors. If a complex system's state space is considered in geometrical terms, attractors correspond to regions of stability in that space. The language of topology is commonly invoked to describe attractor behavior. Given a state space in three dimensions, attractors correspond to watersheds, basins that draw the system out of unstable points (i.e. high ridges) and into regions of stasis (i.e. lakes). Kauffman argues that attractors enable self-organization. Because not all points in state space are equally viable, systems move into basins of stability more quickly and efficiently than would be predicted in systems governed solely by natural selection.

This study describes the role of attractor behavior in the self-organization of informative systems. In particular, we study attractors in the context of information retrieval (IR) systems. Taking a linear algebraic approach, we argue that eigensystems arise naturally in many IR problems, and that they in fact constitute attractors in the "evolution" of a given system. Eigensystems play an important role in several IR applications. However, their theoretical merit remains unarticulated and ambiguous. By understanding the IR task as an evolutionary process that benefits from self-organization, the utility of eigensystems becomes clearer. Proponents of eigen analysis in IR argue that eigensystems manifest the "latent semantic structure" of an informative system. We argue that this manifestation arises from the eigensystems' function as attractors in state space.

2 The Theory of Attractors

Self-organization occurs when systems (of any kind) develop patterned behavior without recourse to external design or intervention. According to Stuart Kauffman, intentionality is not a *sine qua non* for ordered behavior. "Our intuitions about the requirements for order have," he argues, "been wrong for millennia. We do not need careful construction; we do not require crafting. We require only that extremely complex webs of interacting elements are sparsely coupled" [8, p. 84]. Proponents of self-organization argue that natural selection alone cannot account for the explosion in biological diversity that appeared during the Cambrian period, for instance. Nor does Darwinian theory satisfactorily explain the discrete quality of biological taxa [13]. They argue that natural selection has always been aided by systems' tendency to organize themselves [7].

Crucial to the theory of self-organization is the idea of attractors. Kauffman describes attractors in terms of a system's movement through state space. "In the language of dynamical systems, the state cycle is an attractor and the collection of trajectories that flow into it is called the basin of attraction. We can roughly think of an attractor as a lake, and the basin of attraction as the water drainage flowing into that lake"[8, p. 78]. In other words, not all points in state space are equally likely to be occupied by a system. As an example, consider a drop of ink deposited into a glass of water. As time passes, the ink will disperse, creating designs that are random, but also guided by the laws of thermodynamics. Eventually, the law of entropy guarantees that the ink will

$$\mathbf{P} = \begin{array}{cl} \begin{matrix} urb & non \end{matrix} & \\ \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix} & \begin{matrix} urb \\ non \end{matrix} \end{array}$$

Figure 1: Transition Matrix for US Population

disperse more or less evenly through the water. This state of dispersion corresponds to the system's long-term stable point–its strongest attractor. Thus it is highly unlikely that the ink drop will re-form itself into a globe, separate from the water.

Two things are particularly important with regard to attractors. First, they guide the long-term behavior of a system. Second, their form is independent of the initial condition of the system. By definition, attractors guide the long-term behavior of their systems. For instance, Wesson cites as evidence of attractor behavior the "general rule that selective breeding for a trait reaches a plateau in 30 to 50 generations in plants and animals..." [13, p. 179]. This characteristic of attractors is related to their resilience to fluctuations in initial conditions. Kauffman suggests that a system under the sway of an attractor is likely to remain under its sway regardless of externally imposed alterations of its condition [8, p. 78][7, p. 201].

Attractors thus define stable points in the state space of a given system. We will argue that this function is identical to the role of eigenvalues and eigenvectors in systems described as numeric matrices. Like attractors, eigensystems describe long-term behavior of networked elements, regardless of their initial configurations.

