Sabelli, H. (2002). Arrangement, a measure of nonrandom complexity.
***Systems Analysis Modeling Simulation* 42: 395-403**

Arrangement, a Measure of Nonrandom Complexity.

Hector. Sabelli
Chicago Center for Creative Development
2400 Lakeview Avenue, Chicago, IL 60614. U.S.A.
hsabelli@rpslmc.edu

Abstract: This article presents a measure of nonrandom complexity. The recurrence of isometric vectors of N consecutive terms in a numerical series is measured, and the ratio of the percentage of consecutive recurrences to the percentage of total recurrences, which we call arrangement, is calculated. Arrangement is high in time series of physiological recordings and of economic processes, and in biotic series generated by recursions of bipolar feedback. Arrangement is low in random, periodic and chaotic series. Thus arrangement correlates with an intuitive notion of complexity, while many other measures of complexity assign the highest value to random series.

Key words: complex systems; medicine; chaos and bios.

1. Complexity and Creativity

This article presents an empirical measurement of complexity, one of the defining features of creative processes. Complexity is a focus of current research interest. Many measures of complexity assign the highest value to random series. This counterintuitive result might be acceptable if, and only if, creative processes result only from random change, such as biological mutation. This is indeed a widely held hypothesis, because mechanism cannot account for the emergence of novelty and complexity. Thus, from Epicurus' swerving of the atoms to Monod' s origin of life as a unique and aleatory event, creative evolution has been attributed to random events. However, apparently random patterns, such as chaos, can be generated deterministically. Likewise, the complex aperiodic patterns found in physiological [Sabelli et al, 1995; Kauffman and Sabelli, 1998], economic [Sabelli and Kauffman, 1999], and meteorological [Sabelli, 2000] processes, can be generated not only probabilistically using random walk models, but also deterministically with recursions of bipolar feedback. This is bios, a pattern that, in addition to chaotic characteristics, also displays features similar to those found in creative processes (see below). As every process receives both positive and negative responses from its environment, bipolar feedback may be expected to occur in may if not all natural processes [Sabelli, in press]. Complexity may thus be created by deterministic interaction between simple processes (co-creation) rather than solely from random events.

Nonrandom creative processes may be expected to imprint defining features in the time series they generate, namely episodic pattern ("complexes") [Sabelli et al, 1995], novelty [Sabelli, 2001 a], diversification [Sabelli et al, 2000], interacting opposites [Sabelli, 2001b], and complexity, the focus of this article. To develop valid measures of complexity, novelty, and other features of creativity, we take as a model time series of heartbeat intervals, because they are generated by a complex process, and are influenced by physical, biological, and psychological factors; hence they continually display new patterns. The measure of nonrandom complexity to be described here is based on empirical studies [Sabelli et al., 1995] in which we found that the ratio of recurrences to consecutive recurrences, which we called *arrangement*, is high in heartbeat interval series and low in random series. This observation suggested to us that arrangement may provide an intuitive, valid measure of complexity. This possibility is examined here by comparing time series of various degrees of complexity. Intuitively, complexity increases from random to periodic, to chaotic, to biotic processes. Mathematically, periodic, chaotic, and biotic patterns emerge at increased values of the control parameter in mathematical recursions of bipolar feedback [Kauffman and Sabelli, 1998].

2. Methods.

Time series are studied with the recurrence quantification method [Webber and Zbilut, 1994]. Sequences (vectors) of N successive members of the time series are constructed, starting with each data point. The number N is called the `"embedding". In this article, series are studied with 1, 2, ..., 400 embeddings, allowing one to study the data in frameworks of these many different dimensions. The Euclidean norm (the square root of the sum of squares of its members) of each vector is measured. Two vectors are considered <u>isometric</u> when the absolute value of the difference between their Euclidean norms is smaller than a chosen cutoff radius (0.1 in this study). When two vectors are isometric, a recurrence is counted. The percentage of isometries (as a function of the total number of possible recurrences in the sample) is a measure of repetitiveness; the percentage of consecutive recurrences, i.e. of recurrences between consecutive vectors, increases with patterning, and its is considered by Webber and Zbilut to measure determinism. Arrangement is the ratio of consecutive isometries (as percentage of all isometries) over the total number of isometries (as percentage of all the possible isometries). This term was chosen to indicate a measure of organization and its complexity which is different from those previously defined.

