Information and Entropy in Cybernetic Systems

Heikki Hyötyniemi

Cybernetics Group
Helsinki University of Technology, Control Engineering Laboratory
P.O. Box 5500, FIN-02015 HUT, Finland
`heikki.hyotyniemi@hut.fi`
`http://www.control.hut.fi/hyotyniemi`

Abstract. It has been shown that the cybernetic approaches can efficiently be used for analysis and design of complex networked systems. Still, the earlier discussions were bound to the actual application domain at hand. This paper gives more intuition in what truly takes place in a cybernetic system from another point of view. Information theory, and specially the concept of *entropy,* offer a yet more general perspective to such analyses.

1 Introduction

There are many approaches to address the problems of real-life complex systems, each of them concentrating on some specific issues and more or less ignoring the others. The approach that is introduced in [19], or *neocybernetics,* is a new theoretical framework that differs from the existing ones, compensating for the shortcomings:

– In *complexity theory* one studies structurally complex nonlinear functions as independent entities; now the emphasis is on structurally simple large-scale systems where complexity is caused by high-dimensionality and system-wide interactions.
– In *system theory* (or *general system theory*) discussions also embrace the whole system, but they are limited to abstract, holistic studies; now, on the other hand, concrete down-to-earth analyses give substance and semantics to the discussions.
– In *control theory,* also being a branch of system theory, studies are similarly oriented on concrete rather than abstract systems; however, as compared to the current approach, there the emphasis is solely on centralized rather than distributed control structures.
– In *traditional cybernetics* (being rediscovered various times under different names like *autopoiesis* and *synergetics*) the decentralized structures also communicate with each other; however, the studies are stuck with the mechanisms, and the issues of *emergence* are not addressed. The assumption now is that without explicit emphasis on emergence it is impossible to understand the overall behaviors — the actual essence of cybernetic systems.

In *neocybernetics,* the emphasis is on the emergent *pattern* rather than on the individual interaction processes that finally lead to that emergent pattern (see [19]). This emergent pattern is the (hypothetical) *dynamic equilibrium* after all transients in the system have decayed. Stabilization of the system, or finding the balance among tensions, is the ultimate goal of the self-regulatory structures that are based on interactions and (negative) feedbacks within the system. Indeed, these balances (or *higher-order balances*) are the most characteristic feature that is common to very different cybernetic systems. The balances can be characterized as minima of some cost criteria; imbalance means tensions or *free energy* still remaining in the system.

The discussions in [19] where this kind of concepts were introduced were rather heuristic and high-spirited. To keep things in control, a strong conceptual structure is needed; such framework will be studied in this paper.

Wittgenstein observed that a language is needed for expressing ideas: Without strong concepts and grammar structures not everything can be said, and discussions remain vague ands descriptive. There are differences between languages; whereas human languages are practical in everyday life, it is mathematics that is the natural language of nature. Mathematics is pure syntax, but when the syntactic structures and manipulations are applied to semantic entities, new interpretations can be made ...

> *What kind of conceptual structures can be constructed when the cybernetic concepts are "thought of" applying the language of mathematics?*

When deriving the cybernetic models, it turned out that down-to-earth analyses brought necessary substance and semantics to discussions. And, again, it turns out that when the nonidealities of natural domains are actively taken into account, new intuitions can be gained. Counterintuitively, applying more and more specific toolboxes, more and more general conclusions can be drawn. Earlier, the power of system theory as a source of conceptual tools was praised — now the role of *control engineering* is emphasized as a source of intuition. It turns out that without understanding the ideas of *model-based control* and *filtering of signals* the behaviors in cybernetic systems cannot truly be mastered.

This paper discusses the relations between the information theoretic concepts and cybernetic models; and, truly, it seems that some deep paradoxes can be seen from fresh points of view. For example, why does it seem that the complex systems seem to become more and more complex? Why do there exist systems that seem to defy the arrow of entropy? What are the mechanisms for complexity cumulation?

In the Middle Ages, it was thought that there needs not exist common principles in nature: The *sublunar* and *translunar* phenomena, for example, followed their own laws. The breakthrough in physics came when it was observed that the planetary motions were also governed by earthly, not by divine principles. Similarly, today it seems that the systems are divided into two classes: Normally, the systems follow the thermodynamic principles; but then there exist systems where these principles seem not to hold. It seems that in the framework of cybernetic systems, these dilemmas can be studied in a coherent way. To put it

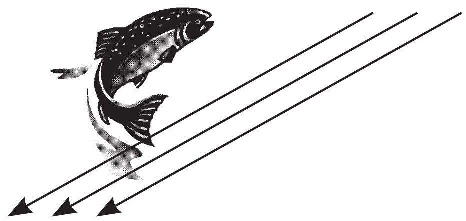

Fig. 1. Cybernetic systems: Framework for systems fighting against the flow of entropy

informally: The stronger the flow of entropy is, the more probably there also exist countercurrents, or “whirls”, in that flow (see Fig. 1). It turns out that *cybernetic systems are systems where the arrow of entropy is inverted.*

2 From physical constraints ...

Before concentrating on the highly abstract analyses concerning information and entropy, it needs to be noted that there is always a physical nonideal system beyond the abstractions. Here, such nonidealities are first briefly studied — and, counterintuitively, it turns out that these concrete analyses give the concrete means to carry out the highly abstract and general discussions.

2.1 Modeling of complex systems

The engineering-like approach to understanding any real-life problems is *modeling,* constructing simplified, abstracted representations capturing the relevant system properties in a compressed form. The key question here is whether a complex system *can* be simplified without missing its essence: If the whole is more than parts, reductionistic methodologies collapse.

This intuition seems to be the mainstream attitude among complex systems theorists — meaning that there is scepticism against the potential of traditional modeling and mathematics in general. For example, in [13] the starting point is that whereas simple systems are modellable, complex systems *by definition* defy modeling attempts: The possibility of “complexity”, in an intuitive sense, only arises when a system acts in unexpected ways; that is, in ways that do not match the predictions of models. Another formulation for this thought is given in [18], where it is claimed that all descriptions of complex behavior necessarily are still more complex than the original behavior: The simplest representation of a system is the system itself, and no models can be constructed. Such pessimism has resulted in predictions concerning End of Science [5].

However, the end of science has been prophesized many times in the past: Always when old approaches have been exhausted, there has been scepticism before new paradigms are found. The huge successes of the scientific paradigm, and the ever regenerative power of mathematical tools, motivates the *Pallas Athene Hypothesis,* in the spirit of the Gaia Hypothesis: Whereas Gaia supports the life processes on the Earth, Pallas Athene supports scientific work. Indeed, more than being a matter of fact, this is a confession of faith. But this optimistic attitude pays back: There is still plenty to study, and it is the Old Science that suffices when studying complex systems.

To start with, Wolfram's arguments [18] are indeed undeniable: If the modeling framework is too powerful, the mathematical analysis framework collapses. Even if studying more and more complex systems, the model structures also must *not* become more complex beyond a certain limit. How can this incompatibility be resolved? When the belief of not-yet-exhausted power of mathematical modeling is adopted as a starting point, it is easier to redirect analysis efforts. The only way to avoid the deadlock is to take another route: When the approach starting from chaos theoretical problem setting turns out to lead nowhere, the opposite way to proceed is to study *emergent behaviors.* Rather than looking at the *process,* consisting of nonlinear iterations, one can look at the *pattern,* the final outcome of the iterations when the steady state has been reached. One does not need to follow the individual behaviors if one knows the goal where the system is going towards; in mathematical terms, this final goal can be characterized in terms of cost criteria that are being minimized (maximized) in an optimization process. The dynamic processes are caused by tensions in the underlying pattern that has not yet converged. Whereas the formulations based on cost criteria are often mathematical, with no physical relevance, in the case of cybernetic models the cost criteria have direct interpretations, implementing a connection between abstract and concrete, between local agents and global systems.

Perhaps because of the marvellous successes of computer technology, the process view of thinking about systems dominates today. For example, in the contemporary *agent* frameworks, the agents are software constructs. The problem here is that the topological structure beyond the algorithms is too vague to offer any added value beyond the explicitly implemented functionalities. It is important to observe that the algorithmic approach is not the only alternative, at least not in all cases.

These discussions can be extended also to more abstract cases. There are many domains where the shift from immediately-observable processes to fundamentally-existent patterns has not yet taken place, and where it is difficult to see what this transition would mean in the first place. It is typical that modeling is first implemented by mimicking the observed surface-level behaviors or intuitions: For example, the self-organizing Kohonen SOM algorithm was first defined as a procedural algorithm; only later, it was observed that this algorithm is just a way (gradient descent algtorithm) for minimizing a special energy function, and new possibilities for further analysis and algorithm development opened up [6]. It is easy to see developments in retrospect — but, for example, there may be possibilities of seeing also the processes of *developmental biology,* or in *biological evolution,* in the cybernetic pattern perspective. Such issues will be later elaborated on.

2.2 Role of individuals

When the cybernetic models were derived (see [19]), the discussions were carried out in a purely information theoretic setting. It was assumed that information can be transferred in a causally unidirectional manner, without "output" affecting the "input" coming from outside the cybernetic system. However, when the information carriers are populations, for example, and information transfer takes place in the form of lower level prey being eaten by higher-level predators, this idealized view necessarily collapses. There is no such thing as information without some physical carrier, and in some environments this fact becomes specially acute.

The cybernetic models were derived by closing all loops inside the system; now the final loops are closed, coupling the system output with the system input.

First, study an ecosystem where the species on a specific trophic level compete for resources on the previous level. As presented in [19], the cybernetic model becomes

$$\frac{d\,x}{d\,t} = -A\,x + B\,u. \tag{1}$$

Here, x is the vector of relative population activities (not simply the number of individuals, or biomass; how the actual biomass is distributed, and what are the losses, is not concentrated on here), and u is the vector of resources coming from outside. The matrix A contains the interaction factors among the same-level actors, revealing the patterns of competition, and B contains the "forage profiles", revealing how the predators exploit the resources. The model can be interpreted so that some kind of "life force", or *elan vital,* emanating from the environment and cumulating in the populations. Whereas this "dissipative flow" never ceases, a cybernetic equilibrium among populations is found when the opposing forces are in balance.