3 Eigensystems and Attractors

To motivate our discussion of attractor behavior in information systems, consider a particularly simple example–a Markov chain for predicting demographic shifts in a population.[1]For the sake of simplicity, let all members of a given population live in one of two situations, an urban or non-urban setting. Based on previous research, Williams [14] offers the Markov model shown in Figure 1. Matrix $\mathbf{P}$ describes the probabilities associated with various transitions of the system. Thus there is a 96 percent chance that a person living in a city this year will remain in the city next year. On the other hand, there is a 4 percent chance that a city-dweller will move to suburbia or the country this year. Williams supplements this model with a vector $\mathbf{x}_0$ which contains initial counts (in millions) for urban and non-urban environments. Given this information we may predict the population distribution next year via formula 1.

$$\mathbf{x}_1 = \mathbf{P}\mathbf{x}_0 \tag{1}$$

[1]This example is adapted from [14]

$$\mathbf{x}_0 = \begin{pmatrix} 58 & urb \\ 142 & non \end{pmatrix}$$

Figure 2: Initial State of the System

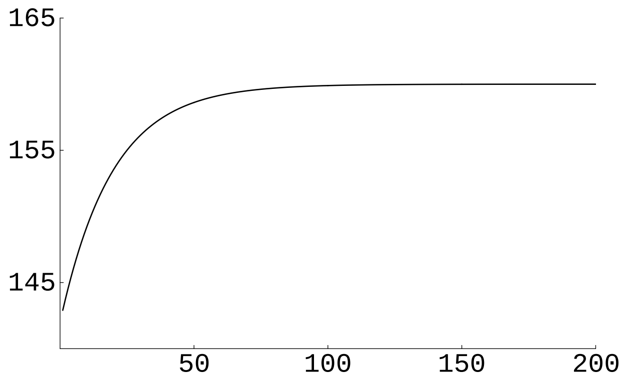

Figure 3: Non-urban Population Convergence

Next we may predict the distribution after two years by multiplying $\mathbf{x}_1$ times $\mathbf{P}$. Continued iterations of this kind allow us to make long-range predictions regarding the population. By repeating this process many times, our long-range predictions eventually stabilize. We reach a point where $\mathbf{x}_{n+1} = \mathbf{x}_n$. For instance, after 200 iterations, the system appears to stabilize at

$$\mathbf{x}_{steady} = \begin{pmatrix} 40 \\ 160 \end{pmatrix}$$

Figure 3 illustrates this process of stabilization. This figure plots the count of the non-urban population as computed at each iteration of the Markov process.

3.1 Eigensystems and Points of Stability

The steady state vector $\mathbf{v}_{steady}$ is directly related to the eigen-system of the input matrix $\mathbf{P}$. To see that this is the case, we briefly review the theory of eigenvalues and eigenvectors, referring the interested reader to more thorough treatments such as [12]. An eigenvalue of a square matrix $\mathbf{A}$ is a scalar λ such that

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

where $\mathbf{x}$ is a non-zero vector. The vector $\mathbf{x}$ is known as an eigenvector of $\mathbf{A}$ corresponding to the eigenvalue λ. We find the eigenvalues of a matrix by beginning with the equation

$$\mathbf{A}\mathbf{x} - \lambda\mathbf{x} = 0$$

which yields

$$\mathbf{Ax} - (\lambda \mathbf{I}_n)\mathbf{x} = 0$$

This system of homogeneous equations only has non-trivial solutions if $\mathbf{Ax} - (\lambda \mathbf{I}_n)\mathbf{x}$ is singular, that is, if $|\mathbf{Ax} - (\lambda \mathbf{I}_n)\mathbf{x}| = 0$. Thus we find the eigenvalues of $\mathbf{A}$ by finding the roots to this equation. We then find the corresponding eigenvectors by substituting each root for λ and solving the resultant set of homogeneous equations.

Above we observed that as we multiplied the vector $\mathbf{x}$ times the matrix $\mathbf{P}$ many times, the series converged on a steady state. That is, we found a point at which $\mathbf{Px} = \mathbf{x}$. Of course this is identical to the eigenvalue problem, with $\lambda = 1$. The Markov model described above is a special type of matrix, known as a regular stochastic matrix. This guarantees that it will have $\lambda = 1$ as its largest eigenvalue. A well-known theory of eigenvalues states that the steady state of regular stochastic matrices corresponds to this eigenvalue. Thus to find our steady state vector analytically, we solve

$$\mathbf{P} - (\mathbf{I}_2)\mathbf{x} = 0$$

$$\begin{pmatrix} 0.96 - 1 & 0.01 \\ 0.04 & 0.99 - 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$$

This system of equations yields a family of solutions of the form

$$r \begin{pmatrix} 1 \\ 4 \end{pmatrix}$$

where r is any scalar. To find the correct solution we obtain

$$r + 4r = 58 + 142$$

$$r = 40$$

Thus the steady state vector $\mathbf{x}$ is

$$\mathbf{x} = \begin{pmatrix} 40 \\ 160 \end{pmatrix}$$

which is precisely what we found by iteration.

