Synthesis of a Macro Sound Structure within a Self Organizing System

Sinan BÖKESOY*
* CICM, University of ParisVIII, FRANCE, sinan@sonic-disorder.com

***Abstract* — This paper is focused on synthesizing macro-sound structures with certain ecological attributes to obtain perceptually interesting and compositionally useful results. The system, which delivers the sonic result is designed as a self organizing system. Certain principles of cybernetics are critically assessed in the paper in terms of interdependencies among system components, system dynamics and the system/environment coupling. It is aiming towards a self evolution of an ecological kind, applying an interactive exchange with its external conditions. The macro-organization of the sonic material is a result of interactions of events at a meso and micro level but also this exchange with its environment. The goal is to formulate some new principles and present its sketches here by arriving to a network of concepts suggesting new ideas in sound synthesis.**

I. Introduction

Generally human made acoustical instruments produce compositionally meaningful sounds, which can be described by some interactions such as excitation and resonance. In such a structure, the excitation consists of a temporal energy input, a sort of some disturbance introduced to the system [1].When an acoustic resonant system is excited, it does attenuate some frequencies and emphasize certain ones. The resonance produces a pattern of energy dissipation. Such resonant systems reach a final stable state because energy is not generated within the system but it is received from an external source. And since the external source comes as an event trigger mechanism with a temporal energy supply having a transient behavior, the response of the system is finite. The interaction of such instruments with its environment happens at this excitation level. Although there might be involved chaotic dynamics on the micro level components of these systems, the macro level output should have a distinctive character as a response to a certain input. The output of the system should be predictable as it makes sense for the performer who would like to control it precisely with the input parameters of his performance. The interaction becomes in what ways the resonant system responds to its excitation system.

Our system which we are going to introduce as a sound synthesis instrument has a much more complicated structure than the one explained above. It will have a control mechanism which serves to the user interaction, but also a self organizing system to deliver dynamic behavior of eco-systemic kind. All this will be exploited within the principles of cybernetics.

II. System structure

Ecological systems have generally a hierarchy of multiple levels of organization on multiple time-rate scales in order to be ecologically valid [2]. Compositionally meaningful sound objects are subjected to a temporal change, spectral variation and envelope, which results as a pattern of change. This pattern becomes a structure establishing a form of the sonic identity perceivable by our ears such as we could be able classify the character of the sound source. In the case of ecologically valid sound objects, the patterns of change are expected to form higher-order percepts. In other words, within the hierarchy of multiple levels, the interaction of low level elements shows emergent properties at higher levels. This interaction occurs among all levels. Organization at one level influences the others [3].

In our system, we think the smallest element as a granular micro-sound event. Sounds constructed with the granular technique require high densities of short events to produce aurally convincing sound textures. Therefore, computer music composers have adopted algorithmic approach to handle granular synthesis with statistical controlled distributions combined with tendency masks, probabilistic functions and other methods while exploring the possibilities of controlling the granular streams [4] [5] [6]. In general, the lack of these applications for creating macro-temporal patterns was the employment of a mono-layered time structure, which was missing a hierarchical organization of multiple layers and complexity.

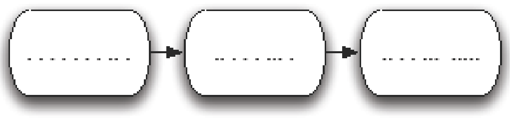

Fig. 1. The event distribution inside consecutive time cells. Since there is no information exchange between these cells and no feedback from any output to input, no evolution is possible in this organization. The distribution in each cell is independent and stochastic.

An alternative event distribution system operating on multiple time-scales within a self similar structure has been proposed [7] [8]. The event distribution mechanism on each level of Cosmos is using deterministic or stochastic functions for making decisions on each events onset time and duration parameters on that level (Figure 2). Each event opens a subspace with other events on a lower level in the structure.

