From Trophic Webs and Hierarchical Energy Transformation Systems to Multilayer Neural Networks

Peter WINIWARTER (1) and Czeslaw CEMPEL (2)

Address - *(1) Bordalier Institute, 41270 Droué, France, fax + 33 5480 1937, e-mail : winiwarter@compuserve.com*
(2) Poznan University of Technology, ul. Piotrowo 3, 60-965 Poznan, Poland, fax +48 61 78 2307, e-mail : cempel@pozn1v.tup.edu.pl

Abstract.
The key idea put forward in this paper is the formal equivalence between energy transformation processors undergoing birth&death processes (Birth&Death processors) and formal neurons. Both can be described as binary threshold automata. The organization of Birth&Death processors in hierarchical feedforward structures (trophic webs) is equivalent to the organization of multilayer feedforward networks of formal neurons. The error-backpropagation in such neural networks can be interpreted as equivalent to the feedback of downgraded energy in trophic webs. Assuming a correlation between feedback and the overall web performance in terms of global energy throughput, an ecosystem can be modeled as a feedforward artificial neural network with error backpropagation being trained through gradient descent. Hence all features of this class of ANNs concerning memory and learning can be applied to the trophic web of an ecosystem and to hierarchical natural and « artificial » energy transformation systems in general. Pareto-Zipf distributions are observed for feedback symptoms in all natural and « artificial » energy transformation hierarchies. Pareto-Zipf distributions are also observed for the backpropagated errors of formal neurons learning along gradient descent.

Key Words ***- energy transformation, information transformation / transmission, Birth&Death processors, threshold automata, Pareto-Zipf-Frechet-Weibull distributions, trophic webs, neural networks, memory, backpropagation, adaptation, learning, evolution***

1 Introduction: where to locate Intelligence?

1.1 The reductionist approach to intelligence

zooming-in to internal causes located in subsystems

According to the reductionist paradigm, causes are looked for by zooming into the subsystems of a system and its constitutive parts. Human intelligence is located in the brain and its functioning is explained in terms of the functioning of its constitutive parts, the neurons with their dendrites, cell-body, axons and synapses ...

The formal model of a neuron (a binary threshold automaton) and its aggregation to artificial neural networks has developed into an autonomous field of research. So far most practical applications of ANNs try to replace tasks which are normally accomplished by human intelligence (pattern recognition, image processing, language and speech processing, neural control systems ....).

Recently ANNs have also been used to replace scientific theoretical predictions. Gernoth and Clark [1995], for example, present multilayer feedforward network models, that predict empirical nuclear decay probabilities extremely well for the entire range of nuclei measured. So far no global nuclear physics theory exists, which produces results of equivalent quality.

However Gernoth and Clark use the ANN approach only as a mathematical black box mapping inputs (nuclear mass and charge) to outputs (nuclear alpha-, beta-, electron capture and spontaneous fission decay probabilities or stability). They do not see any «physical» analogy between the processes going on within a system of nucleons and the ANN model. ANN vector-mapping, like mathematical curve-fitting, may be a very powerful empirical data analysis and prediction tool; but so far it lacks « explanatory » power .

1.2 The holistic approach to Intelligence

zooming-out to external causes located in metasystems

According to the holistic approach (or systems approach), causes are looked for by zooming out analyzing the relations of systems amongst each other and understanding the control mechanisms of the metasystem and meta-metasystems etc.

In our view, the holistic approach does not replace the reductionist approach. It is only a complementary view, which allows to discover (uncover) relationships and causes impossible to detect with a reductionist approach only.

Applying the holistic approach to intelligence in nature, we propose in this article a conceptual framework, which locates intelligence not in microscopic cells or molecules, but in macrostructures of energy transformation processors and their hierarchical organization.

The ANN approach, as used in this paper, is completely freed from its original anthropomorphic brain and neural cell analogy. When we talk about ANNs, we consider only the abstract formal model of binary threshold automata, which are organized in a specific topology, which process data according to a set of well defined rules, and for which consecutive processing cycles are controlled by a formal strategy.

2 Structures of natural Intelligence : massive parallel energy transformation hierarchies, death processing, predator-pry relations

2.1 Local structure: from Birth&Death processors (energy transformation) to binary state transitions in formal Neurons (information processing).

In recent papers [Cempel, 1987, 1991, 1992; Winiwarter and Cempel, 1992] we have presented and elaborated a theory of generalized Energy&Information processors. The model is characterized by energy input, upgraded as well as degraded energy outputs and a limited internal transformation capacity (see fig. 1 below as an example). In addition to the traditional energy flows our model is based on two very simple postulates :

1) the internal accumulation potential is finite and irreversibly filled up to a threshold value.

2) the internal accumulation level regulates the internal accumulation rate through positive feedback (autocatalytic behavior of internal accumulation).

This very simple model results in important statistical features concerning the behavior of a single processor over its life-time and the statistical behavior of a population of similar processors.