Another recurrence method (adopted by Webber and Zbilut in the RQA 5.2 program) defines vectors as recurrent when the Euclidean norm of the distance between them falls below the selected cutoff radius. For the sake of clarity, this is named “similarity recurrence” [Sabelli, in press]. Recurrence isometry indicates quantitative equivalence between vectors, regardless of their direction; recurrence similarity requires in addition similar directionality.

The mathematical time series studied here are random series (uniform, Gaussian), periodic series (period 2, period 16, sinusoidal, zig-zag, repetitive poems), chaotic attractors (Henon, Lorenz, Sprott), periodic and chaotic series generated by the logistic equation $A_{t+1} = A_t * R * (1 - A_t)$, and periodic, chaotic and biotic series generated by the process equation $A_{t+1} = A_t + g * \sin A_t$. modeling bipolar feedback. The empirical time series studied were heartbeat series (R to R intervals, RRI) obtained from the CCCD-Rush Psychocardiology Data Base of 24 hour electrocardiographic recordings obtained by Holter monitoring from adult subjects [Carlson-Sabelli et al., 1994, 1995; Sabelli et al., 1995 a, b], electroencephalogram, electromyogram, respiration [Webber and Zbilut, 1994], water temperatures in the Pacific ocean during El Niño, speech (“e” vowel sound), paleoclimate indicators from coral samples obtained from the NOAA/NGDC Paleoclimatology Program, Boulder CO, USA), the Dow-Jones Industrial Average (DJIA), prices of commodities, and exchange rates of currencies [Sabelli and Kauffman 1999; Sugerman et al 1999].

3. Results

Arrangement, the ratio of consecutive isometries over isometries, is high in time series of physiological recordings (electrocardiogram, electroencephalogram, electromyogram, respiration), economic time series (Dow-Jones Industrial Average; prices for crude oil, corn, gold, silver; exchange rates for American, British, Canadian, Danish, Japanese, European, currencies), and biotic series generated with recursions of bipolar feedback $A_{t+1} = A_t + g * \sin A_t$, that is to say, complex processes (figure 1). Arrangement is low in simple series, including simple speech sound, oceanic temperature change (El Niño), random data, periodic series, and chaotic series. In periodic series, arrangement increases with the complexity of the periodicity (period 2 < period 4 < period 16, etc). In series generated by the process equation (figure 2), arrangement increases with the complexity of pattern (periodic < chaotic < biotic), in contrast to isometries and consecutive isometries that decrease from periodicity to chaos to bios. Neither the ratio of consecutive recurrence to recurrence nor either one of these measures alone correlates with complexity when similarity recurrence is measured.

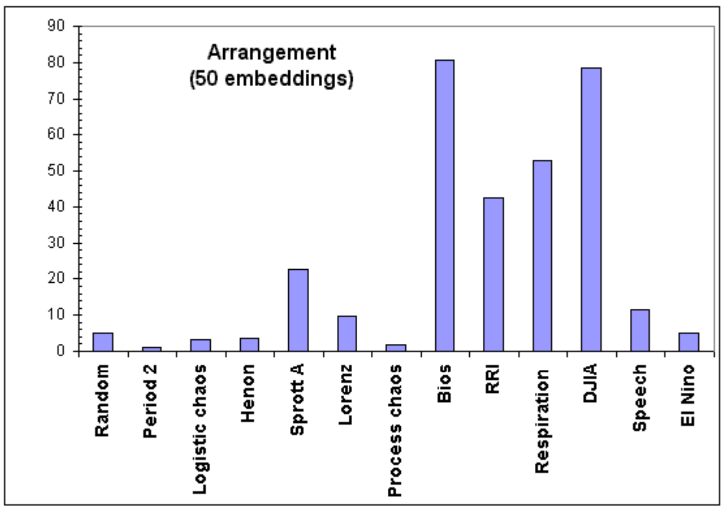


Figure 1: Arrangement for series of 1000 terms calculated with 50 embeddings and cutoff radius 0.1.