As we see on Figure 3, the mechanism of sound synthesis starts by injecting a sound element as the input on the micro level (forming the excitation system) and the bottom-up procedure constructs the meso-spaces and finally the macro-space. The micro level characteristics of a sound grain influence the meso and macro properties of the sound event. The output of the system is fed back again to the input, as the micro-event sample data. This recursive system uses its own output and creates a multi layered development in the sonic structure.

The particular transformations at the macro, meso and micro levels are user defined but they are fixed during the operation and non-adaptive. For example if any of the transformations features some 'destructive' processing, meaning that it leaves something out, the result of the feedback process is degradation.

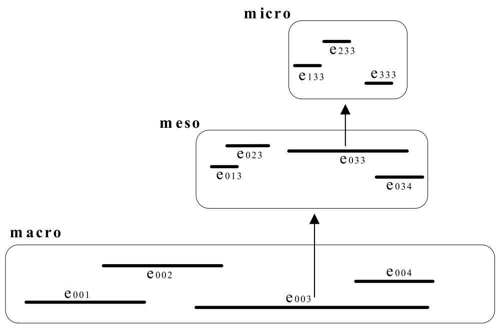

Figure 2: The event distribution mechanism in Cosmos. It is based on a self similar structure and interdependency between layers.

If the transformation is asynchronous granulation, the degradation results to a sonic powder; or if the output gains a noisy character then the feedback reinforces this behavior and the degradation of harmonic structure follows rapidly. But if there is any slight regular pattern in the function, it would be replicated at the output so it is not a drift towards white noise or distortion, and will have present a similar phenomena such as in multi-layered pulsar synthesis [6].

This behavior is the result of the positive feedback, which creates an attractor strong enough winning out all of the other features in the system. The feedback system soon points to its end-states. What is being experiencing with degradation behavior here, is a lack of "structural coupling", a lack of "adaptation" to the input (although, for the moment, the input = the system output). In other words, the transformations, or their composition together at the macro level, are insensitive to the input/output, they just work the same way independent of what comes in. When such a system eliminates the destructive relationship to its input/output, it becomes an eco-system [9]. For establishing this, we should use the selected external conditions of the environment, which will be derived from the analysis of the output to maintain the perceptual attributes of sound such as; noisy/harmonic character, intensity, spectral distribution, rhythmical behavior, particle density etc…

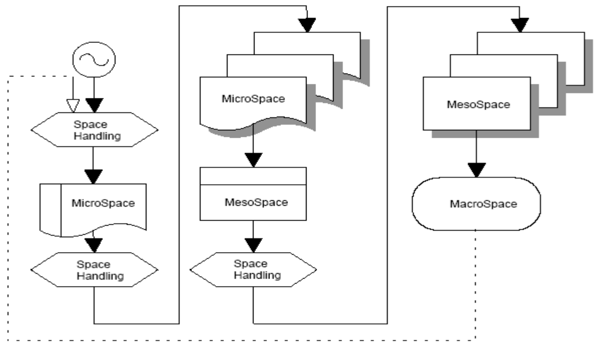

Figure 3: This is the schematic overview of the audio routing and the event space organization inside the application. The audio enters top-left. The dotted line is the feedback line.

III. STRUCTURE&ENVIRONMENT COUPLING

Non-equilibrium systems are driven away from a stable position and exhibiting dynamic behavior. In many cases the transition period between states is significant (e.g. metamorphosis in insects takes time). This period, called a transient, is a non-equilibrium state (equilibrium here refers just a constant state, not only to the lowest energy state familiar from physics). It is the transients that are the actual behavior. What we have here is a closed stochastic system, so the steady state in our system is irrelevant (not counting the deterministic distribution functions). Complex systems of this sort never settle to a fixed status, maybe the only fixed state is the silence in our case. It is subject to constant perturbation (due to the input sound to the Cosmos, which drives bursts of transient behavior).

This instability with order is what we call 'Edge of Chaos', a system midway between stable and chaotic domains (also called self-organized criticality) [10]. It is characterized by a potential to develop structure over many different scales and is an often found feature of complex systems whose parts have some freedom to behave independently such as in Cosmos. For 'edge of chaos' behavior we need some constraints – too many dynamics will die out, too few and absolute order will not be sustained.