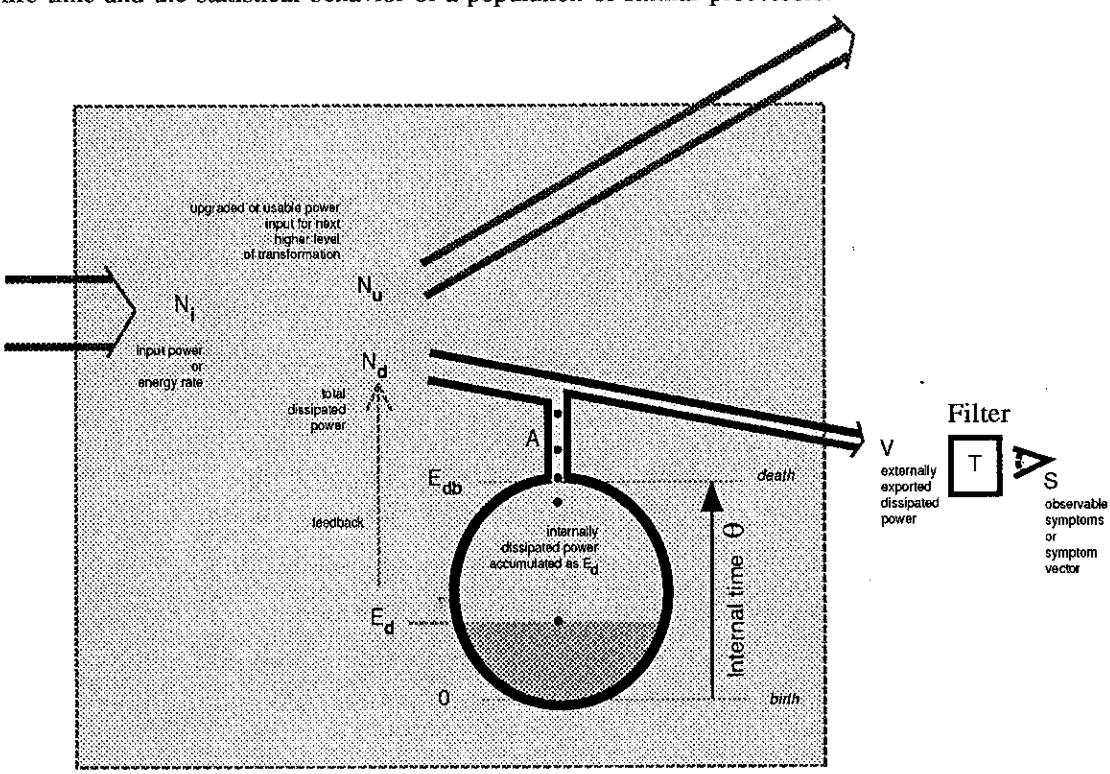

Figure 1. Elementary Birth&Death processor, the building block of the self-similar hierarchical Energy&Information transformation processing in operating systems.

Symptom statistics over the life-time of a single Birth&Death processor (temporal cross-section of an individual) :

Observing one processor over its entire life-time (from its birth at internal time $\theta = 0$ to its death at breakdown time $\theta = \theta_b$) we find its asymptotic behavior of transformed energy. Hence when the breakdown of the processor is approached, the transformed energy goes to infinity (see figure 2a).

However, in most cases, we can neither directly observe nor measure this transformed energy. We can only measure some covariant symptoms by means of some filter T as illustrated in figure 1.

If we observe a given symptom of one processor at regular time intervals over its life span we can make a cumulative probability of symptom occurrence. It can be shown analytically and experimentally [Cempel, 1992], that such a cumulative probability is of the Zipf-Pareto-Frechet-Weibull type (called ZPFW distribution hereafter) [Pareto, 1897; Winiwarter 1987; Cempel, 1992].

Symptom statistics for a population of Birth&Death processors (snapshot or spatial cross-section over a population of individuals at different life-advancement) :

Observing a population of energy transformation processors, which belong to the same operating meta-system and which are in different advancement of processor life, we can make a cumulative probability distribution of symptom occurrence within the population. It can be shown also [Cempel, 1992; Cempel and Natke, 1992], that the symptom distribution is again a long-tailed distribution of the ZPFW type as above.

Moreover, there is strict equivalence between the type of symptom life curve observed for a single processor and the observed symptom probability distribution for a population of processors. In the case of processors which are identical (or of similar « hardware technology » in a first approximation) the parameters defining the ZPFW distribution are identical in both domains; i.e. in the symptom life domain of a single processor and in the symptom population domain. Therefore we can make inferences about the evolution of a single processor based on parameters derived from measurements of a population of processors, and vice versa.

Co-transformation : the equivalence of Energy and Information

So far we have considered only the energy transformation aspect of Birth&Death processors during their lifetime (from birth to death). As illustrated in figure 2a) below, this generalized life process at a given level can be described by a continuos function with an asymptotic behavior (singularity) at the breakdown of the processor.

The same life-process observed by the metasystem can be described as a binary information process. As long as the lower level processor is in the operating state (*life*) the metasystem observes no change or 0-state. Only at the breakdown of the processor (it's *death*) the metasystems observes a binary state transition to 1-state (see figure 2b).

This double-sided or Janus-face like view, « seeing » one and the same process either as a continuos energy transformer or as a discrete binary information transmitter is the important idea put forward in this paper .

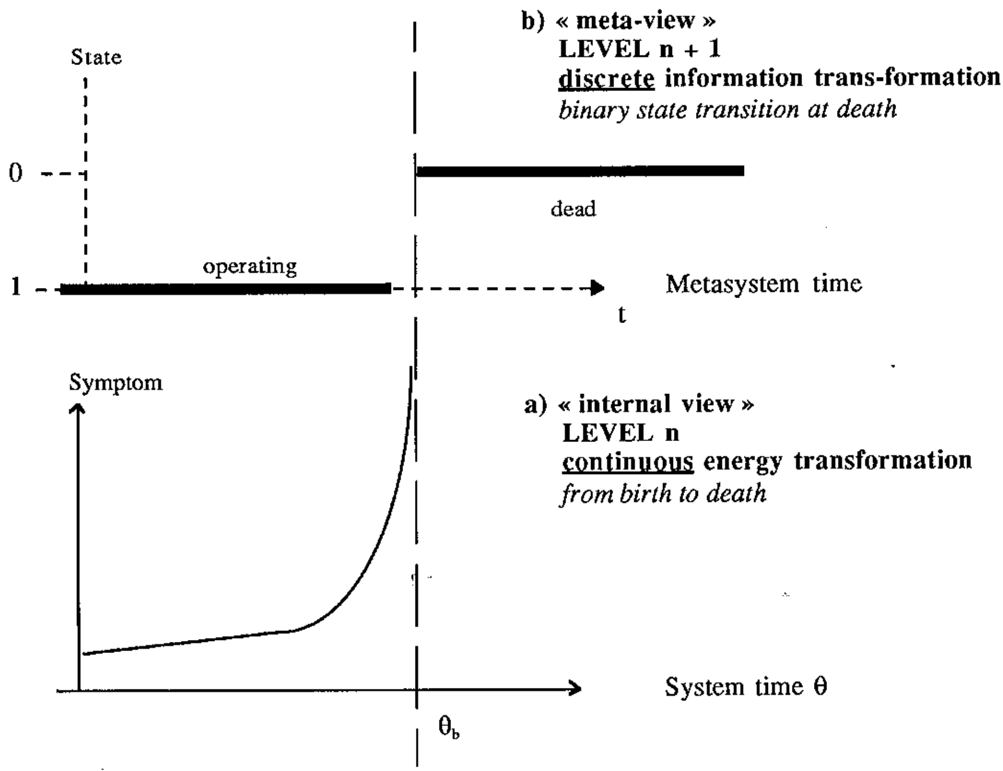

Figure 2. The equivalence of continuous energy processing and binary information transmission : continuos energy-transformation processes on one level are equivalent to state transitions or discrete binary information transmitted to the metalevel. (*death is a squashing function)*

Based on figure 2 we can draw an analogy between an energy transformation processor (Birth&Death processor) and a formal neuron.