Another important result from the theory of eigenvalues states that multiplying the matrix $\mathbf{P}$ times itself repeatedly converges onto a matrix $\mathbf{Q}$, whose columns are identical, each being an eigenvector of $\mathbf{P}$ corresponding to $\lambda = 1$. Thus the long-range model specified by $\mathbf{P}$ ultimately converges on a matrix $\mathbf{Q}$ of the form

$$\mathbf{Q} = \begin{pmatrix} s & s \\ 4s & 4s \end{pmatrix}$$

Because the columns of $\mathbf{Q}$ sum to 1 (this is a stochastic matrix, after all), the solution becomes

$$\mathbf{Q} = \begin{pmatrix} 0.2 & 0.2 \\ 0.8 & 0.8 \end{pmatrix}$$

Thus the long-term probability of living away from the city is 0.8. Particularly important here is the fact that these long-range probabilities are independent of the initial conditions of the system. That is, the one-to-four ratio of country to city-dwellers is a characteristic of the system itself which emerges as the system iterates through its state space. Eigen-analysis reveals structure that is inherent in the system, manifesting the region of stability into which the system will settle if given the opportunity to do so. A matrix's eigensystem thus corresponds to the basins of attraction that propel the system toward stability.

4 Evolving IR Systems

To understand how the theory of attractors applies to information retrieval, it is useful to consider the generation of an information system as an iterative process. Due to the batch processing style common to most indexing approaches this perspective seems foreign. In this section, however, we argue that the methods for parameterizing many IR systems do, in fact, undergo an evolutionary process. The corollary of this argument is that their evolution is directed by the influence of attractors. In the next section, we demonstrate that these attractors are closely related to the eigensystems of the matrices that describe each system.

4.1 Kleinberg's Hubs and Authorities Algorithm

Consider the hubs and authorities algorithm due to Kleinberg [9]. Systems based on hubs and authorities attempt to identify resources that are particularly authoritative on a given topic, assuming that they will prove more relevant than less authoritative sources. The algorithm discovers these authorities by examining the hyperlink structure among documents. Key to Kleinberg's algorithm are two measures. For each document i we record a_i, it's authority score, and h_i, its hub score. a_i and h_i are mutually reinforcing. That is, a document is considered a strong authority if it is linked to by many strong hubs. Likewise, a page is a strong hub if it links to many authoritative pages.

To compute hub and authority scores, Kleinberg begins with a matrix $\mathbf{A}$, an adjacency matrix that describes the graph structure in the subset of the Web under analysis. Thus $a_{ij} = 1$ if document i contains a hyperlink to document j, or 0 otherwise. Assume that $\mathbf{A}$ has been normalized on its columns. The algorithm also uses two vectors $\mathbf{a}$ and $\mathbf{h}$, which contain the authority and hub scores for each page, respectively. We begin by initializing $\mathbf{a}$ and $\mathbf{h}$ with ones. We may thus compute an initial set of authority scores by taking the matrix product given in formula 2.

$$\mathbf{a}_1 \leftarrow \mathbf{A}^\mathbf{T}\mathbf{h}_0 \tag{2}$$

Likewise, hubs are defined in terms of authorities, as in formula 3.

$$\mathbf{h}_1 \leftarrow \mathbf{A}\mathbf{a}_0 \tag{3}$$

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 2 & 1 & 2 & 0 \end{pmatrix}$$