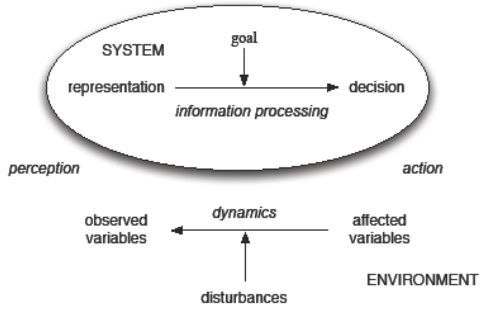

Figure 4: A typical control system in cybernetics. This model can be applied to many possible dynamic structures in different disciplines.

The question is how to establish an evolutionary process of self-organization possible with this system. Because we enter the world of cybernetics, the system design should be considered within its principles [11]. Cybernetics was defined by Wiener as "the science of control&communication, in the animal and the machine", in a word, as the art of steermanship. Our system concept includes; complexity, system-environment boundry, process, state, hierarchy, feedback and network of coupled variables.

The aim would be first of all to create a dynamical system which will represent a structurally closed but organizationally open system [Figure 4]. What is fixed here is the structure of Cosmos, its integrity with its interconnected components. Cosmos should determine its state by the interactions with the environment and among its system components by showing an adaptive behavior to the external parameters, which is essential to self-organization. This approach of cybernetics is taken universal and is valid for any sound-environment interaction, not specifically for the Cosmos model [9]. But the stochastic complex behavior of the Cosmos model introduces some interesting possibilities which will be discussed below.

A. *The Observing Part*

In our case, the system listens to its output, which becomes its 'environment interaction' (Figure 5). The output of the system itself is the perturbation introduced to the system against which it should react and organize its state. In an ideal system, to every class of perturbations there corresponds a class of adequate counteractions. This correspondence might be represented as a homomorphism from the set of perturbations in the environment to the set of compensations.

The observer part keeps tracking certain features of this environment and does an analysis on the sound for certain perceptual attributes, which are set at the design level. It becomes in the end a self-observing system, ultimately using information on its input/output, decides for the emergent behavior to take against the external conditions in order to re-organize itself.

If one implements some flexible behavior in any or all of the three levels, macro meso and micro, the representation of the selected external conditions will be evaluated and actions will be taken inside the new self organizing structure. The sound output of the system itself can be interpreted with useful analysis methods to extract the following criteria (Figure 5). The sound analysis can be further interpreted with statistical analysis tools to obtain the descriptive characteristics of the data sets with the arithmetic mean value $\overline{x}$, the standard-deviation σ, which is the most common measure of statistical dispersion, and the skew, which is a measure of the asymmetry of the distributed values for the incoming stream of observed values. The standard calculation method for these quantities is as in the equation (1).

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \ , \quad \sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2} \qquad (1)$$

$$skew = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{x_i - \overline{x}}{\sigma}\right]^3$$

There can be different integration times on these observed values for corresponding to macro, meso and micro level time scales. The feature extraction process generates control rate data x_i, which is interpreted for the reaction of the system only possible after a certain delay time depending on these integration times.

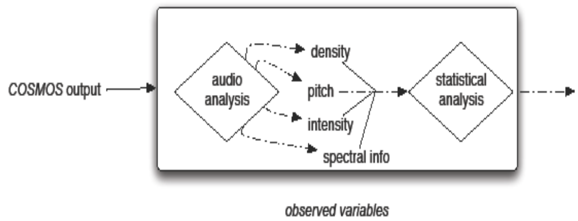

Figure 5: The Cosmos output is being analyzed and the observed variables are extracted in order to be perceived by the system and represented inside the system. Audio input comes in and control rate data comes out.

- *intensity analysis* : At the very low level, the intensity detection is the basic observed parameter, where one can follow the envelope of the sound input to the system.