The model (1) is generic, and not very much can be said about it in general terms. Assuming that the system is truly cybernetic (see [7]), there holds

$$A = \rho\tau \mathrm{E}\{\bar{x}\bar{x}^T\}, \quad \text{and} \quad B = \rho\tau \mathrm{E}\{\bar{x}u^T\}, \tag{2}$$

where ρ and τ are adaptation parameters, and $\bar{x}$ denotes the steady-state values of x corresponding to the "almost constant" u:

$$\bar{x} = \phi^T\,u = \mathrm{E}\{\bar{x}\bar{x}^T\}^{-1}\mathrm{E}\{\bar{x}u^T\}\,u. \tag{3}$$

Assuming that the dynamics of x is truly much faster than that of u, and if the behavior of u is smooth, one can substitute $\bar{x}$ with x in the formulas.

Whereas $x = \phi^T u$ reveals the population balance, that is, the distribution among population activities in steady state as a whole, in reality there is each day a new competition among the individual actors for the available resources. To make the rather abstract model (1) more concrete, this everyday struggle will now be modeled. It is these everyday actions that also affect the available resources, thus constituting the feedback structure between the system and its environment.

Whereas there is the *elan vital* coming from the environment, the inverse effect in the feedback loop caused by the individuals exhausting the resources could be called *elan letal*. An individual tries to take the resources it needs, pushing competitors away. As it is the individuals that are the agents for the inverse action, first study a single individual and its effect in the environment. Define the vector v that reveals the "effective activity" of a population; the contribution of a single individual i to the whole grid of population activities can be expressed as

$$\left(\frac{d\,v}{d\,\tau}\right)_i = -A\,v + \mathbf{1}_i. \tag{4}$$

The vector of *elan letal* being caused by the single individual is here defined as

$$\mathbf{1}_i = (\ \underbrace{0\cdots0}_{i-1}\ \ 1\ \ \underbrace{0\cdots0}_{n-i}\ )^T. \tag{5}$$

It needs to be recognized here that there can be numerical parameters adjusting the overall behavior; it can still be assumed that qualitatively the effects remain essentially intact. The individuals belonging to a certain population affect other populations as revealed by the "competition matrix" A. The time axes differ in different submodels — the dynamics of a single individual is fast (rate of change in variables being typically of order from hours to days), the dynamics of a population is slower (from days to weeks), and the dynamics of the environment is still slower (from weeks to months). The speeds of individual dynamics can be emulated in different ways; here it is assumed that all integrators operate on different time scales.

Again, as seen in the wider perspective, when elan letal cumulates, a cybernetic equilibrium is eventually found, and the effective activities reach a balance:

$$(\bar{v})_i = \mathrm{E}\{xx^T\}^{-1}\,\mathbf{1}_i, \tag{6}$$

so that the effect of an individual on the resources can be expressed as

$$(-\Delta U)_i = B^T\bar{v} = \mathrm{E}\{xu^T\}^T\mathrm{E}\{xx^T\}^{-1}\,\mathbf{1}_i, \tag{7}$$

where the forage profiles in B are utilized. Here the "$-$" sign is explicitly included to emphasize that the resources are exhausted (however, the consumption of some resource type can also be negative, meaning that this resource is *saved* — this is caused by the screening effects among individuals).

Because of system linearity, the effects of one individual can be freely scaled; the overall effect of all individuals in the ecosystem can be calculated as a weighted sum:

$$\begin{aligned} -\Delta u &= \sum_{i=1}^{n} x_i\,(-\Delta U)_i = \sum_{i=1}^{n} \mathrm{E}\{xu^T\}^T\mathrm{E}\{xx^T\}^{-1}\,x_i\mathbf{1}_i \\ &= \mathrm{E}\{xu^T\}^T\mathrm{E}\{xx^T\}^{-1}x. \end{aligned} \tag{8}$$

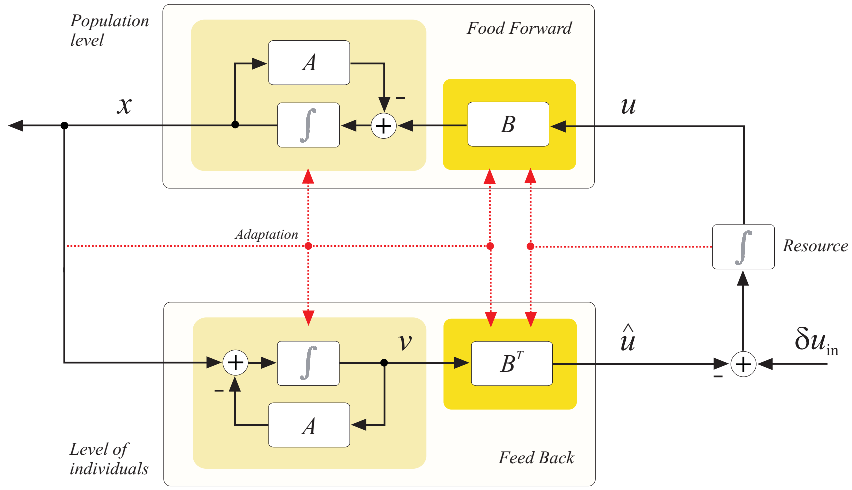

Fig. 2. Action (upper part) and reaction (lower part) automatically implemented by a cybernetic system. The structure looks strangely symmetric — remember Heraclitus: "The way up and the way down are the same"

This is how much of the resources are being used by the whole system, revealing the rate of exhaustion. Assuming that the pool of resources is large, it will not be exhausted immediately; instead, a dynamic structure is instantiated:

$$\frac{du}{d\tau} = -\Delta u + \delta u. \tag{9}$$

where $\delta u(t)$ is the natural rate of growth in the resources; the time scale, or how τ is related to real time t, is dependent of the system. The overall system structure can be expressed as shown in Fig. 2. Assuming that the inner dynamics are much faster, the outermost loop is always stable (phase shift being only 90 degrees).

The mathematical formulation of the feedback structure (8) looks strangely familiar. Indeed, it equals the *multilinear least-squares regression model* between x and u, that is, whatever x is, the best (linear) mapping from it to output (in terms of minimized error variance) is obtained by the MLR formula

$$\hat{u} = \mathrm{E}\{xu^T\}^T \mathrm{E}\{xx^T\}^{-1}\, x. \tag{10}$$

If ϕ spans the principal subspace, as has been shown in [19], mapping from the corresponding x to output applying MLR implements *principal component regression* (see [8]). It has been observed that PCR is an efficient way to reach robust regression structures, filtering out noise and exploiting the correlations in data optimally.

Finally, note that if the natural reproduction δu_{in} in (9) is constant, and if the resource variations have been appropriately captured by the system, a balance is found where $\hat{u}$ compensates the growth in u; the system remains in balance where $\hat{u}$ equals u. If there is variation in the growth rates, the principal component regression from input back to input still does a good job trying to eliminate the variation; and it can be assumed that this principle applies — at least approximately — also to other cybernetic domains. This issue deserves closer analysis.

2.3 Models and control

The principal subspace representation of the input data is, indeed, a rather sophisticated model of the environment. This fact needs to be emphasized:

> The mirror image that is constructed by a cybernetic system (see [19]) is not just any storage of observation data; it is an optimized *model* of the properties of the environment (as measured in terms of covariations).

Why does nature construct such sophisticated models? The deep insight necessary here is offered by *control theory.* There it is well known that the best control result, or the best elimination of external disturbances can be implemented if a model of those disturbances is available. Now the key observation concerning the nature of cybernetic systems can be expressed as follows:

> The interactions implemented by a (truly) cybernetic system are not whatever feedback structures; it constitutes an optimized *model based control* trying to eliminate the disturbances coming as input from the outside environment.

Traditionally in cybernetic (synergetic, autopoietic, etc.) studies, it is assumed that feedback alone is enough to implement the polished behaviors observed in complex dynamic systems. However, it is not so that whatever negative feedback would do a good job; indeed, the whole field of control theory studies this subject of constructing good feedback strategies. What is more, there are also semantic consequences: It is the whole conceptual universe of established theories and tools that become available. For example, applying the intuitions from control engineering, one can further elaborate on the true essence of cybernetic systems:

> The controls implemented by a (truly) cybernetic system are not whatever control structures; it constitutes an optimized *adaptive control* continually changing its behavior according to the changes in the environment.

Adaptive controllers are capable of automatically adjusting their behaviors based on the properties of the observed environment. Because of their promises, adaptive structures have been studied actively in control theory [1]. However, adaptive controllers have not become very popular in practice — it has been recognized that there emerge behaviors that are not beneficial and are difficult to master. The difficulties are fundamentally caused by *loss of excitation* in the system. Adaptation of behaviors can only be carried out if there exists some kind of a model of the environment, identified on the basis of the available information as delivered by the measured signals. Adaptive control is fundamentally based on model adaptation. Problems emerge when this adaptive model is connected in the feedback loop to implement model-based control. Whenever some behavioral pattern in the signals is detected, or when an appropriate model for it is found, it can be exploited and the pattern can be compensated; after this elimination, the behavioral pattern *no more exists* and the model becomes obsolete, so that a new model needs to be found. But when the previous obsolete model is forgotten, information is lost, and the original patterns of disturbance can again freely pop up. It is not so that this model adaptation would be a refinement process; it is a more or less cyclic process between states of bad behavior and good identification versus good behavior and bad identification, making the overall system behavior non-stationary and very complicated.

This adaptive control view opens new horizons to analysis of behaviors in cybernetic systems. Again, as in technical control systems, the structure of the excitation — whatever is the phenosphere — changes, as the originally most relevant input disturbances are compensated and less evident ones remain to be discovered. In this sense, as the cybernetic control system has evolved so much that it has enough momentum for essentially affecting its environment, eliminating the relevant input variation, the system adaptation starts getting astray, concentrating on the less relevant patterns. The final outcome can be that the model becomes ruined altogether and the boundary to chaos is crossed: If for some reason the original disturbance pattern is evoked, in the streamlined system there are no more mechanisms for tackling with it. Only after a turbulent period, the necessary structures are reconstructed — but this reconstruction of the control structures can be based on very different underlying subsystems and agents after the collapse.