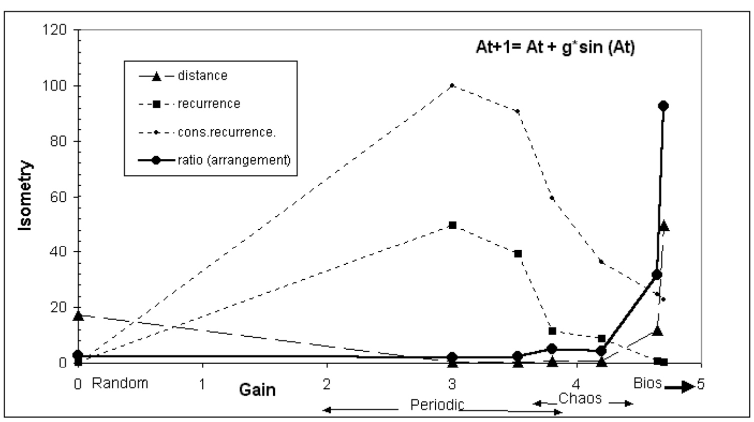


Figure 2 A: : Isometry recurrence, consecutive isometry, and their ratio (arrangement) for patterns generated by recursion of bipolar feedback $A_{t+1} = A_t + g * \sin A_t$ (initial value = 100) at increasing gain. Analysis performed with 3 embeddings; cutoff radius 0.1.

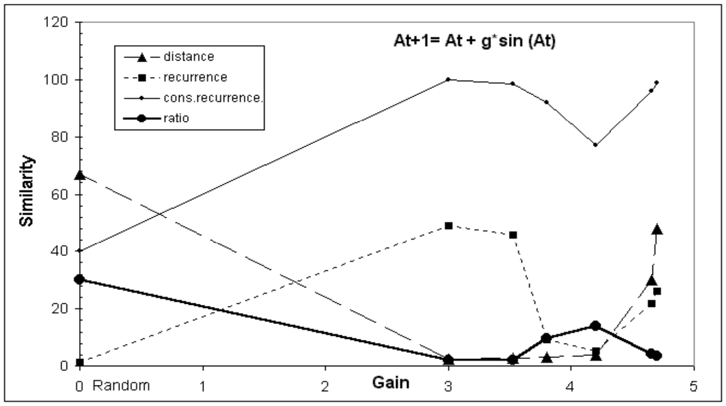


Figure 2 B: Similarity measurements for patterns generated by recursion of bipolar feedback.. For comparison, data for a uniform random series (range 50 to 150) are also plotted (at gain 0). Analysis performed with 3 embeddings; cutoff radius 1.4.

Arrangement correlates with complexity within a wide range of embeddings, even when isometry and consecutive isometry vary in a nonlinear manner with the number of embedding dimensions. Examining a wide range of embeddings allows one to consider the organization of the time series at various temporal dimensions. Periodicity can be demonstrated only with embeddings that coincide with the period (figure 3); e.g. 365 embeddings are required to demonstrate annual periodicity. Simple periodic series show 100% isometry and consecutive isometry (arrangement = 1) when the embedding coincides with the period; at other embeddings, arrangement remains low, even when recurrence is also low. The embedding plot of sine waves likewise shows periodic peaks of recurrence, at which arrangement is 0. Plotting recurrence measurements as a function of the number of embeddings allows one to detect coexisting patterns in more complex time series (figure 4). I n many empirical time series, such as heartbeat intervals, as well as in chaotic and biotic aperiodic series, isometry, consecutive isometry, and arrangement are high at very low embeddings, denoting the similarity between consecutive terms of the series, absent in random data. Recurrence decreases with increasing embeddings up to a minimum (circa 6 embeddings for heartbeat intervals and for biotic series). Biotic and chaotic series cannot be readily differentiated at these low embeddings. As the number of embeddings is increased further; isometry, consecutive isometry, and arrangement increase, up to a maximum, and then decrease. The same obtains for random data. There is, however, a distinct difference between biotic series, with high arrangement, and random and chaotic data, that show low arrangement.