- *pitch analysis* : When talking about the discrete event pitch, the value can be centered and for an event cluster, the pitch analysis can be interpreted in a range from deterministic output to noisy character. Also when the event distribution happens with a regular rhythmic order, then with fast distribution rates, the rate itself is being perceived as a pitch. So the rhythm becomes to continuous pitch value as an emergent property.

- *spectral processing* : Filtering methods emphasize a certain region of the spectra, therefore increasing the redundancy. Applied in a feedback loop, it is analogous to the Larsen tone effect, which is the true acoustical feedback. The use of several BandPass filters would allow following and matching certain regions in the spectrum of the incoming audio such as in the classical vocoder synthesis. The spectral centroid is also an elementary parameter, which can be observed for finding a correlating value for the spectral brightness of the incoming audio.

- *density* : The density here is the quantity of events in a certain time span. But when they are overlapping, the detection of the distinctive events becomes difficult although perceptually its existence is evident depending also to their pitch distribution. When the pitch distribution is wider, the perception of discrete events becomes easier. The more density of the overlapping events, the more becomes the perceived intensity. Therefore density and intensity are correlated.

For the implementation on Max/MSP, the 'analyze~[1]' object can be used to determine the *pitch*, *loudness*, *noisiness* and *brightness*. The loudness is already an average value, because the STFT is being used with an overlapping window function which does extract the spectral contour along the bins by averaging their intensity along the window size. The object 'lp.stacey[2]' on Max/MSP can be used to report the statistical analysis for the *mean value, standard deviation* and *skew* quantity; but also one can implement these functions easily with JavaScript using already existing source codes.

Furthermore one can apply low-pass filter for smoothing the noisy data if needed with a more efficient method then by a simple moving window average. The Savitzky-Golay smoothing filter [12] is an example and can be used before the statistical description stage of the data.

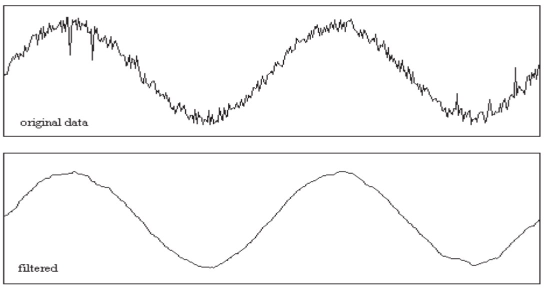

Figure 6: An example to the smoothing filter process by Savitzky-Golay method. A second order polynomial has been used to filter the noisy part from the data, which is easily implemented on MaxMSP.

We are not going to dive here into details and problematic in extracting the features like the precise pitch, amplitude and density values on the incoming data in terms of reliability. It is the reason, why we have suggested having a descriptive analysis of the data with the statistical functions. It is more efficient in our case to use the statistically significant output of the analysis and make use of this on the decision part. The results we intend to compare are interpreted on the high level perceptual attributes, not on the audio level information such as the waveform. The artifacts of the analysis represent the incompleteness regarding the representation of the external conditions.

But also according to the cybernetics principle of 'incomplete knowledge'[13], the model embodied in a control system is necessarily incomplete, the system cannot represent itself completely, and hence cannot have complete knowledge of how its own actions may feed back in to the perturbations.

B. The Decision Part

Adaptive behavior can be at any time scale, and can follow the perceptual attributes and react by changing the system variables depending on these conditions. In order to adequately compensate perturbations, a control system must "know" which action to select from the variety of available actions. This is the law of 'Requisite Knowledge of Cybernetics' [13]. Without the knowledge the system would act blindly to the external conditions. The analysis results should be represented in the Cosmos model with the relevant system variables in order that when the system looks to "itself", "itself should contain the representation of selected decisions in reaction to the external conditions. Likewise the adaptive behavior leads and forces the system to change itself [2].