As an indication of what kind of analyses will be carried out later in tis paper, let us apply the above technical studies to non-technical cases. It is tempting to extend these considerations also to very complex domains, like human societies: For example, after the Pax Romana was reached, the army was no more so essential, and not so much resources was put to it; finally, after centuries of peaceful status quo, the barbarians managed to conquer the empire. It seems that this is inevitable in all "too developed" systems; one can try to maintain the old good structures, but as the Soviet example reveals, the tensions just become stronger and stronger until the structures necessarily collapse. And it needs to be recognized that this process of decay is comprehensive: The loss of excitation is experienced by all subsystems that independently optimize themselves. It is hopeless to fight against the natural tendencies. It is not so that concentrating on some specific detail and fixing it the inevitable could be avoided (for example, putting more resources to military expenses does not necessarily help if the soldiers are unmotivated because of the missing threat). How could the collapse be avoided, then? In practical control systems it has been observed that the adaptive controls must not be *too good,* one should not implement the best possible control, allowing strong enough disturbances excite the systems — or extra noise can explicitly be added in the system. It seems that robustness and optimality in a system are contradictory goals. Again drawing bold conclusions: In the case of a culture, this means that new ideas, etc., need to continually enrich the society. A continuous turmoil is better than the complete collapse after a stagnation!

The above leaps from simple, concrete, low-level systems to societal systems consisting of humans were huge and heuristic. Too wild speculations are useless. The field of complexity theory is a notorious example of systems where "the *hole* is larger than its parts", resulting in empty words with no substance. However, the framework of cybernetic systems offers *concrete concepts and tools to motivate and functionalize the intuitions.* This claim will be elaborated on in the rest of this paper.

A good control has to be based on a good model of the system behavior, capturing the essence of the domain. Above, in the concrete ecological case, the feedback that was based on the latent PCA-type variables implements a control of resources that is optimal in a simple environment; however, as more and more complex systems are being discussed, when it is not only the second order statistical properties of data that are needed to capture the system dependency structures, more and more sophisticated models are needed for implementing good feedback. The control theoretical intuitions still apply when analysing such more complex systems. Indeed, in real life the models can become extremely complicated, and the agents implementing the control strategies, trying to drive the system towards balance, can be equally complicated. This applies also to cybernetic systems consisting of humans as agents: These actors have more or less thorough understanding (model) of the environment. Depending on this understanding, they have a vision of how a "non-tension" situation could be reached, and the actions are implemented accordingly ...

However, it needs to be admitted here that the necessary conceptual tools that are needed to turn intuitions into verifiable/falsifiable theories are not yet available. To extend the intuition to new, less structured domains, beyond simple population models studied above, more powerful ideas need to be employed. When modeling complex systems (biological, cognitive, etc.) the details are so overwhelmingly complex that the overall picture cannot be seen when looking at the individual processes alone. The number of underlying structural alternatives is essentially higher than the number of visible patterns, so that there exists an infinite number of alternative structures to choose from when explaining behaviors. To capture the most appropriate model structure, one needs strong modeling principles — stronger than what exist today. It turns out that to reach this one has to answer not only the *how* questions, but also the *why* questions. Traditionally, such teleological problem settings have been thought to belong to *metaphysics,* not to natural sciences. So, what is the essence of Logos, or the underlying *elan vital* in cybernetic systems, large and small?

3 ... Towards universal principles ...

It is also a non-trivial control system that is being implemented by the cybernetic system. How can this intuition be exploited?

3.1 About *entropy* and *order*

The concept of *entropy* is among the most fundamental ones in nature, and when searching for universal laws governing cybernetic systems, these issues need to be addressed.

Applying the thermodynamic interpretation (as defined by Clausius), entropy reveals the extent to which the energy in a closed system is available to do work (as defined in a somewhat sloppy manner). The lower the entropy level is, the more there is *free energy.* In a closed system, entropy level cannot decrease; it remains constant only if all processes within the system are reversible. However, because the natural processes typically are irreversible, entropy in the system increases, so that energy becomes "inert". Even though the total amount of energy remains constant, according to the *first law of thermodynamics,* it becomes less useful, according to the *second law of thermodynamics.* Ultimately, the system ends in a thermodynamic balance, or "heat death", where there is no more free energy available.

The cybernetic systems, as defined in [19], are also characterized by balances: First, the determination of $\bar{x}$ is based on finding the dynamic equilibrium as determined by the system model (1). Second, the matrices A and B, as defined in (2), are also dynamic equilibria as determined by the statistical properties in u (see [8]). Indeed, in a cybernetic system there are balances at each level — and, in this sense, the convergence towards a steady-state model is completely in line with the second law of thermodynamics.

However, the above observation is not yet intuitively sufficient, and some more analysis is needed. There are many definitions for the concept of entropy, and it seems that the corresponding intuitions are to some extent contradictory, or at least obscure.

In statistical mechanics (by Boltzmann and Gibbs), and analogously in information theory (by Shannon), entropy is related to probability: More probable states (observations) reflect higher entropy than less probable ones. In a sense, entropy is the opposite of information — less probable observations contain more information about the system state. In such discussions, the second law of thermodynamics, or the increase in entropy, is reflected so that systems tend to become less ordered, and information becomes wasted. This probability-bound interpretation of entropy is intuitively appealing, but it seems to result in paradoxes: For example, a symmetrical pattern is intuitively more ordered, containing more information, and consequently having lower entropy than a completely random pattern — on the other hand, symmetric pattern can be seen to contain *less information* than a random pattern, because the redundancies caused by the symmetricity can be utilized to represent the patterns more efficiently, so that the entropy level should be now *higher* now. Indeed, as discussed in

[11], the "algorithmic entropy" is higher in a symmetric pattern than in a non-symmetric one. To confuse concepts concerning order and symmetry even more, or, rather, to reveal the inconsistencies in our intuitions, think of the following claim: A totally unordered system can be said to be extremely symmetric as the components cannot be distinguished from each other.

However, here it is assumed, according to the original intuition, that *orderliness is a manifestation of low entropy.* The key point here is that the simplicity of symmetric patterns, or ordered patterns in general (loss of information in them), is just an illusion: *The missing information of the pattern is buried in our pattern recognition capability.* If the same data is to be presented without the supporting underlying mental machinery, or specialized interpretation and analysis tools, there is no handicap — the redundancy cannot be exploited, and no compression of data can be reached. In general, a higher-level representation makes it possible to *abstract* the domain area data; in other words, a *model* is the key to a compressed representation.

Similarly, in cybernetic systems one seems to be facing paradoxes, or fundamental inconsistencies that are not only of semantic origin. How is it possible that in some systems — being equally subject to real-life constraints and laws of nature — the arrow on entropy seems to be *inverted,* so that rather than getting disordered, new order emerges in them? For example, in cognitive systems, in social systems, and in living systems in general, more and more complicated structures are introduced in the course of evolution. Of course, there are no outright contradictions here (the subsystems getting ordered are open systems, the overall entropy in the universe all the time increasing), but why do the systems not select the "easy way", exhausting energy for simply increasing entropy? Why are there such countercurrents in the flow of entropy?

As in the case of symmetric patterns above, it is a higher-level structure representing the lower-level data that looks more ordered and "smart", as interpretated by our perception machinery. The PCA-based cybernetic system distinguishes between random noise and correlated variation in the data, thus compressing information so that it contains less noise. As this correlated variation in the environment is interpreted as information, the cybernetic system seems to act like a Maxwell Demon, distinguishing between two "containers" of information and noise, compressing information and pumping "negative entropy" into the emerging structures (see Fig. 3).

Applying the discussions in Sec. 2, the mystery of cumulating complexity can be resolved: It is the control system intuition that is needed to solve the "arrow of entropy" paradox. Even though it seems that entropy level goes down in some subsystems, when one looks at the structure among the systems more closely, the contradictions vanish: The system of decaying entropy is the supersystem, or control system, driving the subsystem *more efficiently* towards increased entropy, or "local heat death", as characterized by the stable balance (see Fig. 4). The minor decrease in entropy in the supersystem is compensated by the major increase of entropy in the subsystem, so that the second law of thermodynamics is obeyed — and, indeed, this entropy is now pursued more efficiently

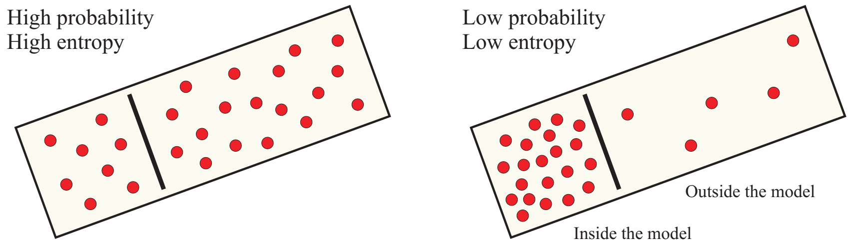

Fig. 3. Schematic illustration of how information, or the "units of covariation", are condensed in a cybernetic system

than otherwise would be possible. If the data properties are stationary, the cost of the constant higher-level structure becomes negligible as compared to ever-cumulating entropy on the lower level. More explicitly — it can be claimed that the second law of thermodynamics is the *motivation* for emergence of order in complex enough environments (see Sec. 3.2).

Perhaps the most important consequence of the new interpretation of the cybernetic systems is that reductionistic approaches become possible: Traditionally, the only systemically consistent level of studying entropy-decaying systems was the holistic level, the whole Earth as one entity, now each subsystem as studied alone is also thermodynamically consistent.