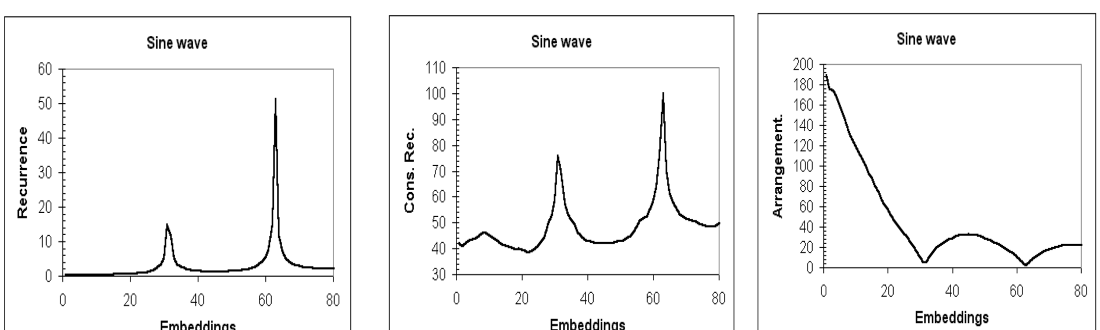
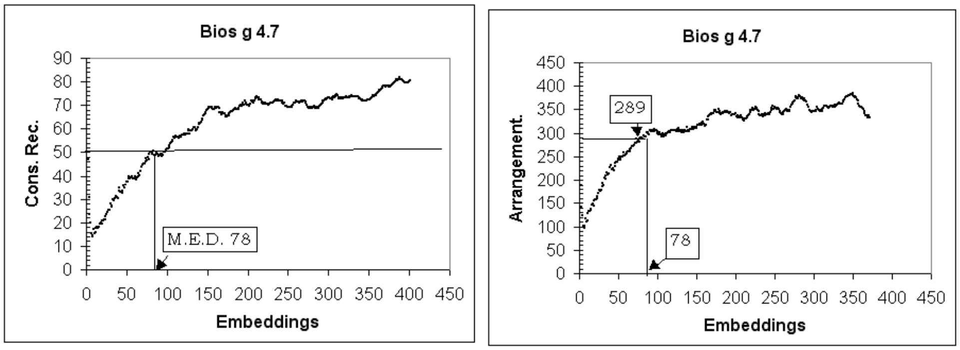


Figure 3: Embedding plots showing the value of isometries, consecutive isometries, and arrangement at increasing embeddings for a sine wave (sine time/30).



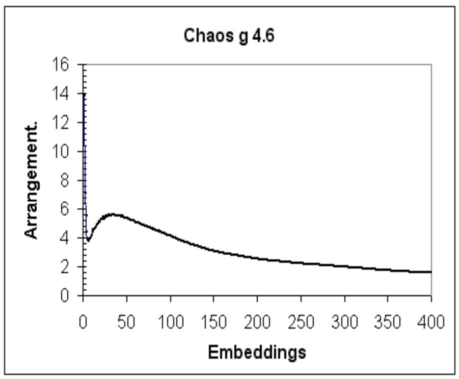


Figure 4 Embedding plots of time series generated by the process equation at gain 4.6 (chaos) and 4.7 (bios). Arrangement is high in bios at all embeddings, and low for chaos, except at very low embeddings. The figure illustrates the identification of the median embedding dimension (MED), used to measure arrangement in Table 1.

Useful comparisons among numerical series with widely different range of values can be made by measuring recurrence and arrangement at the Median Embedding Dimension (MED), which is defined as the embedding at which 50% of the isometries are consecutive [Sabelli et al, 1995]. As shown in table 1, the MED of empirical time series are of the same order of magnitude as those for biotic series, and much lower than those observed with random and chaotic series, and with 1/f noise. Regardless of their MED, arrangement is high in complex series, including 1/f noise, and lower for simpler series, including paleoclimate markers. Although arrangement and the MED obviously measure different types of complexity, both markers indicate that heartbeat interval series recorded from healthy persons are more complex than those recorded from patients with psychiatric illness. This is consistent with the view that cardiac rate and variability are largely determined by higher central nervous system activity.