Both can be considered as **threshold automata with binary state changes** as explained below :

- An energy transformation processor sums up internally stored transformation energy up to its internal capacity E_{db} (see figure 1). When this capacity is reached, (at breakdown-time θ_b for death processes) the processor breaks down, which is perceived by the metasystem as a binary state transition. Note, that the internally accumulated energy is a weighted sum of the input energy.

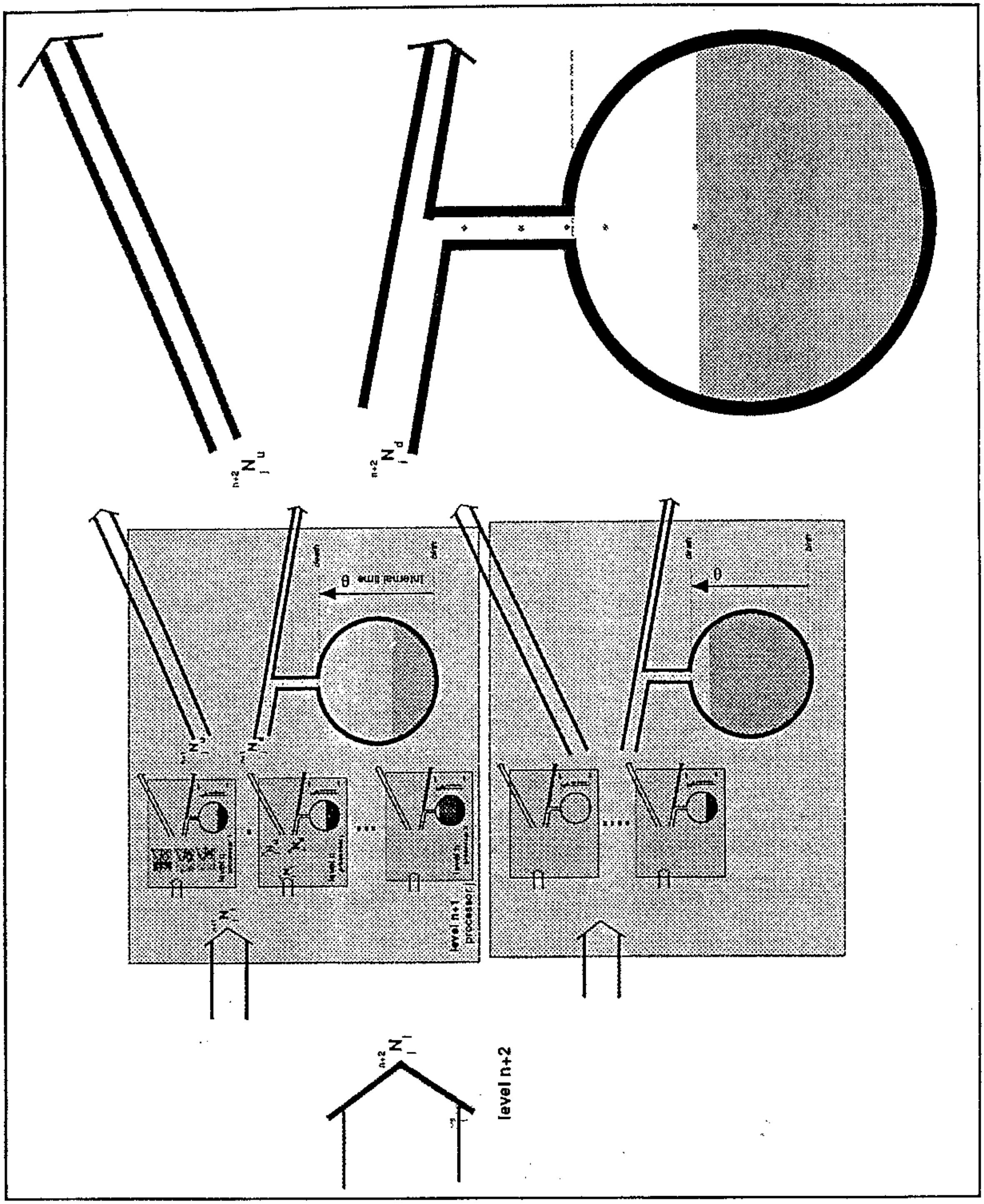

Figure 5. *Operating system* : a trophic hierarchy of Birth&Death processors. The self-similarity of the models permits to zoom in to lower levels and to zoom out to higher levels using the same conceptual and formal tools.

2.2 Global structure: from Trophic Webs (energy transformation hierarchies) to Multilayer Neural Nets (information transmission hierarchies)

In the elaboration of the concepts below, we limit ourselves to the modeling of biological ecosystems called trophic webs, since they are easy to conceptualize. Our arguments however are more general and can be applied to any hierarchically organized network of energy transformation (Birth&Death) processors.