Figure 4: Example Adjacency Matrix

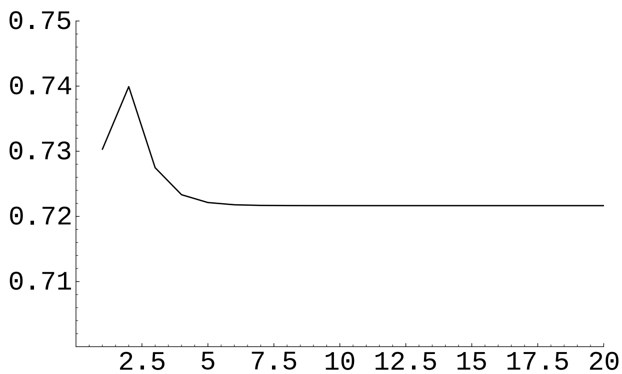

Figure 5: Authority Score for Document 1

However, the initialization of our hub and authority vectors was probably inaccurate. That is, we doubt that a vector of ones captures the true relationships among documents. The vector $\mathbf{a}_1$ provides a better picture of this information that $\mathbf{a}_0$ did. With this in mind, we may repeat the authority computation by taking $\mathbf{a}_2 \leftarrow \mathbf{A^T h}_1$. An analogous step is taken to find the second iteration of the hub scores. By repeating these iterations n times, we eventually derive $\mathbf{a}_n \leftarrow \mathbf{A^T h}_{n-1}$. It can be shown that $\mathbf{a}$ and $\mathbf{h}$ will converge onto steady states if n is allowed to grow sufficiently large. Moreover, $\mathbf{a}$ and $\mathbf{h}$ will converge on the principle eigenvectors of $\mathbf{A}^T\mathbf{A}$ and $\mathbf{A}\mathbf{A}^T$, respectively.

Deriving the hub and authority scores from an adjacency matrix is thus an iterative process. If we consider each iteration as a step of the system through its state space, the final vectors $\mathbf{a}^*$ and $\mathbf{h}^*$are those points in the space where the system eventually stabilizes.

We now demonstrate this process with a small example. Consider the adjacency matrix $\mathbf{A}$ shown in Figure 4. By applying Kleinberg's algorithm to this matrix, the elements of $\mathbf{a}$ and $\mathbf{h}$ quickly converge into a steady state, as shown in Figures 5and 6. Figure 5 traces the authority score associated with the first document across 20 iterations of the algorithm. Figure 6 shows the hub score for document one as the iterations proceed. Both hub and authority scores appear initially to be "wrong," and to be coerced toward stable values as the process moves forward. The convergence of other documents moves similarly, but we

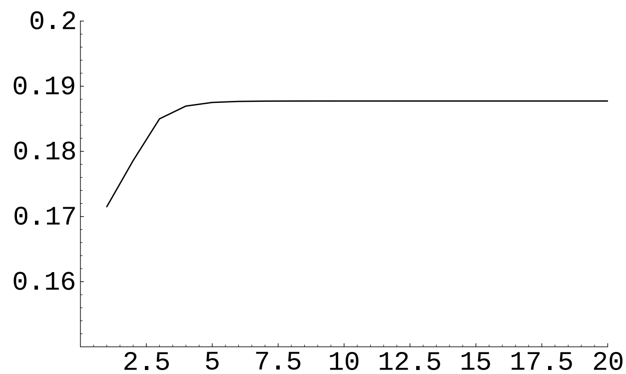

Figure 6: Hub Score for Document 1

omit discussion of them here. We may thus understand each iteration of the hubs and authorities algorithm as a phase in the system's movement through state space, with the final vectors $\mathbf{a}^*$ and $\mathbf{h}^*$ corresponding to the points of greatest stability in the space, the principle eigenspace of the original matrix.

4.2 The Vector-Space Model of IR

Keyword retrieval may also be understood iteratively. The traditional vector-space model of IR imagines documents and queries existing in a high-dimensional vector-space represented by the term-document matrix $\mathbf{A}$. Under this model, documents are ranked against a given query $\mathbf{q}$ by means of the cosine between $\mathbf{q}$ and each column of $\mathbf{A}$. Assuming that the columns of $\mathbf{A}$ have been normalized, a document ranking $\mathbf{r}$ is derived by means of formula 4.

$$\mathbf{r} = \mathbf{A}\mathbf{q} \tag{4}$$

However, the standard vector-space model fails to account for patterns of co-occurrence evident in the matrix $\mathbf{A}$. This defect is most noticeable in the inability of keyword-based IR systems to recognize semantic relations between terms. Thus queries about *cars* fail to retrieve documents about *automobiles* despite a rough equivalence of these terms to authors and end users.