The new view of cybernetic systems as pursuing balance is fundamentally different from traditional intuitions. It has been assumed that "interesting" complex systems are *at the edge of chaos.* For example, when studying the processes of *life,* the mainstream view is expressed by Ilya Prigogine: Life is "as far as possible" from balance, whereas death means final balance. Erwin Schrödinger [14] phrased this as "What an organism feeds upon is negative entropy; it continues to suck orderliness from its environment". Also in cybernetic systems, static balance means death — but a living system is characterized by *(thermo)dynamic* balances. Now the roles are essentially inverted: Whereas a living thing is traditionally assumed to play an active role, now it just has to adapt to its environment; it is the environment that pumps disorder into the system, and life processes try to restore balance. It is not imbalance but the *homeostasis,* as explained by Bernard and Cannon, that is the essence of life processes; this balance is not only an emergent phenomenon but the very kernel of the relevant functionalities (see [19]). It is not about minimization of entropy, but, on the contrary, it is explicit maximization of overall entropy in a cybernetic system — this entropy increase is just channelled in a smart way!

Heraclitus claimed that the Logos running the world is *fire.* Perhaps a better characterization of cybernetic processes is that the universal mind running them is a *fire extinguisher* — incoming excitation is being attenuated. It seems that, again, the Eastern tradition is perhaps deeper than the Western is: The underlying vitality principle beyond the Chinese philosophy and medicine is based on balancing and ordering; see Fig. 5. — On the other hand, in Indian philosophy

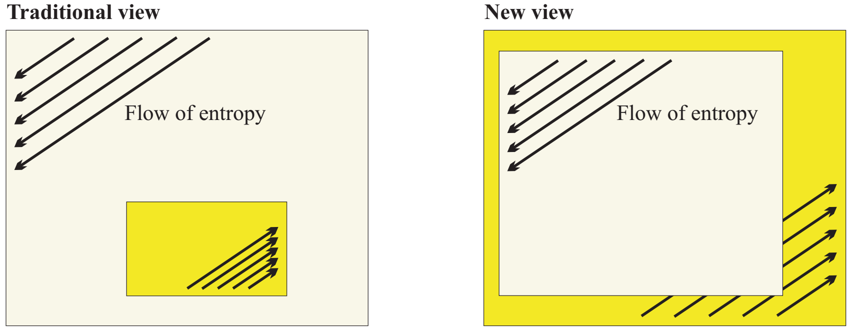

Fig. 4. The roles of subsystems and supersystems need another look (see text)

many principles (stationarity, desire and consequent suffering, etc.) also reflect the cybernetic ideas.

To summarize the above discussion, one can say that cybernetic (sub)systems constitute a framework for connecting successive system levels in the same framework, so that the emergence of structures within the overall "entropic" model can be explained. Structure (model) on the higher level means higher entropy on the lower level. It is the available resources that are the driving force keeping the system running, and maintaining the dissipative, irreversible processes in a cybernetic system; however, this underlying machinery is another thing from the actual *essence* determining the cybernetic functions. It is not so that the available resources would simply constitute the supply of "free energy" in the thermodynamic sense, so that the amount of resources would directly stand for amount of "neg-entropy"; the opposite of entropy, or incoming information, is revealed in the form of *variations* in resources. This is the same thing as in thermodynamics where free energy is buried in temperature *differences* rather than in temperatures themselves. As Gregory Bateson intuitively puts it [3]: "Information consists of differences that make a difference". Whatever is the interpretation of the resource vector, whatever are the physical dimensions of the input, the driving force in the cybernetic information cumulation is variations coming from outside; the system does what it does trying to eliminate these variations. The Shannon's formula just defines a static entropy measure, connecting information theory to thermodynamic domain in a formal way; it may be that the deepest interpretation and the most fruitful framework for cybernetic studies comes from a combination of these two fields in a more fundamental way. Perhaps one could here speak of "information theoretic thermodynamics". This is not only jargon: It turns out that the above discussions can also be put in practice.

3.2 "Principle of least difference"

Traditionally, the second law of thermodynamics is thought of as being a universal, more or less metaphorical principle. The existence of systems with inverted,

Fig. 5. Chinese symbol for the *ordering principle,* also denoting *air* or *vapour*

entropy-decaying nature has made it difficult to motivate explicit utilization of this principle in practice: It seems that the entropy principle cannot be applied in a reductionistic way for analysis of concrete large-scale systems.

Now, according to the above discussions, the entropy in a subsystem always increases when seen from the higher-level system. In a cybernetic system, entropy increases in a consistent manner, there is "balance pursuit" at all levels, completely in line with the second law where thermodynamical balance is the ultimate goal. Because of this consistency, any subsystem at any level can be studied separately, and also holistic systems can be analyzed in a reductionistic manner. In this sense there is no more difference between different kinds of complex systems: Living systems and non-living ones, for example, can be modeled in the same framework. Whereas the first law of thermodynamics (energy principle) offers powerful tools for deriving static models, it seems that the second law (entropy principle), being a fundamentally flux-based concept, offers generic tools for deriving *dynamic models.*

The entropy level, or, rather, changes in levels, can be applied as the measure for tensions in a cybernetic system. To make the very abstract entropy principle applicable in practical modeling tasks, one first has to define what *free energy* is in general terms; that is, one cannot rely on the particle-level considerations, but a system-level characterization is needed. The starting point here is that as it is the balance, or match with environment, that is being pursued in a cybernetic system, free energy originates from *mismatch* between the system and its environment. In this sense, one can speak of "principle of least difference", extending the principles of *least action* or *minimum energy,* as originally proposed by Maupertuis, and later extended by Euler, Lagrange, and Hamilton. In concrete terms, one can define a cost criterion J, or the energy function, as being the inverse of the "fitness":

$$J(x,u) = \frac{1}{2}\,(u - \phi x)^T\, W\,(u - \phi x)\,. \tag{11}$$

Here, u stands for the state of the environment, whereas x is the internal state of the system; it needs to be recognized that the state variables (elements of the vectors u and x) operate in different domains, and in between, a mapping matrix

ϕ is needed. Matrix W makes it possible to weigh data components in different ways. Typically the dimension of x is lower than that of u, so that no complete match can be reached. Now, free energy in the system state x can be defined as the difference between $J(x, u)$ and its minimum value $J(\bar{x}, u)$ — in minimum, free energy is exhausted.

The vector u can not only represent the real observed state of the environment — it can also stand for some *hypothetical* state. As presented in the next section, the same cybernetics-motivated modeling principles can be applied also for the iterative technical product development processes (and also in economical optimization processes); assuming that, for example, the goal is to reduce some quantity (weight, elapsed time, etc., depending on the device), in the goal state this quantity should be zero, and this deviation then constitutes the free energy, giving raise to technical development. As long as zero-delay, zero-cost, etc., has not been reached, the system will evolve. Of course, this optimum state is never reached, and the nice results concerning the cybernetic balances cannot be directly exploited. However, there are clear patterns here, too — even though the developments are random and sporadic, on average the developments take place in the direction of maximum gain, and typically the optimum state is being approached in a more or less consistent manner (typically following the exponent curve).

Mathematical formulations sound rather trivial, and it is difficult to see how any more complicated problems could be formulated in this vector-based framework. However, it needs to be rocognized that also the most complicated artificial intelligence methodologies have traditionally been formulated applying "problem spaces" and "goal states". Reasoning and planning tasks, for example, can be formulated so that a goal state is being searched for. Even though the problem space in AI cases is not necessarily a mathematically compact vector space, it needs to be recognized that the structural complexity can be substituted with dimensional complexity — that is, introducing enough new variables, the search problem in the high-dimensional space becomes simpler (remember the idea of *support vector machines*). Indeed, if correct variables are selected, so that local minima can be avoided, also reasoning problems become pattern recognition tasks in the high-dimensional space.

Entropy is also always being maximized in a cybernetic system. But powerful principles not only explain behaviors — they can also help to predict behaviors beyond the range of the original model. Indeed, the principle of least difference can be strenghtened: It can be assumed that *a system tries to maximize entropy as fast as possible.* Ideas concerning maximum (generalized) entropy have been studied a lot, but mainly for determining static optima; only in limited application domains such ideas have been applied for dynamic analyses. In the field of "autokatakinetics", as proposed in [15], the idea of maximum entropy production is intuitively proposed. Now the concrete, simple formulations make it possible to reach practical results and powerful modeling tools.

What is, then, the fastest route to maximum entropy, characterized by a dynamic balance with "hidden tensions"? It needs to be recognized that there is no centralized “master mind” in nature. Nature does not know the global optimum, or where to go to reach it. The optimization strategies that nature implements are decentralized, distributed to very local agents that only see their local environments, and do not know the “big picture”. Generally, the direction of fastest local decay in cost criterion is revealed by the (negative) gradient — and, indeed, at least in simple environments, the processes proceed so that higher densities, concentrations, temperatures, etc., are discharged towards lower ones. To emulate such locally consistent behaviors one can first write the gradient for the criterion (11):

$$\frac{d\,J}{d\,x}(x(t)) = \phi^T W \phi\, x(t) - \phi^T W\, u. \tag{12}$$

Now the continuous-time version of the *steepest descent* gradient algorithm can be written in the state-space form:

$$\frac{d\,x}{d\,t}(t) = -\Lambda A\, x(t) + \Lambda B\, u, \tag{13}$$

There are physical restrictions that determine what is this rate of adaptation, and these factors are collected in Λ. Combining (12) and (13), the matrices A and B are defined as

$$A = \phi^T W \phi, \quad \text{and} \quad B = \phi^T W. \tag{14}$$

To apply the simple dynamic model, the global criterion J is not explicitly needed; to justify the linear dynamic model, it is just assumed that it exists. On the other hand, if the criterion *is* known, the entropy principle offers a practical way to determine dynamic models also for complex systems.

A cybernetic system also carries out pattern matching against its environment, finding a balance between the outer and inner states. At least in special cases the above hypotheses are justified: In a truly cybernetically optimized system, as in a Hebbian/anti-Hebbian neuron system, ϕ defines the basis vectors of the principal subspace of the input u, and W is the input data covariance matrix; this assures that the system obeys not only the first-order balance but also the second-order balance (see [19]). It turns out that the quadratic formulation of the cost criterion can also be motivated in the Hebbian/anti-Hebbian framework — but, from the point of view of universal applicability of the maximum entropy principle, can this intuition be generalized; why should this formulation apply to other cybernetic systems in other domains?