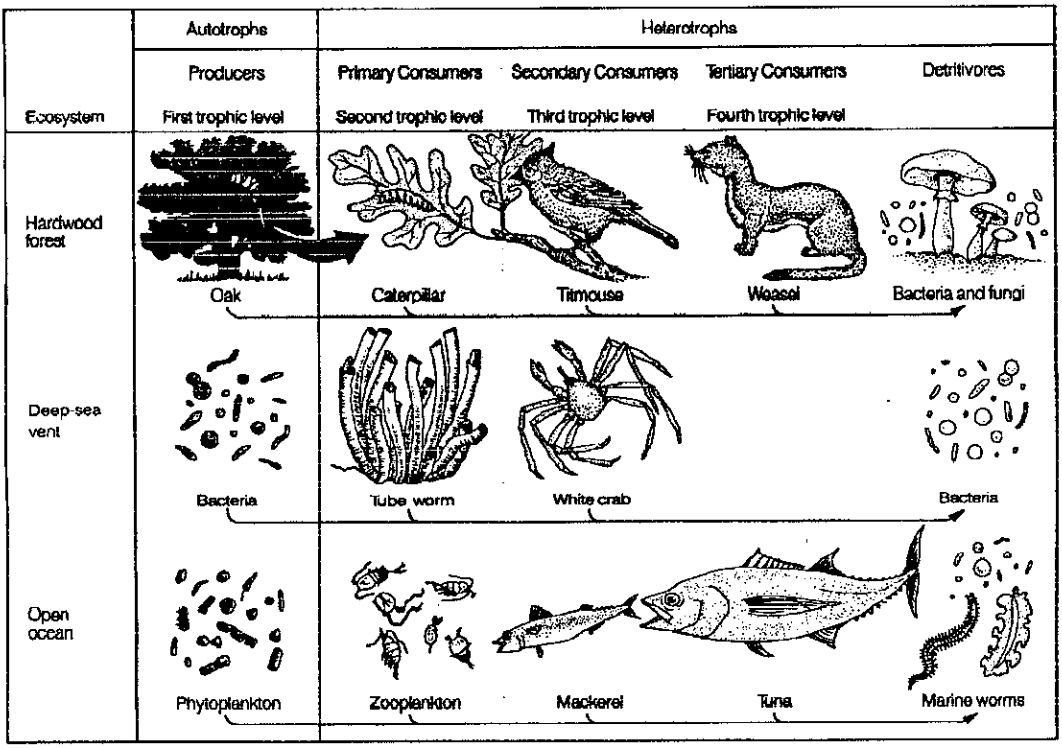

Figure 3. Trophic (Feeding) Levels and Food Chains: Producers and Consumer in Nature. Organisms vary from one ecosystem to another (shown here are those for forest, deep-sea vent, and open ocean), but in each ecosystem, autotrophic producers form the first trophic level, while consumers form the next three trophic levels, and detrivores utilize the wastes from all levels. After Postlethwait and Hopson [1992].

Virtually all recent attempts to model evolutionary systems and self-organization are based on hierarchical concepts [Miller, 1978; Jantsch, 1980; Auger 1986, 1990]. The approach of Auger is based on the concepts of states in the phase space of thermodynamical systems. Applying these concepts to multilevel systems demands to take bottom-up and top-down interlevel couplings into account. Auger's approach permits to generate dynamical models applicable to ecological, biological and economical systems. Despite the attractiveness of the model, the drawback of a purely thermodynamical approach is the lack of Birth&Death processes. The concepts of thermodynamics are based on ever-living entities (particles) - that never appear, nor disappear - moving around in system space.

It is the merit of Howard Odum [1988] to have focused our attention on the energy transformation aspect of hierarchically organized ecosystems. He already stated that " observing self-organization in nature suggests how energy is related to hierarchy and information ... The details of the energy transformation mechanisms are quite different in ecosystems, chemical reaction systems, turbulent hydrodynamical systems, social systems, and stars, but energy and mathematical characteristics are common to all".

According to Odum, the hypothesis of energy transformation in self-organizing open systems, quite beyond classical energetics, involve concepts of selfdevelopment such as systems learning and evolution in energy terms. "Characteristic ecosystem designs result from self-maximizing energy flow".

The typical system design is a hierarchical web of energy flows (pathways) and transformation processes (box units) represented in Figure 4. below. Energy flows are given on the pathways at steady state with inflows equal to outflows. At each transformation (box), most of the available energy is degraded and dispersed. A smaller amount of energy is upgraded energy to the right.

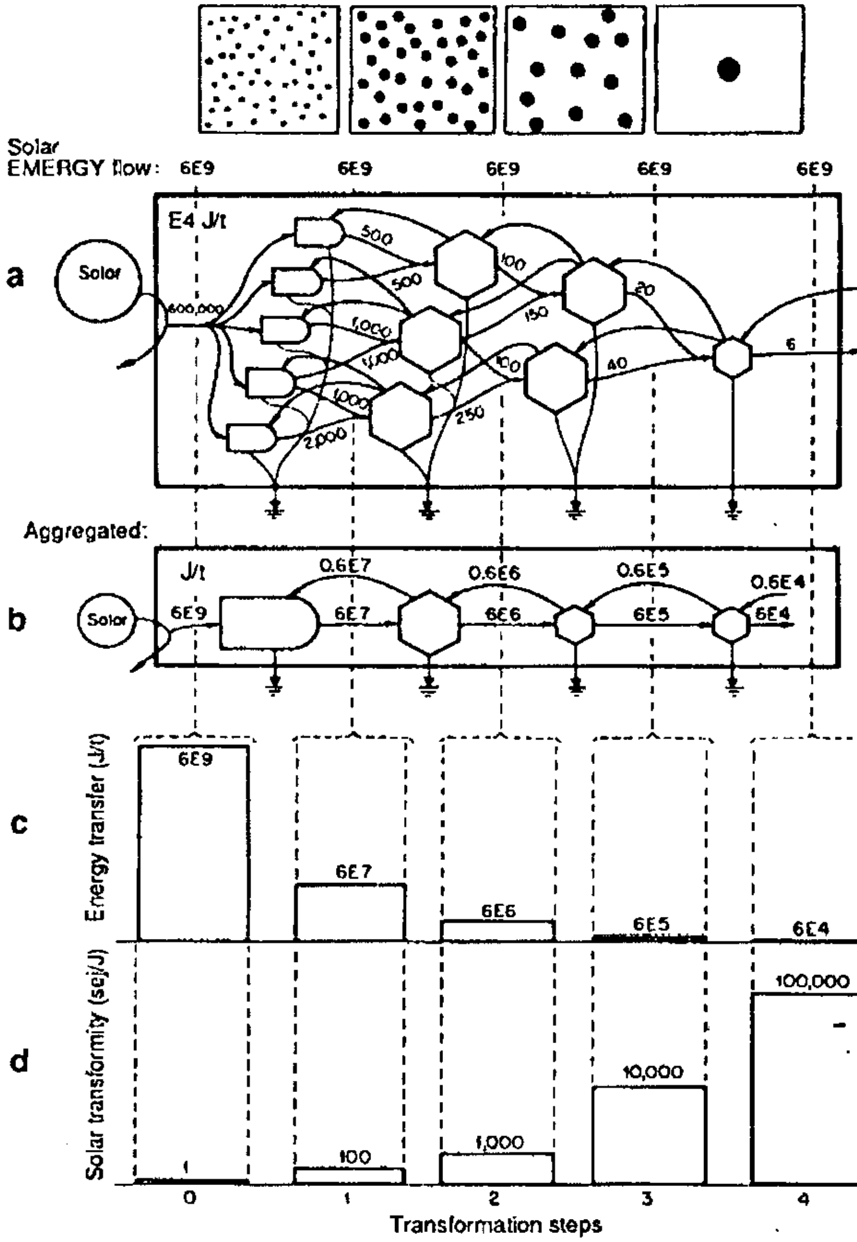

Figure 4. Energetics of an energy transformation hierarchy with feedback and recycle pathways omitted. Sketches show the distribution of size and territories of units in each category. (a) Web with energy flows indicated, (b) energy transformation chain formed by aggregating the web, (c) graph of energy flows at each stage in the energy hierarchy, and (d) solar transformities for each level in the hierarchy. After Odum [1988].