To improve the vector-space IR model, many systems employ query expansion. Thus, terms are added to a query vector prior to search execution. Information regarding term-term correlation usually motivates the selection of expansion terms. We may thus imagine an augmented system that takes account of such relations by means of a third matrix $\mathbf{T}$. Rankings under such a model would take the form shown in formula 5

$$\mathbf{r}_T = \mathbf{A}\mathbf{q} + f(\mathbf{T}) \tag{5}$$

where $f(\mathbf{T})$ is some function applied to the term co-occurrence matrix $\mathbf{T}$, derived

$$
\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}
$$

Figure 7: Sample Term-Document Matrix

$$
\mathbf{T} = \begin{pmatrix} 4 & 1 & 2 & 2 & 1 & 2 & 2 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 2 & 0 & 2 & 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 2 & 2 & 1 \\ 2 & 0 & 1 & 0 & 2 & 3 & 1 \\ 2 & 0 & 1 & 1 & 1 & 1 & 2 \end{pmatrix}
$$

Figure 8: Term Co-Occurrence Matrix

by formula 6.

$$
\mathbf{T} = \mathbf{A}\mathbf{A}^{\mathbf{T}} \tag{6}
$$

Each cell $\mathbf{t}_{ij}$ of the matrix $\mathbf{T}$ contains the number of times that terms i and j co-occur in a document. Thus the main diagonal of $\mathbf{T}$ contains the total frequency of each term in the corpus.

If we consider the matrix $\mathbf{T}$ as a system in its own right, we may thus be interested in discovering its stable state. That is, we may improve our estimation of the relations among terms by allowing $\mathbf{T}$ to move through its state space, paying attention to any attractors that appear evident in this motion. Let motion through the state space of $\mathbf{T}$ correspond to multiplying a column-normalized $\mathbf{T}$ by itself. Thus $\mathbf{T}_i = \mathbf{T}^i$. As in the hubs and authorities algorithm, $\mathbf{T}_i$ converges onto a steady state.

As an example, consider the term-document matrix $\mathbf{A}$ shown in Figure 7 The term co-occurrence matrix for $\mathbf{A}$ is given in Figure 8. The normalized version of $\mathbf{T}$, $\mathbf{T}_{norm}$ appears in Figure 9. By computing $\mathbf{T}^n_{norm}$, with sufficiently large n, we derive the steady matrix $\mathbf{T}^*$, shown in Figure 10. As dictated by the theory of eigensystems, the system converges onto a matrix whose columns are the original matrix's principal eigenvector. The motion of this system through its state space is depicted in abbreviated form in Figure 11. Here we see the value for the first cell in the matrix as the matrix is self-multiplied repeatedly. As expected, the value falls into a steady state and remains there without change.

What we have found regarding term co-occurrence matrices also holds for

0.685994	0.57735	0.603023	0.603023	0.316228	0.458831	0.57735
0.171499	0.57735	0.	0.301511	0.	0.	0.
0.342997	0.	0.603023	0.301511	0.	0.229416	0.288675
0.342997	0.57735	0.301511	0.603023	0.	0.	0.288675
0.171499	0.	0.	0.	0.632456	0.458831	0.288675
0.342997	0.	0.301511	0.	0.632456	0.688247	0.288675
0.342997	0.	0.301511	0.301511	0.316228	0.229416	0.57735

Figure 9: Normalized Term-Term Matrix

0.639996	0.639996	0.639996	0.639996	0.639996	0.639996	0.639996
0.121051	0.121051	0.121051	0.121051	0.121051	0.121051	0.121051
0.332998	0.332998	0.332998	0.332998	0.332998	0.332998	0.332998
0.302913	0.302913	0.302913	0.302913	0.302913	0.302913	0.302913
0.260886	0.260886	0.260886	0.260886	0.260886	0.260886	0.260886
0.412021	0.412021	0.412021	0.412021	0.412021	0.412021	0.412021
0.36781	0.36781	0.36781	0.36781	0.36781	0.36781	0.36781