The key notion here is that of a *diffusion process:* In diffusion systems the behavior is characterized by an explicit search for balance, or maximization of entropy. The tensions causing interactions come from the free energy that is manifested in the form of imbalances in concentrations, temperatures, or other distributed quantities. The diffusion process is internally balanced, negative feedbacks being automatically built up. Different kinds of diffusion processes are typically linearly dependent of the differences in the system — and this linearity can be interpreted in terms of quadratic cost criteria in the form (11). If finding

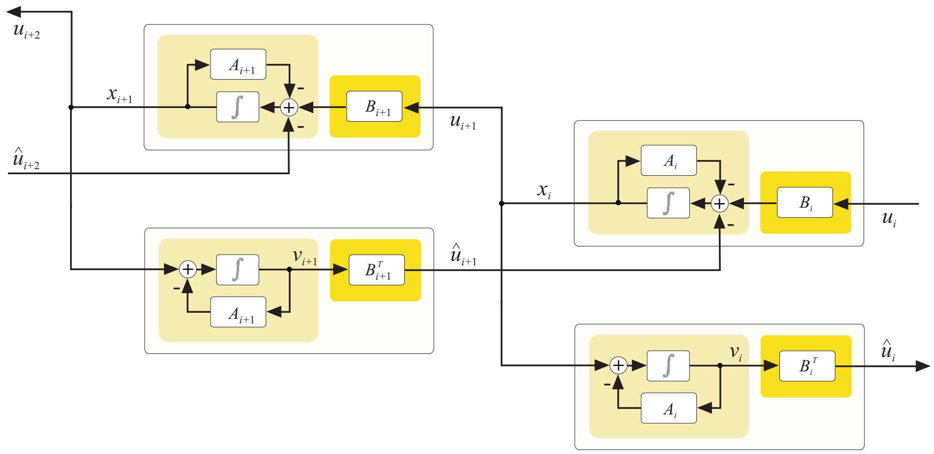

Fig. 6. Connection of trophic levels

the balance in a cybernetic system can be seen as a generalized diffusion process, the same framework (13) can always be utilized. When looking at a (truly) cybernetic process in a static perspective, it can be said to constitute a higher-order balance system; on the other hand, when seen in a dynamic perspective, it is a system characterized by *higher-order distributed diffusion.*

As discussed in [16], the exponential model is very plausible according to theoretical analyses and practical experiences, at least when studying population dynamics. The problems arise when the constraints have to be taken into account, resulting in very non-universal and intuitively non-appealing model structures. Luckily enough, in cybernetic models integrate the constraints in the same simple model structure. The model $\dot{x} = -Ax + Bu$ with real-valued eigenvalues in A is a general form of interacting (perhaps cascaded) diffusion processes with exponential behavior.

Another point worth mentioning here is that, again, one should take into account all feedback loops: It can well be so that the developments are not, after all, so straightforward as the diffusion intuition might suggest. For example, after a structrural "experiment", carried out by some of the system agents, it is not only the internal parameters of the system that get adjusted as the new balance is restored, but perhaps also the external parameters. In retrospect, some ecological/economical selections may look as the best choice merely because those decisions determined the "parallel universe" that was realized. It is not necessarily so that the system is an image of its environment; a strong enough system can make its environment reflect its own structure!

3.3 Escaping from a domain to another

Where does the free energy in nature, or variation in resources, originate from? The energy producing processes in the Sun are relatively stationary, providing

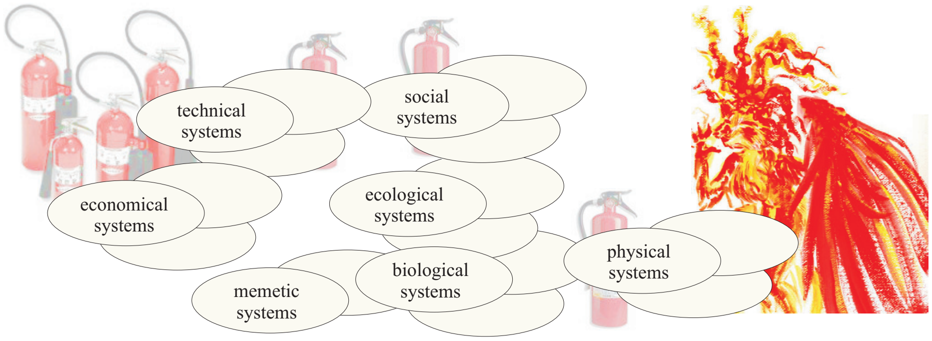

Fig. 7. Schematic illustration of cybernetic systems extinguishing the "fire"

a rather constant flow of energy, no variations. However, the orbiting of planets causing day and night, summer and winter, introduce the first variations; after that, there are the "locally unstable" (chaotic) processes further creating environmental variation. "First there was chaos": Chaos truly creates "information" as unanticipated fluctuations are generated in such processes — information that can subsequently be exploited by other systems (for analysis of information, see Sec. 4). For example, the fast enough fluid flows based on the Navier-Stokes equation introduce turbulence and nontrivial spatio-temporal distribution in energy. These physical processes are still "sub-cybernetic" — they remain stable because of the inherent nonlinearities, not because of the actual feedback structures. Or, even though the interactions can be interpreted as feedbacks, these feedbacks are physical, bound to concrete location, whereas the feedbacks in actual cybernetic systems are more like information theoretical. For example, planetary motion is not cybernetic; Newtonian action and counteraction is not sufficient to implement negative feedbacks in the cybernetic sense. What is more, there is no structural adaptation or evolution in these systems; qualitatively nothing different emerges as there exist no learning mechanisms (sand dunes, for example, even though virtually being manifestations of new order, have no negative feedback effect on the wind). However, as claimed in [4] the principle of maximum entropy production can be applicable also to such sub-cybernetic processes (the convection patterns of the *Bénard process* being used as an example).

This incoming variation is then processed by the actual cybernetic systems, trying to exhaust it. The complex dynamics, in the form of chaotic-like oscillations, is not an inherent property of the cybernetic model structure; the dynamics arises as a reaction to the external excitation. When different kinds of cybernetic systems are combined, qualitatively new levels of model complexity can be reached, resulting in more and more smart-looking systems. For example, when a cognitive system is integrated in a biological system, as is the case with individual humans, much more efficient exploitation of the environment is reached. Similarly, as a single cognitive system is integrated in a population system, as in the case of an "intelligent organization" where individual humans are appropriately networked, the resulting system can outperform the capabilities of the individuals. In any case, it is the entropy principle that still applies, all systems trying to eliminate the incoming variation (see Fig. 7). Thus, one could perhaps distinguish between different kinds of cybernetic systems: There is a continuum from *physical systems* through *natural* (or *normal*) *cybernetic systems* to *constructivistic systems,* and further to *technical systems,* depending of how much of the emergent structure is determined directly by the environment; in physical systems, there are no degrees of freedom, whereas in constructivistic systems the emergent structure is practically free of the environmental constraints. In technical systems it is not only the system that is (more or less) artificially constructed, as in constructivistic systems, but also the optimum state where the system tries to balance itself is imaginary (see later). Indeed, in economical systems the imaginary goal ("maximum amount of money", or "minimum cost") is visible in its most explicit form.

Entropy is universal, but it is not centralized: It separately governs the behaviors of the smallest systems as well as cosmic ones. There are no centralized processing or control units in nature; the only information available is the information delivered by the system itself, and the only energy available for self-organization is the energy engaged in the tensions of the system. The entropy-pursuit machinery is thus localized, characterized by distributed negative feedback mechanisms. Ultimate functionalism cannot be reached in such a distributed system of systems. This means that even though the model-based control can optimally attenuate variations in some subsystem, the optimality is lost in wider perspective. Indeed, this loss of optimality is a very fundamental phenomenon: Paradoxically, it turns out that locally optimized entropy maximization, or elimination of information, results in *maximal preservation of overall information.* This can be explained as follows: A good controller is aware of what is happening in the system being controlled, or what is the state u of the outer environment; this information that is used for controlling the environment is captured in the internal state of the controller x. This x containing the essence of the environment is available for yet higher level systems to be exploited as input resource! And, further, this relaying of information from former systems to latter ones can be repeated; there can exist an arbitrary number of successive levels in the chain of cybernetic systems, and the information is maximally transferred up to the top layer (see Fig. 6).

After the long slow evolution of natural systems, resulting in more and more complex structures, developments became much faster when a new machinery was once introduced: The human, and specially his/her cognitive capabilities offer a general-purpose control platform for implementing different kinds of higher-level systems. In this case the model is in another domain (phenosphere) as compared to the system being modeled. Nature relies on completely distributed strategies when implementing control systems, and, as was observed, this results in extreme complexity and illogicality; the human is needed to implement a more streamlined system, where the non-optimalities are ripped off. From the ecological point of view, this simplification and loss of diversity in the form of increased functionality and consistency of course has to be seen as impoverishment.

In human-constructed systems, the balances need not be “fight-or-perish” equilibria, they can also be negotiated. The human societies can be based on ideas by John Forbes Nash, not only by Adam Smith; the welfare society can, after all, be a good idea — if the faith of the underlying dynamics can be appropriately foreseen in advance.

It is not only “optimization” that is carried out by humans — completely new resources of free energy are also released by humans. In concrete terms, new variables are introduced in the resource vector by human activity. In a sense, the human has the role of a *catalyst:* New resources become available because of human activity, and processes that otherwise would never take place are activated. For example, the oil reservoirs would never have been exhausted if the human culture had not done that. Thus, after all, exhaustion of natural resources and destruction of the environment is inevitable and predestinated by the entropy law. And the success of the human can be measured in terms of increased entropy, or the rate of consumption!

It needs to be recognized that in more complex domains the solutions to the optimization problems, or models of how the free energy in the environment can be used, are by no means unique. For example, if there is *money* available in the market, there are many different ways of exhausting it. According to the selected strategy, a process can be instantiated to produce it; this process is further divided in subtasks with simpler goals, employing human staff for running the subprocesses, keeping the subcontrols in balance. The original push (economical pressure) is thus divided into investments, constituting a delicate structure. Within a selected (sub)structure, the implementation is more or less unique, but there are always many ways to select the structure.