Odum already clearly speaks about control mechanisms through feedback : "As in a military hierarchy, services and controls are moving back down the hierarchy (to the left). The higher quality but smaller quantity energy types feed back as controls, reinforcing (amplifying) the production processes. For example, large animals control the more numerous smaller organisms by their behavior, placing of waste products, ways of eating, and control of pollination, seeding, and reproduction."

Compared to Odum's energy transformation hierarchy our proposed model of a trophic web as a hierarchy of Birth&Death processors is very similar (see Figure 5 below). The major difference between Odum's elementary energy transformation units and our elementary Birth&Death processors is the internally irreversibly accumulated dissipation energy part which finally leads to the break down or death of a processor. It is the death process which allows to describe a processor as binary threshold automaton.

Hence, based on the above argumentation that Birth&Death processors accumulate weighted energy inputs up to a threshold, a trophic web of energy transformation processors (with lower level processors feeding into higher levels) can be considered as a neural network.

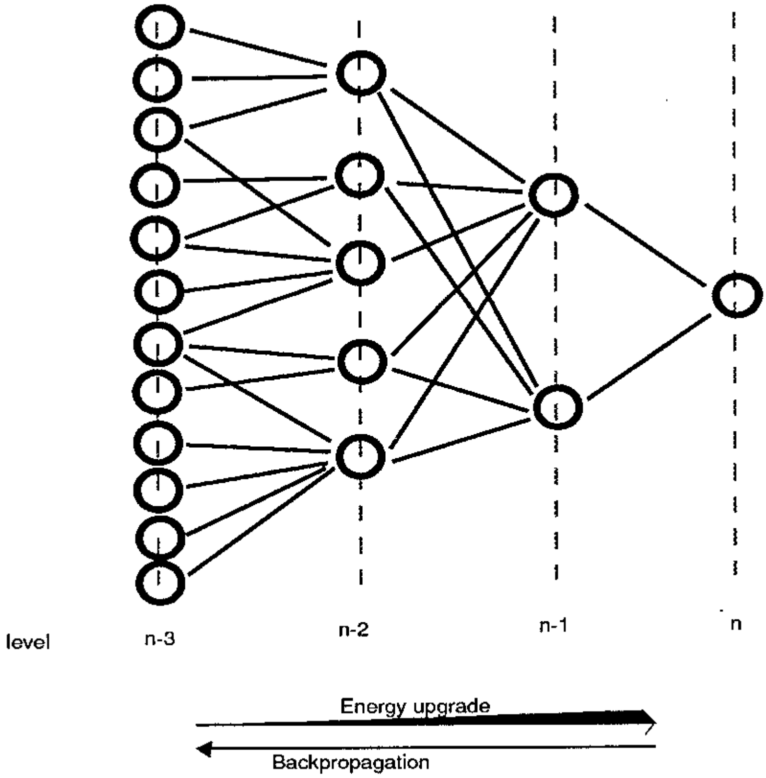

Figure. 5 Simplified representation of an operating energy transformation web formally equivalent to a feedforward neural network with backpropagation. Note the schematic nature of the graph. In ecological trophic webs up to 100,000,000 processors of a lower level feed into a single processor of an upper level.

As it was postulated in [Winiwarter, 1995b], we can interpret a node of the graph as energy processor : the link to the upper level is constituted via upgraded energy output of the processor, while the downgraded energy output from the processor to the web constitutes the feedback. On the other hand we can interpret a node of the graph as information processor and it can be considered as formal neuron : inputs are integrated over its life time and firing a binary signal to the next higher level occurs at death.

Hence the trophic web (ore any hierarchical energy transformation system), when looked at from an energy point of view, can be described as a trophic hierarchy of Birth&Death processors, and as a network of binary threshold automata (ANN), when looked at from an informational point of view.

3 Processes of natural Intelligence : weighted feedforward

3.1 Local process : from feedforward of (bio)-Mass to Weights in ANNs

Our analogy between Birth&Death processors and formal neurons is based on the concepts of weighted summation up to a threshold and its subsequent *squashing* to a binary value. There is no conceptual difficulty to accept death as a binary state transition from operation to non-operation. A swallowed pry just undergoes a binary state-transition from life to death as it becomes the input for the predator, the next level of energy transformation. The predator however does not simply add his inputs; he weights them. We put forward the simple hypothesis, that the neural input weight is equivalent to the physical weight or the bio-mass of the pry.

3.2 Global process: from (bio)-Mass-storage to Memory

Based on the neural network analogy, a web of Birth&Death processors should reveal memory features :

- In the case of a neural network, the memory consists of the threshold values, the input weights and the topology of active connections of all neurons in the network.
- In the case of a trophic web (or an Energy&Information transformation web in general), the memory consists of the bio-weights of the Birth&Death processors and the topology of active (trophic) connections within the web.

It is the bio-mass distribution of the massive parallel Birth&Death processors on each trophic level which constitutes the « knowledge » or memory of a trophic web of given topology at a given moment in time.

4 Controls of natural Intelligence : feedback training cycles and gradient descent

4.1 Local control: from Feedback of downgraded energy to error-Backpropagation *(from bullshit to green grass)*

Again, based on the above analogy, the learning / adaptation of energy transformation webs organized hierarchically, as illustrated in figure 5, could be modeled like the learning in feedforward neural networks with backpropagation.

From an energy point of view each processor transforms input energy. Upgraded energy is fed forward to the next higher level (from left to right). Downgraded energy is fed back to the lower levels (from right to left). This feedback (backpropagation) modifies also the processor parameters of a given level for the next operation/transformation cycle. Downgraded *bullshit* increases the bio-mass of the next generation of grass!

Learning and adaptation could be understood here as repeated processing cycles of the energy transformation web. At each trophic cycle the feedback modifies the parameters of the next generation of newly born processors in the web. When the feedback and the processor replacement policy is random, the web would show chaotic behavior. However, when the feedback is correlated to the overall performance of the web (e.g. the gradient descent method for error minimization), the web will reveal goal oriented behavior and learning.