The system variables in the Cosmos model are the onset and duration distribution functions (a range from deterministic to various stochastic functions), density distribution functions, and stochastic modulation generators which do affect pitch, intensity, filter parameters on macro, meso and micro levels independently. The overall goal of these functions is to achieve control on each event space and perform the process of change on the appropriate operation level. This organized spatial distribution of events and modulation functions are reminiscent to *morphogenesis* in developmental biology, where it is the study about understanding the processes that control the organized distribution of cells during the development of an organism. This change is controlled by the genetic program and can be modified by the environmental factors. The decision part in our system, should deliver this genetic code in controlling the 'organic' function of Cosmos, which develops the morphological aspects of sound.

The challenge of the user is how to describe the organic character and translate it in the decision mechanism with the available parameters. This involves the classification of the macro-sound structures.

The problematic in this classification effort is also the definition of the spectral sound morphology, which is a process of change, a transition between states in the timbre space [14]. Everything is a matter of degree. Within this scheme, we suggest to use fuzzy logic operations in its linguistic form to allow partial membership in a set of macro sound characteristics. Therefore, this fuzzy inference step takes control in the decision module, where the statistical analysis coming from the observer module is the input parameter and the macro sound representation is the output. One can use the fuzzy logic control kit [15] for the implementation on MaxMSP.

1 *analyze~* object developed by Tristan Jehan

2 *The Litter Pack* objects developed by Peter Castine

Implementing and regulating the system behavior in Cosmos is complex since the interdependencies among these system variables are subject to create unexpected emergent behavior and they indirectly implement the system dynamics. There are two possible situations of emergent behavior here;

- Within Cosmos; unpredictable emergent behavior because of the stochastic self similar structure. It does emerge and maintain itself at the 'Edge of Chaos'.
- The action which the system takes for organizing itself in the direction of the decision mechanism. It does produce order bottom-up.

The decision mechanism, which is designed as an external application, will decide for the manipulation and application of these i/o functions according to the desired goal [Figure 7]. The intervening user could assign the settings at the initialization point. With regard to any parameter observed, the decision part can be a conformer or a regulator at the most basic level. The parameter which is compared by the decision module, is going to be observed inside certain boundries for an statistical inspection of the value. This distribution range and the boundries will be set also by the user.

parameter value + distribution range ⟶ boundries of parameter space

- Regulation tries to maintain the parameter in the system at a constant level, regardless what is happening in the external conditions.
- As a conformer, it allows the environment to determine the parameter, therefore applying a positive feedback in adapting itself to the external conditions.

The output of Cosmos represents the emergent behavior including the unpredictable elements in the bottom-up development. It is also the perturbation which enters to the system and causes the transient action. Also represented in Figure 7, the chain of actions will be like below in the system.

- The observer does the analysis;
- The decision module compares the results with its internal conditions and decides what action it should take against the incoming external conditions;
- The reactions are mapped to the Cosmos parameter space in order to represent them inside the system;

The internal conditions of the decision module are set by the user, which reflect a compositional function introduced by the user.

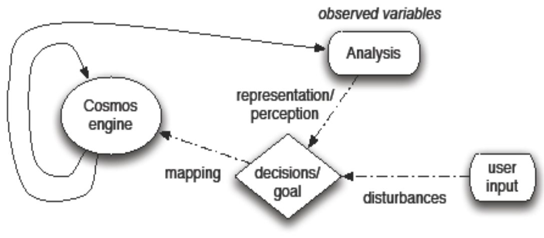

Figure 7: The Cosmos output is being analyzed and the observed variables are extracted in order to be perceived by the system and represented inside the system. Audio input comes in and control rate data comes out.

The delay between the decision point and the observation beginning is merely based on the observation process length depending on the macro, meso and micro level processes in Cosmos. For instance, if we would like to analyze the process of change on the macro-space in Cosmos, the data will arrive after the macro-space cell time, and the decisions will be taken and applied at the start of the next macro-space cell. The more the delay time, the more difficult establishing the precise control. At some point, transients may not have enough time to settle down because of the late reaction of the decision mechanism and the system will exhibit chaotic behavior.