In all systems, the bottleneck is caused by the scarcity of information and understanding. When something is better understood, or when a better model exists, more or less immediately after that (or, at least, when the new balance after the transient has been reached) it is exploited and new feedback structures are constructed. In this sense, one is always on the edge between the known and the unknown. This is easy to see in technical and economical systems, where the whole closed loop between the observations and actions is explicitly optimized; understanding is exploited in a straightforward way and more streamlined systems are implemented immediately when it is justified in terms of economical etc. considerations. However, also in politics the control of the society is implemented in a rather straightforward way by applying legislation, according to estimates and assumptions about future that are based on more or less accurate models of the society dynamics. The current state of the societal system is measured in terms of statistics, and also by opinion polls.

Today, the models can be constructed proactively rather than reactively: The company hierarchy can be designed for some production task, and the customer demand is created only afterwards. In complex domains there are no straightforward patterns to be matched in the environment, or, more accurately, the space of available variables is not fixed beforehand. In such cases, evolution is not a random but a goal-directed process; this evolution is much faster than in biological, etc., environments because it is based on human intermediating agents: Best strategies are searched for in a goal-directed manner, and spreading of new strategies can be instantaneous.

A human with such freely "reprogrammable" modes of behavior is a powerful agent for implementing different kinds of cybernetic systems. The human mind is a platform for societies, economical and scientific systems alike. It is not only the hardwired animal instincts but also the needs and desires that affect the human behaviors. Because of the flexibility of the human mind, the rules of the game common to all agents need to be imposed consistently upon everybody; otherwise, if the agents do not behave in a coherent way, there is no orchestration, and no cybernetic functionalities emerge in the overall system. In short, this coordination is implemented in the form of some kind of moral imperatives or codes, either religious, or philosophical/political. Indeed, it seems that religiosity of humans had an evolutionary advantage when organized societies and cultures were to emerge; otherwise, uncoordinated anarchy would have prevailed. Instead of rigid moral codes, nowadays the "moral control" is based on much more flexible and efficient means: Today's humans are controlled to a large extent by fashions, and specially by *money.* It is money that offers a universal measure for quantifying very different kinds of things in a transparent way, making the control actions by the "system" very fast and unambiguous; in concrete terms, money makes things commensurable, so that the above proposed vector-form coding of complex phenomena is justifiable — however ethically objectionable such rude evaluating and assessing might feel like. However, it is with human agents as it is with ants travelling on their paths: There is the mainstream flow, but there also exist rebellious dissidents, introducing necessary variation in behaviors so that there is possibility of changes in behaviors.

It is in the human nature that one searches for new frontiers to discover and conquer, to understand and exploit — or, to model and control. Putting it poetically, one can admit that God gave man dominion over the world — but there is catch: Man first has to detect and identify that world. The human builds models, but, simultaneously, he is also a model of his environment; man is an image of God — but this God is Gaia. And it is the lure of money that implements the whispering of the modern secular conscience.

"Grasping" a thing means that a model capturing the relevant phenomena has been successfully constructed. Arthur Schopenhauer already observed that human understanding and intellect are there just for biological reasons, or rational thinking is a tool for fulfilling the physical needs. But Schopenhauer claimed that *art* is a key to escape the rat race: Aesthetic experience is, according to him, free of any concrete utilitarian purposes. However, following the above discussions, it turns out that also the aesthetic experience can be explained in the same framework of better fulfilling ones needs: It is an efficient means of constructing more versatile models of the reality. Remember that there often exist analogues among systems in different domains, that is, there can exist common patterns among very different systems; or, more accurately, the same model structures can be applicable in different domains. The more general the modeling principles are, the more probable it is that there are similarities. For example, learning to see symmetries in concrete patterns can help to see corresponding structures also in more abstract domains, so that such pre-created model structures can be reused for reinterpreting, that is, for finding creative new associations. In this sense, art (or, actually, practically any field of special expertise) can truly help in seeing nature in new ways, and constructing more efficient (subconscious) models for other systems.

3.4 Basics of constructivism

The claim here is that, really, it is the entropy principle, as implemented in cybernetic control structures, that governs also complicated cybernetic systems, like all human behavior and activity. The purpose of all human information gathering, for example, is gaining knowledge and understanding; understanding is the route to exploitation of deposits of variation that exists in the natural resources. The feedback from understanding to explotation is seen as a control loop. What makes this often difficult to recognize is the fact that the implementations of cybernetic control can be so distributed, and the application domains are typically so non-mathematical.

In cybernetic systems it is populations of agents that determine the system behavior; in ecosystems, etc., the agents themselves are a part of the system, operating in the same domain, whereas in more complex environments, the agents operate in some *other* domain. Typically this means that there are more degrees of freedom available, and the laws governing the adaptation are not so stringent. In such constructivistic systems different kinds of approaches are necessary; however, it can still be claimed that the operations still maximize overall entropy production.

Technical product development processes are typical constructivistic cybernetic systems, where the dynamics is caused by tensions determined by external constraints and the "technological drive". How does development of *computer technology,* for example, boost entropy in the overall system, then? Word processors, typical software products, are used to construct descriptions of complex domains; these models are (more or less balanced!) views of the domain field. More sophisticated word processors make this model construction process faster; and the new hardware and software tools make it possible to share this model, and compare its virtues with competing models. Specially, when doing scientific research, the Internet technology has made the "paper production" process much faster as the information availability has increased. Computer technology also boosts the delivery of new models (ideas and theories) from one domain (cognitive system) to another (scientific community). A scientific paradigm is determined as a balanced interplay between such theories, being a cybernetic combination trying to explain (model) the subject domain. Straight after such modeling is satisfactory, technical applications are introduced, where the developed model is applied for controlling, or manipulation of the domain field.

In this sense, development of the computer technology finally result in better mastering of the environment, in practice meaning that the available resources become more efficiently exploited, so that the free energy decays faster.

Quite concretely, the programming languages are tools for easier construction of models of complex domains; the better the tool, and the more polished the program is, the better match with the reality can be reached. Simulation tools, on the other hand, are tools for transforming the information in static descriptions into functionally more applicable form. The computer is really a universal machine, offering a general-purpose platform for implementing any kinds of models and for functionalizing them, to be applied in different ways for finally somehow affecting the environment.

The key point here is that there is a huge number of subprocesses intertwined in a higher-level cybernetic domain; indeed, there seems to exist a *fractal structure* of cybernetic systems involved. It is not only the original system; it is also the information availability, message transfer speed, quality in terms of accuracy, etc., that are typically under the process of ever-continuing development. This fractality of parameters cannot easily be formalized, because all substructures can further be split in subparts. Indeed, whenever a subsystem has become appropriately identified, it becomes a subject of further development. Typically in technical environments, this development means making the system faster, smaller, cheaper, etc. How does the agent doing the development work know that there are some cosmic consequences, about entropy, or control of environmental variations? The agents in ecolocigal systems do not know anything about their control function, and neither do the humans know anything about the big picture — the overall behavior is again an emergent phenomenon. It is as it is with simpler cybernetic systems: If the agents just follow the common principles (go towards resources, avoid competition), the higher-level functionalities automatically emerge, and, as a system, the adopted strategy has evolutionary advantage. However, as compared to ecosystems, etc., now there are no immediately evident resources or competitors, there is no immediate benefit or threat, and these simple principles are no more applicable as such. The goals in a constructivistic system are defined in terms of a not-yet-existing environment; more powerful "hardwiring" of the agents is needed.

In short, the new "constructivistic imperative" that is needed to explain the development in complex cybernetic systems can perhaps best be paraphrased as *citius, altius, fortius.* The urge to make things in some sense better is reflected in the technological drive, and in human ambition in general, whatever is the branch of activity. As motivated later, in constructivistic systems this principle is just another formulation for good modeling of not existing but *possible* systems. And, again, because of its success among the alternative strategies, this has also the evolutionary advantage. In concrete terms, what are the properties that are needed so that the agent — the human — could fulfill its task?

– **How systems become constructed?** When escaping from a simgle domain to another, the agent needs to be smart enough, and, additionally, to try something before there is any evident use of it, the agent needs to be curious — that is, the agent needs to be simultaneously *homo sapiens* and *homo ludens.*

– **How systems get adapted?** To make the systems adapt when there is no immediate need for that, a driving force is needed; this constructivistic imperative is not any explicit rule, it is an implicit tendency that is manifested in greediness and ambition.

It has been claimed that one of the main differences when humans are compared to animals is that humans can think of the future, they can plan, and they can imagine what the environment *could be* rather than merely adapting to the current circumstances. Human behavior is *proactive* rather than reactive. Humans can visualize the optimum state and the route to that state.

And, of course, yet another key feature in humans from the point of view of acting as agents in constructing cybernetic systems is their social nature: Because of the *monkeying,* imitating other's behavioral patterns, is so natural to humans, new ways of behaving can easily be instantiated and new kinds of systems are possible. Even though *free will* is said to be one of the main things that characterizes us as human beings, it is the *absence* of free will that characterizes societies. This can also be expressed as (mental) laziness: Avoiding difficulties, going where it is easiest, not against opposing forces but following them, leads to a dynamic balance also in abstract systems. However, it needs to be recognized that always when behaviors become too homogeneous and predictable, there is room for local opportunistic optimizations, perhaps resulting in parasitic strategies.

When does the *group-think,* or the group being dumber than the individuals, change into *system intelligence,* where the society is cleverer than the individuals? Perhaps the key here is — according to the above intuitions — adaptation and balance: Only after the sustem reaches the steady-state, internal tensions compensating each other, the system level structures emerge.

3.5 Return to Natural Philosophy

Neocybernetics is not just another scientific paradigm; no, it goes beyond that, shaking the very foundations of science. Traditionally, it is thought that it is philosophy (logic) that is the basis for mathematics, and, further, mathematics is the basis for technical research (engineering disciplines). Now, in the framework of cybernetic systems, this thinking can be inverted, as shown in Fig. 8: Mathematics (linear algebra) offers the syntax (language) and engineering (control practices) offers the semantics (interpretation) for philosophical considerations (metaphysics). Put in another way, empirism precedes rationalism, giving substance to "das Ding an Sich". This claim is further elaborated on below.