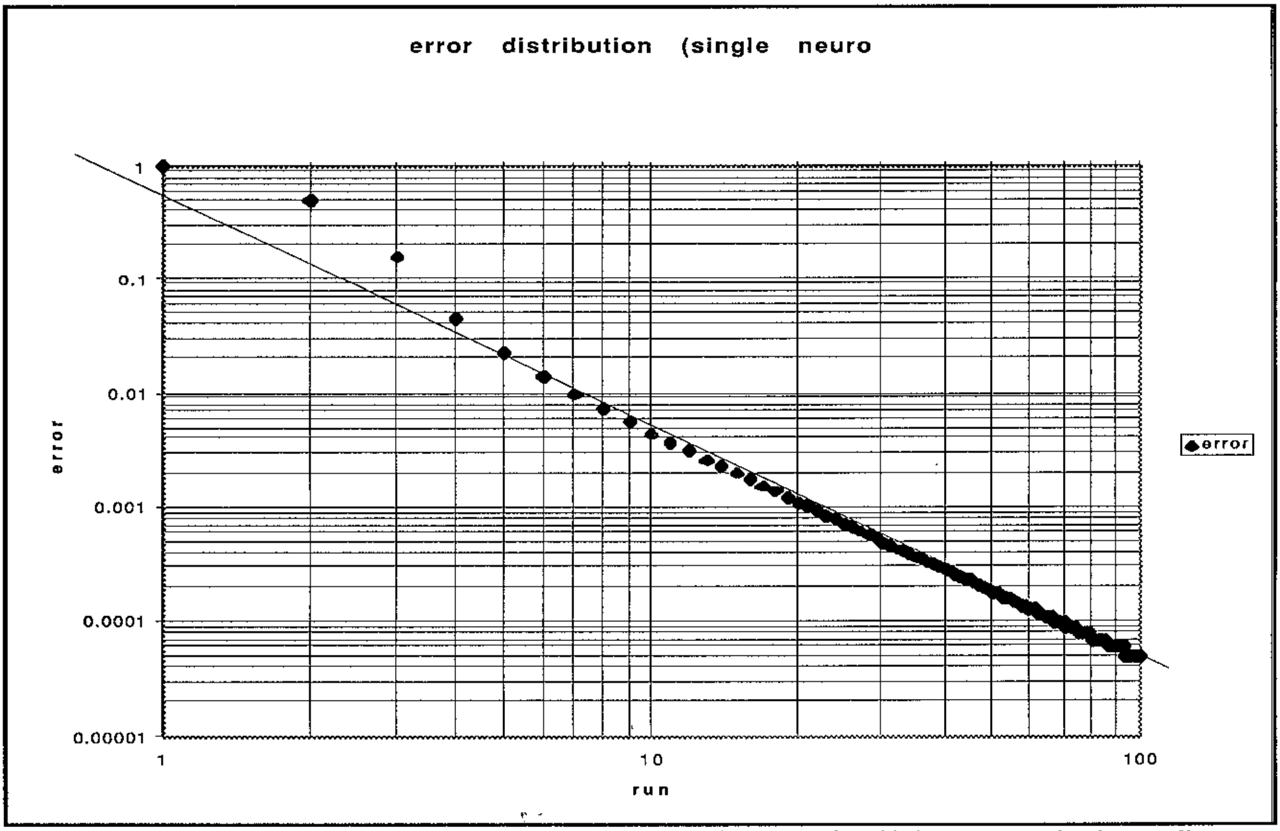

Figure 6. error-backpropagation during consecutive runs of a single neuron trained according to gradient descent. The errors follow a Pareto-Zipf distribution.
(data after McCord Nelson and Illingworth [1991])

When we analyze the error distribution of a single artificial neuron learning according to backpropagation and gradient descent, we observe a Pareto-Zipf distribution as illustrated above in figure 6.

4.2 Global control: From selective natural « fertilization » to Learning and optimization within a potential landscape

Like in the case of Birth&Death processors we can pass from the individual life-time domain to a global population domain and ask the question, what the error distribution would look like at a given moment in time for the population of neurons in a massively parallel processing network.

Together with B. Roehner [Roehner and Winiwarter, 1985] we have shown, that Pareto-Zipf distributions reveal a remarkable property, called stability under addition. This means that the aggregation of Pareto-Zipf distributions yields again a distribution of the Pareto-Zipf type. Hence a sample population of individual neurons following Pareto-Zipf statistics should reveal Pareto-Zipf behavior. I.e. massively parallel processing ANNs learning according to backpropagation trained according to gradient descent should reveal Pareto-Zipf statistics for backpropagated errors.

Now let us consider the observed symptom distributions for Energy&Information transformation systems. The observed symptoms are proportional to downgraded energy. Assuming, that downgraded energy has a positive feedback effect on the next processing cycle, we can interpret the observed symptoms as proportional to error backpropagation.

The fact, that we observe Pareto-Zipf distributions for symptoms on virtually all evolutionary Energy&Information transformation systems (see table 2. below), speaks in favor of our hypothesis of a common underlying process. This adaptive process can be modeled in terms of massive parallel processing multilayer feedforward neural networks learning through backpropagation and trained through gradient descent during consequent processing cycles.

5 Genealogical webs *(reproduction or genealogical hierarchies)*

5.1. Local genealogy: hardware, birth processing (generation), individual attractors

So far we have only considered the self-production aspect of a Birth&Death processor implying, that a single processor runs from birth to death with the processor technology (hardware) remaining unchanged over its entire lifetime.

Recall, that we have modeled death processes as binary threshold automata, triggered by the irreversibly internally accumulated downgraded dissipation energy within a finite potential.

(most human beings will protest to be reduced to binary death threshold automata)

Likewise we can model birth processes as binary threshold automata, triggered by irreversibly internally accumulated upgraded energy resulting in a binary state transition occurring at replication or birth.

(most women will protest to be reduced to binary birth threshold automata)

5.2 Global genealogy: processor design, parent-child generations (heredity), global fields

So far, from a global point of view, we have only considered energy transformation hierarchies at a given moment in time. For biological systems this corresponds to a spatial hierarchical decomposition in terms of *which processor feeds into which*, (or *who is eaten by whom*).

Collective processor memory and learning - given constant processor hardware - can be resumed under the term system **adaptation**. The proposed conceptual framework allows to model memory and adaptive system behavior for massive parallel processors organized in hierarchical feedforward levels undergoing production cycles. Salthe [1985] in his treatise on evolving (biological) hierarchical systems, calls this hierarchy the ecological hierarchy.

In evolving hierarchical systems however, processor technology changes over time (from generation to generation). In accordance with the biological terminology we call this adaptation of hardware technology **evolution**. Hence, in order to model evolution, we need to take into account the hardware changes of processors. These are related to the replication, reproduction or birth aspects of processors. Salthe [1985] calls this hierarchy the genealogical hierarchy. i.e. on each level of the hierarchy the parent child relationships from one generation of processors to the next generation of processors constitute the system's evolution in time.