IV. Dynamic Modes of Operation

What we see in first case is a dynamic complexity, where the structure of a system may be simple, but the behavior unpredictable. It is a property of its behavior. Some authors define that complex systems may be still divided up to complex adaptive systems and complex deterministic systems. In a deterministic system it is always possible to predict the final state, if the initial state is known. In a complex adaptive system it is not possible to know the initial state in such a detailed level, that the final state could be predicted. This is due to the non-linearity and the loops of positive feedback.

It is expected that the implemented control system establishes an attractor, where the behavior of the Cosmos complex stochastic system with unpredictable patterns would approach to a configuration of the phase state characteristic to the control system attractor. The goal-directedness of this system is suppression of these unpredictable patterns and deviations from the basin of the attractor. Musically a steady state tone or silence can be regarded as a stable state and noise as a chaotic state for the system. What is interesting here is the other equilibrium states for the observing mechanism which are at the 'edge of chaos'. So in going from any state to one of the equilibriums (the goal presented by the decision mechanism), the system is going from a larger number of states to a smaller. In this way, it is performing a selection and this reduction in the number of reachable states signifies that the variety, and hence the statistical entropy of the system diminishes. This is called again the process of self-organization.

In general a complex system will have separate dynamic modes of operation. We are considering Cosmos as a dissipative system, which take energy input (its own output as the audio material feedback and the control data from the decision mechanism) to maintain its homeostatic position. It is the flow of matter and energy through the system that allows the system to self-organize, and the exchange entropy with the environment.

Homeostasis is the property of an open system, the self-maintaining nature of systems from the cell to the whole organism [16]. Reactive homeostasis in biological systems is an immediate response to a homeostatic challenge such as predation. This predation is depending on the structure implemented in the event distribution and modulation mechanism of Cosmos; the ability of its morphogenesis. Especially in sound synthesis there can be many low level sound operations which deliver non-linear, non-reversible processes such as the frequency modulation technique, where there is a non-linear relationship between the input spectra and the output spectra. *Homeostatis* is a feedback phenomenon which cannot exist without reaction. In our case with Cosmos, it reestablishes reactive compensation, that is to say reestablishing the desired internal state according to the decision part of the system.

Will the output of the system respect exactly the desired control, or will it be asymptotically close to its destination? (a question regarding the evolution theory) How do we define the attractor from a compositional point of view?

The environment which is introduced by the user on the decision mechanism is a pre-defined artificial space with its statistically significant perceptual attributes. With these parameters one specifies the attractors characteristic for certain musical/sonic features. If the combination of these control parameters offers the variety of an environment which becomes the desired timbre space, the perturbations in the input would lead the system to re-organize itself approaching the attractors within this timbre space.

V. SOME COMPOSITIONAL ASPECTS

Stockhausen[3] states that any separation between acoustics and music is no longer meaningful in this era of computer aided sound design. The line between musically interesting synthetic sounds and digital sound effects can be very thin. Therefore the use of such a system is interesting in both ways. There are intriguing compositional aspects necessary to mention here and we claim that the philosophical structure is compatible with the ideas of musical sound&form having progressed since the beginning of 20th century. John Cage[4] has regarded the form, the structure (the divisibility of a whole into parts) as the expressive content, the morphology of the continuity. The simultaneous morphologies in different dimensions, which focus on particular perceptual attributes of sound, lead to the existence of form.

Luigi Russolo[5] has assessed that the musical art is searching the most dissonant combination of sound, the most strange and strident, namely a musical noise. Algorithmic composition has dealt in the beginning with the generation or transformation of notes and phrases. In this case, the composition of larger temporal forms is a process of both composing the phases and also organizing them into larger structure (Iannis Xenakis[6]). Within the ability to reach the microstructure of sound, the process above has become the composition of sound and then composing with the sound [17]. At this point the note event is no longer assigned to specific sound sources as we call them performed instruments, and particularly the process of composing the sound itself has gained the ability to deliver the formal structure and complexity by accessing the morphologies in different dimensions of the structure.