To start with, it is the same with neocybernetics as it is with other scientific disciplines: One has to admit that *models are always false.* The essence of the real world cannot be captured, and the models should never be mixed with reality. It was the start of modern science when one started doing physics based on

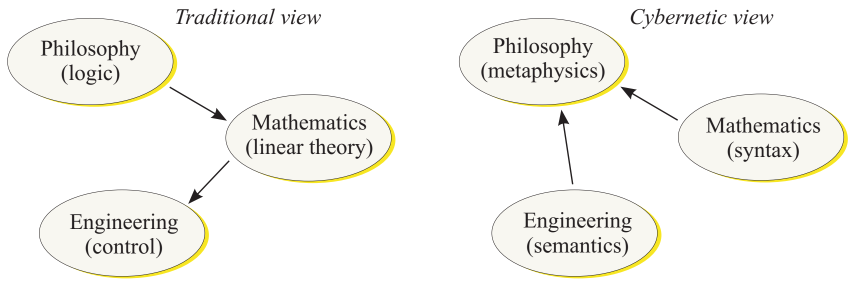

Fig. 8. Relations among theories and metatheories

empirism, only trying to explain the observations, rather than metaphysics, trying to explain the underlying reasons for those observations. However, things are different when one is constructing "higher-order models", or *models for models.* Now it can be claimed that *models are essentially true.* What does this mean?

One has to remember that the model construction of cybernetic systems also applies to the cognitive domain: The mental system constitutes a "mirror image" of the environment as determined by the observations. No matter what the underlying realm truly is like beneath the observations, the mental machinery constructs a model of it. If the same modelling principles are copied in the computer, there will be a fundamental correspondence among the data structures as constructed by the computer, and the mental representations as constructed by the brain in the same environment. This is an extension of the Kantian revolution: Perceptions are not observed but constructed as the observations are matched against what there already exists. This makes it possible to reach *intersubjectivity* of representations, technical or natural: The world models can be essentially the same, not only between humans but also between humans and computers. This makes it perhaps possible to reach Artificial Intelligence in the deep, not only in the shallow sense. Clever data processing becomes possible: The computer can carry out the data preprocessing in a complex environment, and the constructed data structures can be interpreted naturally in terms of corresponding mental representations.

But this intersubjectivity is not all there is; indeed, one can reach *interobjectivity.* If nature itself tries to construct models for eliminating free energy in the system, as presented above, the human trying to model these cybernetic systems can touch not only the shadows of the behaviour (in the Platonian sense), but the actual essence – these models can be *fundamentally the same.* This means that if some naturally evolved cybernetic system (an ecological system, for example) is modeled by a human applying the appropriate principles, this model has a deep correspondence with the system itself; what is more, in environments that are in a transient (an economical system, for example), the cybernetic models can predict what the final system would look like after the stationary state perhaps is reached. In this sense, the new models can perhaps give insight in the true essence of complex systems and in the hidden tensions in such systems. This observation has also cosmic consequences: Whatever are the systems on the other planets like, assuming that those systems are similarly based on local agents and evolution processes towards better exploitation of the environmental resources, they must obey the same universal principles. The celestial ecosystems, or social systems, etc., most probably do not essentially differ from the earthly ones — of course the details differ (like surface patterns), but the principles remain the same.

Universe constructs models — and, after all, models are used for *simulation.* It is not only Douglas Adams who claimed (in his book "The Hitch Hiker's Guide to the Galaxy") that the Earth itself is a huge computer carrying out (distributed) simulation: Edward Fredkin proposed in early 1980's "a new theory of physics" based on the idea that the Universe was comprised ultimately of software.

There are many philosophical issues that can be attacked in the cybernetic framework from the fresh point of view. It may need to be admitted that meta*physics* cannot be addressed in the current framework, physics being based on spatial interactions among particles, but one can perhaps speak of *metabiology, metaecology,* etc. What do these metatheories stand for? There are fundamental questions concerning modeling issues that have not seriously been questioned before: For example, the Pallas Athene Hypothesis mentioned above (or, perhaps more accurately, Antero Vipunen Hypothesis) is not just an unsubstantiated claim but there is some deeper essence there. Indeed, this hypothesis can be expressed in a stronger form: It seems that system complexity and analyzability go hand in hand: If Nature has been able to construct sophisticated model structures, why not us? The clain here also is that *cybernetic systems can always be modeled,* one just needs to find the appropriate model structure. Perhaps a new era of *positivism* is ahead? – And, to mention another deeply philosophical principle: *Ockham's razor* is routinely being applied in modeling ("simplest explanation is the most appropriate"), but the motivation for this idea is typically merely pragmatic. In the framework of optimized cybernetic systems, the models being based on principal subspaces, etc., extreme compactness truly is the final faith of the fully evolved systems.

But this optimality in cybernetic systems applies only when the system is seen in the local perspective. Indeed, the discussions above can be summarized so that it is not, after all, some *intelligent designer* that is responsible for all natural diversity — rather, looking at the immense inconsistence, one could speak of a *hardworking idiot:* The left hand does not know what the right hand is doing. The resulting system of systems is neither systematic nor systemic. However, it may be so that *finalism* will have a renaissance: As explained in connection with entropy, there *exist* goals in natural systems — and such systems are not constrained to biological or ecological domains.

Most human endeavors can also be interpreted as manifestations of the same cybernetic principles. One specially interesting group of cybernetic domains of human activity is that of *scientific research.* Scientific theories are again models of the environment, whatever is the branch, no matter if it is natural or human sciences. The more complex the domain field is, the more there are degrees of freedom, and the less the available data constrains the possible solutions. As the hierarchy of complexity evolves, it is difficult to evaluate the priority among candidate explanations. The latter layers are dictated more by the prior layers than the actual environment being explained; in complex enough cybernetic domains the system starts to create its own meanings not bound to the outer realm. This is not dependent of the field of study, this is more like a property of "too evolved" science, where there are more theories than evidence (for example, take cosmology and its *wormholes, parallel universes,* etc., being manifestations of *ironic science*). The cybernetic view towards doing science makes it perhaps easier to reach *reconciliation* between the "two cultures" within sciences [17]; in these postmodern times, as there is more and more pressure towards new results, the scientific explanations are similarly constructivistic in natural sciences as they are in human sciences. Also natural scientists and engineers should be humble. It is often claimed that science proceeds positively towards higher levels of perfection following its own internal laws — and these laws should be only determined by objective criteria of *truthfulness.* However, it is the human that is always integrated in the loop of doing science — this means that it is not only the match against evidence that alone determines the vitality of a paradigm. The humans determine what is "hot" and what is not. The shifts between Kuhnian paradigms are not so clear-cut, and it seems that also science is on the edge of chaos; the term *ironic science* has been coined [5]. Within a cybernetic system there are subsystems; indeed, science is a fractal structure of cybernetic systems — such embedded systems are studied in what follows.

4 ... And back to concrete domains

The above discussion is not merely semantic jargon; down-to-earth analyses are possible. To elaborate on the abstractions, the concept of *information* turns out to be useful. Information can be studied in mathematical terms, and it can be used as a link between the abstract and concrete ideas.

4.1 About *information*

Regardless of the underlying mechanisms (interacting agents, or explicit constraints), a cybernetic system is characterized by dynamic balances. When seen from outside, and when the phenomena have been quantified appropriately, the system implements *principal component analysis,* or, actually, *principal subspace analysis* of the incoming data. In either case, the model captures the (co)variation in the data in the most efficient and compact way. Simultaneously, the PCA model is capable of structuring and reproducing the variation: The compression of high-dimensional multivariate data is based on the distribution of variation. Now, if this variation in data is called *information,* there are efficient means of mathematically processing and analyzing that information.

Information is here a purely syntactic concept, being measurable in strictly mechanistic means. At this level there is no need for any interpretation or semantics, and information should not be mixed with *knowledge.*

The above definition of information is only technical, but it is also intuitively useful and justifiable. First, there is the pragmatic support: Because of the mathematical benefits, error square criteria have traditionally been applied in modeling — quadratic criteria result in linear optimization methods. Variance also enables *commensurability* of signals in different dimensions: Variances can directly be added together, giving a single scalar measure (assuming uncorrelatedness). Physically, the variation is easily interpretable because it stands for signal power. But perhaps the most important motivation for variance-oriented approaches is that it gives a possibility to capture *semantics* in some limited sense: Assuming that it is *contextual semantics* one is interested in, the covariation among variables carries a scent of the *essence* of the system. This interpretation of *meaning* sounds trivial, but the mathematically solid basis makes it possible to process and reprocess the data, and when enough of it has been carried out, something qualitatively new can emerge. And when anthropocentric interpretations are added, results can look *smart.* It needs to be emphasized here that semantics is not only related to the cognitive system but to all cybernetic domains.

It is this information that is now identified with the free energy that was discussed above: Information carries the correct connotations in this context. It needs to be emphasized that it is not the *level* or absolute value of the signal that is of relevance, but its deviations from the long-term average. If the levels of all incoming resources remain constant, or even if they vary in an identical cycle reflecting the same underlying variable, there is little free energy available: Constant flow of information results in “heat death”, or trivialization of the system, because only one principal component is needed to represent all the variation, so that a single population/species can exhaust all resources in a centralized manner; this means that diversity will be lost. Complex variation of resources, on the other hand, results in nontrivial distribution among actors.

Just as in traditional control systems, also in cybernetic systems there is the physical and the “metaphysical” level: There are the process flows, and there is the flow of information. Contrary to the Prigoginian views, it is abstract information rather than the concrete energy that flows through in the dissipative processes.