	Ecological hierarchy		Genealogical hierarchy
--	---------------	7	total biosphere
6	surface of the earth	6	historical biota
5	bio-geographical regional	5	monophyletic lineage
4	ecosystem	4	deme (set of replicating individuals)
3	population	3	integrated genotype phenotype
2	organismic	2	gene
1	molecular	1	species

Table 1. Salthe's ecological and genealogical hierarchies of nature

The *monster*-abstraction, which reduces replication or birth to a binary state transition, allows us to model genetic webs similar to energy transformation webs: i.e. massive parallel processors organized in hierarchical levels undergoing reproduction cycles.

6 Outlook: towards a general theory of Birth&Death processors evolving in Energy&Information transformation systems

It would go beyond the scope of this paper, to develop a complete theory of evolution based on ANNs of Birth&Death processors. We only want to outline a conceptual framework guiding future research in areas of evolutionary systems for which we can observe Pareto-Zipf distributions.

the emergence of evolutionary Energy&Information transformation webs

In accordance with a recent paper [Winiwarter and Cempel, 1996] we propose a hierarchy of natural and « artificial » evolving systems as below. Here the ordering of the levels is based on the sequence of emergence of new types of Energy&Information transformation processes during the evolution of the universe. By new types of

Energy&Information we understand qualities, which cannot be explained in terms of the qualities of the lower levels

Table 2. below gives an overview of the potential application for our approach. Note that in the table both energy-transformation systems and information transformation systems are mixed. They should be separated like in Salthe's table above.

	Macroscopic evolutionary Systems	**Microscopic Energy/Information Processor type**	**Emerging Energy/Information Fields**	**empirically observed symptoms of energy transformation showing Pareto-Zipf (PZFW) regularities**
21	transdisciplinary systems science community	formal transdisciplinary (meta-scientist)	meta-theoretical fields	backpropagation error-sizes in ANNs
20	computer systems (Internet)	man-meta-machine processor (man-computer)	information technology	produced file-sizes
19	scientific community	formal written symbolic communicator (scientist)	scientific fields (theories, paradigms)	produced research papers
18	industrial production community	man-machine processor	technological fields (technology tradition)	produced dissipation energies
17	alphabetical community	written symbolic (literate) individual	written tradition, literature fields, history	produced word-class sizes
16	economic community	economic symbolic (money) trading individual	economic tradition / trade fields	produced firm-sizes, incomes, profits
15	language community	formal symbolic (verbal) communicating individual	oral tradition, religious fields	produced city sizes
14	agricultural community	plant and animal manipulating individual	fields of agri-culture	produced field sizes
13	hunting community	tool manipulating individual	hunting tradition / culture	
12	fire-place community	fire manipulating individual	cooking tradition (oikos)	produced agglomerations
11	ethological community	symbolic communication individual	social fields / ethological	produced population aggregates
10	eco-system	multicellular organism	central nervous (feeling)	produced population sizes
9	biotope	cell (eukariote)	genetic/sexual fields	produced species
8	bio-sphere/Gaia	self-replicating metabolic unicellular (prokaryote)	trophic (bio-energy)	produced genera
7	atmosphere oceans	macro-molecule	bio-chemical	produced micro-particle sizes
6	earth-crust, tectonic plates	micro-structure	geo-mechanical	produced earthquakes
5	planets	physico-chemical element	geo-thermal	produced chemical compounds
4	galaxy, stars	nucleus	electromagnetic + weak nuclear	produced nuclei
3	proto-galaxy	hadron	strong nuclear + electro-weak	
2	proto-universe	quark	gravitation + grand unified	
1	Planck-era of big bang	graviton	super-grand-unified	

Table 2. Generalized Life symptoms : the emergence of energy/information systems, processors, associated fields of interaction and observed Pareto-Zipf (PZFW) regularities. For a historical overview and detailed sources see [Winiwarter and Cempel, 1992 and 1996]. According to the proposed conceptual framework, each level in the table can be modeled as a multilayer feedforward ANN with backpropagation learning/evolving according to gradient descent.

In the hierarchy, each level is characterized by a macroscopic Energy&Information system composed of microscopic Birth&Death processors which interact under the influence of their own characteristic field, which is generated by the totality of all the processors interacting within the system. For example the bio-chemical field consists of and regulates the interaction of microscopic molecules in the macroscopic ocean as a system.

Intralevel coupling and transformation :

Climbing up the hierarchy we observe diminishing coupling/transformation energies between processors on each consecutive higher level. e.g. binding energies between nucleons are orders of magnitudes higher than binding energies between chemical elements etc.

Interlevel coupling and transformation :

a) a higher level can only emerge and evolve within all previous lower levels; usually when the field in the lower level exceeds a certain threshold value (bifurcation, symmetry break).

b) lower levels can destroy higher levels, but higher levels can only constrain lower levels. (see Cempel's energy hierarchy).

c) lower level processors feed into (or are eaten by) higher level processors.

As in the case of energy transformation processors, where we pushed the abstraction of the « *life* » process to an extreme, reducing the self-production or «life » of a single processor to a binary threshold automaton (death-processor). This enabled us to see collective energy transformation behavior in a new light, allowing for the first time to model global natural system's memory and learning in terms of ANNs.

In terms of energy transformation webs, our conceptual framework can be applied to economic systems (production threshold automata feeding forward finished goods to upper levels and feeding back money to lower levels etc.)

In a similar way we can push the abstraction to an extreme and consider the processes of (self)-reproduction in terms of binary threshold automata (birth processors) organized in a multilayer feedforward hierarchy of information transformation. 0/1 reproduction automata feed the next generation forward to the upper level in the genealogical hierarchy and the global « temperature » of the upper level, feeds back and controls the processing of the next reproduction cycle on the lower level in the genealogical hierarchy.

It is the long term temperature of the biosphere, which « drives » genetic evolution. General « cooling » of the web drives it towards minima within a potential landscape[Winiwarter, 1995b]. Short local heating, followed by subsequent cooling can avoid trapping of the web in local minima of suboptimization.