Shuffling the data prior to calculations decreases arrangement in biotic series, and increases it in the case of periodic data. These results indicate that arrangement correlates with our intuitive notion of complexity. In the case of chaotic series, shuffling decreases arrangement al low embeddings (as for bios) and increases arrangement at high embeddings (as for periodic series), indicating two different forms of organization. Arrangement is reduced by shuffling for physiological and economic data, regardless of the embedding dimension. This is consistent with the notion that many natural processes are creative, and exhibit a biotic pattern, rather than chaotic or random patterns [Sabelli, 1999].

The measurement of nonrandom complexity is part of a comprehensive process method that identifies and quantifies the defining features of creativity in time series [Sabelli, 1999]; computer programs for these analyses are collected in the Bios Data Analyzer [Sabelli et al, 2000] to be published in a special issue of this journal.

Table 1. Arrangement at the Median Embedding Dimension (M.E.D.)
Isometry measured with delay 1, cutoff radius 0.1. * N = 1000 ** N = 3500 *** N = 7000.

	M.E.D.	**Arrangement**
Random (uniform)***	214	16
1/f pink noise **	120	97
Logistic chaos*	363	15
Process chaos g = 4.3**	345	12
Bios g = 4.61*	18	42
Bios g = 4.65*	21	57
Bios g = 4.66**	53	189
Bios g = 4.7**	78	289
Healthy RRI (N=31). **	60 ± 4	61 ± 4
Coronary disease RRI (N =9) *** (mean ± S.E.)	36 ± 12	61 ± 20
Depressed RRI (51 samples)** (mean ± S.E.)	37 ± 2	49 ± 2
Bipolar depressed RRI (N=9) *** (mean ± S.E.)	52 ± 17	62 ± 20
Affective psychoses RRI (N=45) ** (mean ± S.E.)	41 ±3	49± 3
Schizophrenic RRI (N=9) *** (mean ± S.E.)	25 ± 8	34 ± 11
Dow Jones Industrial Average**	10	54
Corn **	35	98
El Niño*	37	58
Paleoclimate marker (Seychelles)*	33	15

Acknowledgements: This research was supported by the Society for the Advancement of Clinical Philosophy. I am thankful to Drs. C. Webber and J. Zbilut for the use of their programs, and to Linnea Carlson-Sabelli for collaboration in earlier stages of this research.

Kauffman, L. and Sabelli, H. 1998. The Process equation. *Cybernetics and Systems 2*9: 345-362

Sabelli, H. 1999. Process theory: mathematical formulation, experimental method, and clinical and social application. *Toward New Paradigm of System Science.* P. Y. Rhee editor. Seoul: Seoul National University Press, pp 159- 201.

Sabelli, H. 2000. In Peoria, the weather is biotic. *General Systems Bulletin* 29: 9-10.

Sabelli, H. in press. Bipolar Feedback. *Encyclopedia Of Life Support Systems.* UNESCO.

Sabelli, H. 2001a. Novelty, a Measure of Creative Organization in Natural and Mathematical Time Series. *Nonlinear Dynamics, Psychology, and Life Sciences* 5: 89-113.

Sabelli, H. 2000b. Complement plots: analyzing opposites reveals Mandala-like patterns in human heart beats. *International Journal of General Systems* 29: 799-830.

Sabelli , H. and L. Kauffman 1999. The Process Equation: Formulating And Testing The Process Theory Of Systems. *Cybernetics and Systems* 30: 261-294.

Sabelli HC, Carlson-Sabelli L, Patel M, Zbilut J, Messer J, and Walthall K. 1995. Psychocardiological portrait. In *Chaos theory in Psychology.* F. D. Abraham and A. R. Gilgen (Eds). Greenwood Pub., Inc., Westport, CT. pp 107-125.

Sabelli, H. Sugerman, A. Patel, M. Kauffman, L.,. Konecki, J.,. Carlson-Sabelli, L., Abouzeid, A., Sween. J. 2000. Bios Data Analyzer. 1. Process Method to Analyze Creative Processes and its Scientific and Medical Application. W*orld Congress Systems Sciences*, Toronto.

Webber, Jr., C.L. and Zbilut, J.P. 1994. Dynamical assessment of physiological systems and states using recurrence plot strategies. *J. Applied Physiol.* 76, 965-973.