The system which we have presented here is, aiming towards shaping the timbre space in that direction too. The indeterminacy is built in due to the unpredictable emergent behavior. We define the restrictions and constraints at the decision module, which demands the system to re-organize itself and shape the complex sonic output according to the compositional needs. The constraints let the sound object still evolve organically within these dynamics, so it would be a compositional idea to define the balance point, the distance from the attractor defined by the environment; the requisite variety versus the requisite constraint. The question would be what is the musical perceptual meaning of this trade point?

The “fitness” of states in the system is determined by how closely they match the formal needs of the particular dimensions set in the decision module. In the sense of the direction of evolution, what would be the average fitness with compositional means? Can we classify this with existing terms like harmonic, non-harmonic, order-disorder, and with general terms describing the macro-sound object? These questions are pushing this research forward and suggesting the planning for future work.

ACKNOWLEDGEMENTS

I’d like to thank Agostino Di Scipio for his inspiring work and for his support in leading me with his suggestions about this research.

REFERENCES

[1] D. Keller, touch’n’go: Ecological Models in Composition. Master of Fine Arts Thesis, Simon Fraser University. Burnaby, BC, 1999.

[2] D. Keller., B. Truax., Ecologically-based granular synthesis in Proceedings of the International Computer Music Conference. Ann Arbor, MI: ICMA. 1998.

[3] Bregman A.S., Auditory Scene Analysis: The Perceptual Organization of Sound.Cambridge, MA: MIT Press, 1991.

[4] I. Xenakis, Formalized Music- Revised Edition. New York: Pendragon Press, 1992.

3 in *Perspectives of New Music* 1(1), 39-48, 1967

4 in his essay on “*Indeterminacy*” 1958

5 *L’Art des Bruits*, Manifeste Futuriste 1913

6 beginning with his *Achorripsis* 1956 and *ST* series compositions 1956-62

[5] B. Truax., “Real-time granular synthesis with a digital signal processing computer” in: Computer Music Journal 12(2), pp. 14-26.,1987.

[6] C. Roads., “Sound Composition with Pulsars” in : Journal of the AES, vol. 49, no. 3, pp. 134-147, March 2001

[7] S. Bokesoy., “The Cosmos model, an event generation system for synthesizing sonic structures”, in Proc. of International Computer Music Conference (ICMC’05), Barcelona, Spain, pp. 259-262, 2005.

[8] S. Bokesoy., “Feedback Implementation within a Complex Event Generation System for Synthesizing Sonic Structures” in Proc. of Digital Audio Effects (DAFx’06), Montreal, Canada, pp. 199-203., 2006.

[9] A.. Di Scipio., “Sound is the Interface” in Proceedings of the Colloquium on Musical Informatics, Firenze, Italy, 2003

[10] C. Lucas, Perturbation and Transients-The Edge of Chaos, available at http://www.calresco.org/perturb.htm, accessed February, 2007.

[11] W.R. Ashby., Introduction to Cybernetics, Methuen, London. 1956-1999 (electronically republished at http://pcp.vub.ac.be/books/IntroCyb.pdf)

[12] W. Press and S. Teukolsky., Numerical Recipes in C, Second Edition, Cambridge,University Press, pp. 650-655, 1997.

[13] F. Heylingen., Principles of Systems and Cybernetics: and evolutionary perspective, in Cybernetics and Systems 92, R. Trappl (ed.), (World Science, Singapore), pp. 3-10, 1992.

[14] D. Smalley., Spectro-Morphology and Structure Processes, in: S. Emmerson (ed.) The Language of Electroacoustic Music Basingstoke: Macmillan, pp. 61-93, 1986.

[15] R. Cadiz and Gray S. Kendall, “Fuzzy Logic Control Tool Kit: Real-Time Fuzzy Control for Max/MSP and Pd” in Proc. of International Computer Music Conference, New Orleans, 2006.

[16] N. Wiener., Cybernetics (Control and Communication in the Animal and the Machine), MIT Press, Cambridge MA, 1948,1961.