7 Conclusion

1. All observable evolutionary systems can be described as hierarchical **Energy&Information transformation** webs .
2. On each level, energy transformation processors **feed up**graded energy into higher levels and **feed** downgraded energy **back** to lower levels. This holds for natural evolutionary systems (astrophysical, geochemical, ecosystems ...) but also for « artificial or man-made » evolutionary systems (city systems, economic systems ...).
3. Energy transformation processors in all observable evolutionary systems show birth and death processes . We presented a simple model for an energy transformation processor, which is general enough to be applicable to any type of energy transformation. A **Birth&Death processor** is characterized by a limited transformation capacity in which dissipation energy is irreversibly internally accumulated until a threshold capacity is reached . (*natural death*).
4. Birth&Death processors can formally be described from two points of view :
   - from an energy-transformation point of view, a Birth&Death processor continuously transforms and accumulates energy up to a threshold (breakdown at death or replication at birth)
   - from an information-transformation point of view however , as observed by the metasystem of the processor, a Birth&Death processor is equivalent to a **binary threshold automaton** or formal neuron with the two possible states : e.g. in the energy processing hierarchy the two states are 0-state (operation or silence) or 1-state (firing at death).
5. Birth&Death processors organized in an hierarchical transformation system (**trophic web**) are as a consequence formally equivalent to a **neural network** of the feed forward type. Lower level processors feed into higher level processors. Hence all the features of neural networks, like memory, adaptation/learning and optimization can be looked at in an analog way in trophic webs of energy/information transformation processors
6. A trophic web has **memory**; it's memory is non-localized, embedded in the web-topology and the ensemble of weights (bio-masses) of the web's processors .
7. **Learning** (adaptation) of the network takes place in the form of backpropagation. The downgraded energy of each processor is re-cycled influencing the processor parameters in the next processing cycle. Training of the web functions according to gradient descent through **positive** (or negative) **feed-back** from higher level processors to lower level processors in both senses of the word (energetic feed-back and cybernetic feedback).
8. The **error-distribution** of a single neuron trained with **backpropagation according to gradient descent** follows a **Pareto-Zipf** distribution.
9. Pareto-Zipf distributions are **stable under addition**, hence we should observe error-distributions of this type also for massively parallel processor networks (like trophic webs) trained with backpropagation according to gradient descent.
10. Longtailed distributions of the Pareto-Zipf-Frechet-Weibull, called **generalized life-symptoms**, are empirically observed for (backpropagation) symptoms of virtually all known types of Energy&Information transformation systems.
11. The above approach may be the basis for a general theory explaining self-organization, self-learning and evolution in nature

8 References

[Auger, 1986] P. Auger, Dynamics of Hierarchically organized Systems : A General Model Applied to Ecology, Biology and Economics, *Systems Research* 3:41-50, 1986.

[Auger, 1990] P. Auger, Self-organization in Hierarchically Organized Systems, *Systems Research* **7**: 221-236, 1990.

[Cempel, 1987] C. Cempel, The wear process and the machinery vibration - a TriboVibroAcustical model, *Bull. Polish Academy of Sciences*, Techn. Sc., **35**, N(7-8): 347-363, 1987.

[Cempel, 1991] C. Cempel, Condition evolution of machines and its assessment from passive diagnostic experiment. *Mechanical Systems and Signal Processing*, **5**, 1991.

[Cempel, 1992] C. Cempel, Damage Initiation and Evolution in operating Mechanical Systems. *Bull. Polish Academy of Sciences*, **40** (3): 1-14, 1992.

[Cempel and Natke, 1992] C. Cempel and H.G.Natke, Damage Evolution and Diagnosis in Operating Systems. In H.G.Natke et al. (ed.), *Proceedings of Safety Evaluation Based System Identification Approaches Related to Time-Variant and Nonlinear Structures*, pp.44-61, Vieweg, Braunschweig, 1992.

[Gernoth and Clark, 1995] K.Gernoth and J.W.Clark, Neural Networks That Learn to Predict Probabilities: Global Models of Nuclear Stability and Decay. *Neural Networks* **8** (2): 291-311, 1995.

[Jantsch, 1980] E. Jantsch, *The Self-organizing Universe*, p.224-240, Pergamon Press, Oxford, 1980.

[McCord Nelson and Illingworth, 1991] M. McCord Nelson and W.T.Illingworth, *A Practical Guide to Neural Nets*, pp.324-327, Addisson-Wesley, Reading Mass., 1991.

[Miller, 1978] J.G. Miller, *Living Systems*, p.1028-1029, McGraw Hill, New York, 1978.

[Odum, 1988] H. Odum, Self-Organization, Transformity, and Information, *Science* **242**: 1132-1139, 1988.

[Pareto, 1897] V. Pareto, *Cours d'économie politique*. (1897) Reprinted as a volume of Oeuvres Complètes, Droz, Geneva, 1965.

[Postlethwait and Hopson, 1992] J.H.Postlethwait and J.L.Hopson, *The Nature of Life*, p. 732, McGraw Hill, New York, 1992.

[Roehner and Winiwarter, 1983] B.Roehner and P.Winiwarter, Aggregation of independent Paretian random variables, *Advances in Applied Probability*, **17** : 465-469, 1983.

[Salthe, 1985] S.N. Salthe, *Evolving Hierarchical Systems*, p.185, Columbia University Press, New York, 1985.

[Winiwarter, 1983] P. Winiwarter, The Genesis Model. Part II: Frequency distributions of elements in self-organized systems. *Spec. Science and Technology* **6** :103-112, 1983.

[Winiwarter, 1985] P. Winiwarter, Iso-dynamics of population-size distributions in hierarchical systems. In B.Banathy (ed.) *System Inquiring, Proceedings of the Society for General Systems Research*, vol.I pp.87-95, Intersystems Publications, Seaside Calif., 1985.

[Winiwarter, 1987] P. Winiwarter, Size-distributions and hierarchical control in biological.systems. In B.Lück (ed.), *Biologie Théorique*, pp. 289-299, Editions du CNRS, Paris, 1987.

[Winiwarter and Cempel, 1992] P. Winiwarter and C. Cempel, Life Symptoms : the behavior of open systems with limited energy dissipation capacity and evolution. *Systems Research*, **9** (4): 9-34, 1992.

[Winiwarter, 1995a] P. Winiwarter and C.Cempel, Birth&Death Processors 1 : time-energy equivalence in System Life and the fractal dimension of System Time in energy-transforming Hierarchies. In B.Bergvall-Dareborn (ed.) *Proceedings of the International Society for the Systems Science*, pp.73-85, International Society for the Systems Sciences, Louisville, Kentucky, 1995.

[Winiwarter, 1995b] P. Winiwarter, Birth&Death Processors 2 : modeling System memory, learning and evolution based on the formal equivalence of energy transforming hierarchies and neural networks. In B.Bergvall-Dareborn (ed.) *Proceedings of the International Society for the Systems Science*, pp.86-100, International Society for the Systems Sciences, Louisville, Kentucky, 1995.

[Winiwarter and Cempel, 1996] P.Winiwarter and C.Cempel, Evolutionary Hierarchies of Energy processing in Nature, to be published in R.Trappl (ed.) Proceedings of EMCSR'96 conference, Vienna, April 1996.

[Zipf, 1948] G.K. Zipf, *Human Behavior and the Principle of least Effort*, Addison Wesley, Cambridge Mass., 1948.