Figure 10: Steady State Term-Term Matrix

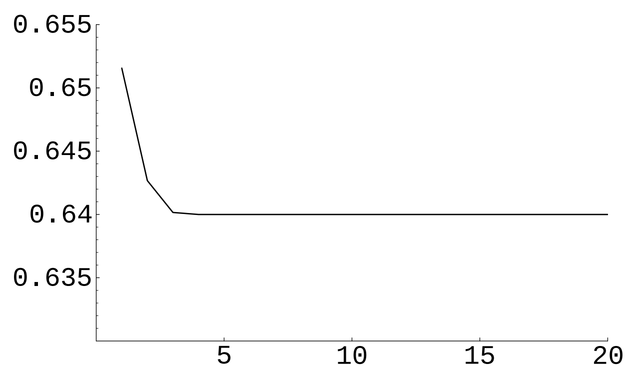

Figure 11: Value of $\mathbf{T}^i_{11}$

document-document matrices. Such matrices show the number of terms that each pair of documents shares. Thus these matrices are useful for purposes of document clustering. In both cases–terms and documents–the co-occurrence matrix constitutes a system whose principle eigenvector acts as an attractor as the system iterates through state space. The eigensystems of these matrices thus characterize the topology of this state space, suggesting useful dependencies and relationships inherent in the data.

5 Conclusion

While it seems plausible that eigensystems function as attractors in information systems, the implications of this argument for retrieval are not immediately clear. We suggest that eigenvalues are useful due to their role as attractors. In other words, they describe the topology of the information space generated by a given system. In the case of Markov chains or the hubs and authorities algorithm, the theory of eigenvalues guarantees the presence of a stable point. Defining hubs and authorities recursively relies on eigenvalues insofar as the method depends on an eventual end to its recursion.

Latent semantic indexing (LSI) [4, 5, 10, 11] uses eigen-analysis to represent terms and documents in a factor space of low-dimensionality. It is widely argued that this space approximates human "concept space" more accurately than a high-dimensional term-space can. Justification for this argument usually relies on the idea of noise reduction; a low-dimensional space captures the major semantic variations among documents, while eliding their less significant, random fluctuations. While this is true and compelling, the argument presented here suggests another justification for using factors derived from eigen-analysis. Insofar as the eigensystem captures the long-term behavior of a system regardless of its initial conditions (insofar as it approximates the system's attractors) a matrix's eigensystem constitutes a map of the system's information space. As in the example of a simple Markov model, eigen-analysis articulates the long-term behavior of an information system. And as in the hubs and authorities example, eigensystems constitute points of stability in state space. In the case of LSI, such stability is crucial because the factor space in which terms and documents are represented is derived recursively. That is, terms are made similar by virtue of appearing in similar documents, and documents are declared similar due to the terms they share. Such a complex information space is immediately present in a term-document matrix. However, in this form, the contours of the space are obscure. By discovering the eigenstructure of such a system, by discovering the topology described by its attractors, the recursive identification of similar documents and similar terms occurs naturally.

References

[1] V. K. Balasubrahmanyan and S. Naranan. Quantitative linguistics and complex system studies. *Journal of Quantitative Linguistics*, 3(3):177–228, 1996.

[2] R. F. Cancho and R. V. Sole. Two regimes in the frequency of words and the origins of complex lexicons: Zipf's law revisited. Working Paper 00-12-068, Santa Fe Institute, 2000.

[3] R. F. Cancho and R. V. Sole. The small-world of human language. Working Paper 01-03-016, Santa Fe Institute, 2001.

[4] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman. Indexing by latent semantic analysis. *J. of the American Society for Information Science*, 41(6):391–407, 1990.

[5] C. H. Q. Ding. A similarity-based probability model for latent semantic indexing. In *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1999.

[6] B. A. Huberman. *The Laws of the Web: Patterns in the Ecology of Information*. MIT Press, Cambridge, 2001.

[7] S. A. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, Oxford, 1993.

[8] S. A. Kauffman. *At Home in the Universe: The Search for the Laws of Self-Organization and Complexity*. Oxford University Press, Oxford, 1995.

[9] J. M. Kleinberg. Authoritative sources in a hyperlinked environment. In *Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 668–677, San Francisco, California, 25–27 January 1998.

[10] T. K. Landauer and S. T. Dumais. A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2):211–240, 1997.

[11] R. E. Story. An explanation of the effectiveness of latent semantic indexing by means of a bayesian regression model. *Information Processing and Management*, 32(3):329–344, 1996.

[12] G. Strang. *Linear Algebra and its Applications*. International Thompson Publishing, 1988.

[13] R. Wesson. *Beyond Natural Selection*. MIT Press, Cambridge, 1994.

[14] G. Williams. *Linear Algebra with Applications*. Jones and Bartlett Mathematics, 2001.