Even though variation in signals has always been applied for modeling purposes, now there is a fundamental shift in thinking: Traditionally, the fresh variation is seen as *noise* that is being minimized — now, on the other hand, variation is explicitly taken as welcome phenomenon. Indeed, measurements always tell about system properties but also about the observation process, and very different results are found when the point of view is changed. This claim is best clarified when one looks at the cost criteria that are being minimized in the modeling processes: Again, study the general criterion (11), where u is the input data and ϕx is its reconstruction applying the model. The weighting matrix

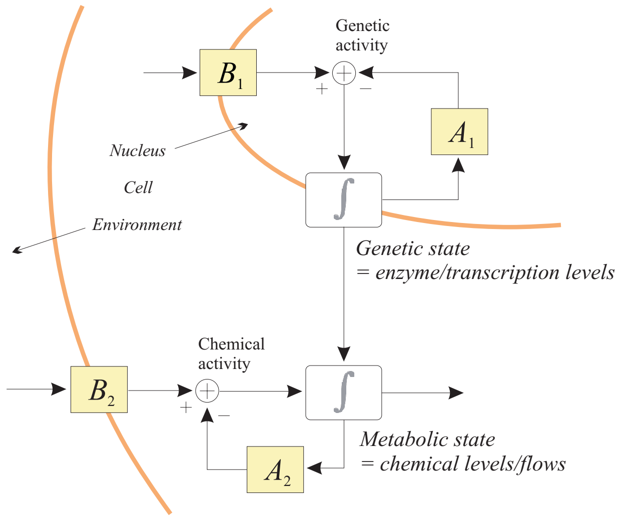

Fig. 9. Levels in a cellular system: Genetics and metabolics

W determines the mutual relevances among data components, thus (partially) determining how the system sees the world. When doing traditional *maximum likelihood matching,* for example, one has

$$W = \mathrm{E}\{uu^T\}^{-1}, \tag{15}$$

meaning that variation in those directions where there is most variation is suppressed. In the cybernetic case one has

$$W = \mathrm{E}\{uu^T\} \tag{16}$$

instead: It is evident that the more there is variation in some direction, the more that variation is weighted in matching. Deviations are now interpreted as valuable resource.

4.2 Emergence of structures

Cybernetic system becomes a mirror image of its environment, the representations being optimized in the local-level interaction processes. All agents experience the same environment; how is it possible that the emergent systems have self-organizing structures where different agents have varying roles? To have more insight, let us study two very different examples.

In Fig. 9, a schematic illustration of the proceses in a living cell are presented (cf. [19]). There are two very different kinds of subsystems: The first metabolic subsystem is based on chemical balance reactions, where the balance is being found as being constrained by the incoming chemical flows through the cell wall. The set of available chemical reactions is essentially dictated by the enzymes that are produced as determined by the active genes. The second level, or the genetic subsystem, is based on very different kinds of elementary processes, like

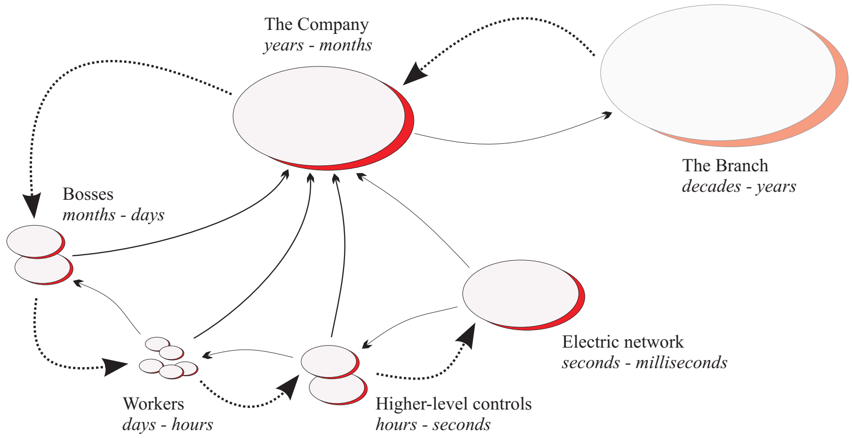

Fig. 10. A cybernetic economic system is constructed of various subsystems that are equally cybernetic

message-RNA production, transfer, and decoding, and it is practically impossible to model this dynamics in a mathematically compact form; the situation becomes still more complicated as there are chains of gene activations, and the interactions esentially form a deeply interconnected network. The *transcription factors* essentially dictating the gene activities are products of some other gene activity. However, again, when one concentrates only on the resulting balance after the transients have decayed, things become much simpler, and it can be modeled as a linear system around the equilibria. In both subsystems, the enzyme / metabolite levels are integrals of the genetic / chemical activities, respectively, and their local dynamics around the equilibria can be modeled applying the basic cybernetic model. Enzymes being produced in the genetic subsystem are catalysts, not being exhausted in the subsequent metabolic processes, so that the link betveen the subsystems is, in principle, unidirectional; however, as there is also metabolic feedback from the outside the nucleus into it, the hierarchy is not complete — rather, the structure is cyclic. In any case, it is a reasonable abstraction to distinguish between the two levels: this is due to the *different time scales* between the subsystems. Chemical balances are reached much faster than the gene activation balances.

In Fig. 10, on the other hand, illustrates the many-level structure of a typical company: On the highest level, there is the outside society determining the operating environment, and the lower levels represent the organizational arrangements within the company. It is assumed here that the company being studied is an electric company — the lowest-level subsystems are again "outside" the company, their behaviors being determined again by the environment, whereas the company actions try to keep these outside systems in some intended balance. Again, each level of subsystems operates in different time scales (as illustrated in the figure). The higher-level subsystem seems to determine the "set points" for

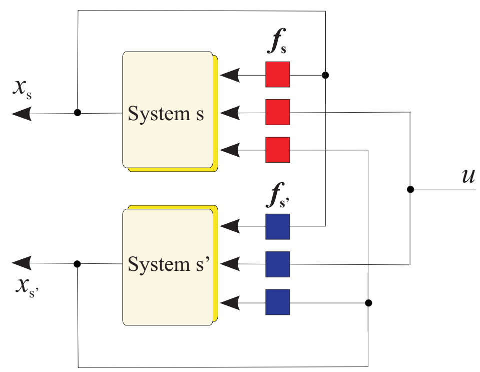

Fig. 11. Where the feature extraction can be implemented

the lower-level subsystems. The same structure can be detected also in industrial automation systems: on the lowest level, there are the natural (chemical) processes being stabilized by the lowest-level controls; the next level of balances is determined by the regulatory controls, trying to reach the reference values; the yet higher level in the cascade structure of controllers is determined by the production optimization. When seen as a single system, the dynamics is "stiff", containing very slow and very fast modes; the system can be seen as a combination of algebraic and dynamic constraints. Such stiffness problems vanish when different levels are studied separately.

These examples show how levels and (fractal) hierarchies are encountered and manifested in real life, the subsystems being more or less independent. Each of the subsystems tries to answer to all demands and needs, resulting in local balances — but the balances in different levels are tightly linked together. How to functionalize the notion of hierarchies in a cybernetic model? Where does the seemingly universal multi-level structure of balances emerge from? It seems to be natural to distinguish between levels, but how can this be implemented in a mathematically solid framework?

4.3 Frequency-domain hierarchies

Whatever the environment is really like, it is visible to the agents only in the experienced observation data. It is not the objective environment that is being experienced by all agents in an absolutely homogeneous way; the agents' "world views" are subjective, depending on how the signals are seen and how they are interpreted. In other words, this can be expressed in the form: How the features are extracted from the data? Within a single subsystem the feature extraction principles remain invariant (see Fig. 11).

In mathematical terms, feature extraction can be simplest implemented (that is, applying strictly linear methodology) by appropriate weighting of the signals. When this weighting is carried out in *frequency domain,* the time scales can naturally be taken into account — and it was these time scales that seemed to be the main difference between subsystems. So, let us introduce a new func-

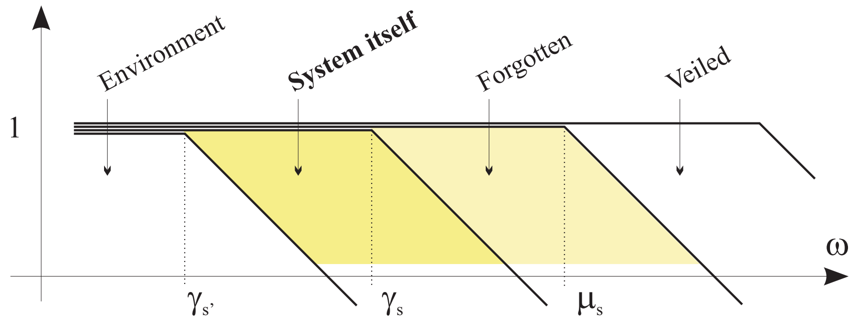

Fig. 12. How information distribution can be analyzed applying control engineering intuitions (see text)

tional structure in the cybernetic model: Assume that the signals coming in the subsystem s are *filtered* through the low-pass filter

$$\frac{du_{\mathrm{s}}}{dt} = -\mu_{\mathrm{s}} u_{\mathrm{s}} + \mu_{\mathrm{s}} u_{\mathrm{in}}, \tag{17}$$

so that the Laplace-transformed input signal is

$$U(s) = F_{\mathrm{s}}(s)\, U_{\mathrm{in}}(s), \tag{18}$$

where the transfer function is

$$F_{\mathrm{s}}(s) = \frac{\mu_{\mathrm{s}}}{s + \mu_{\mathrm{s}}} \tag{19}$$

with time constant $\tau_{\mathrm{s}} = 1/\mu_{\mathrm{s}}$. This means that beyond the angular frequency $\omega = \mu_{\mathrm{s}}$ the visible signals start decaying; the subsystem essentially cannot see behaviors that take place at higher frequencies (they are "veiled", as shown in Fig. 12). The information content (power in terms of variation) in the filtered signals can be expressed in terms of the power spectrum:

$$U(\omega) = H_{\mathrm{s}}(\omega)\, U_{\mathrm{in}}(\omega), \tag{20}$$

where the power spectrum corresponding to (17) is

$$H_{\mathrm{s}}(\omega) = \frac{\mu_{\mathrm{s}}^2}{\omega^2 + \mu_{\mathrm{s}}^2}. \tag{21}$$

Note that more precise cut-off behaviors can be implemented applying higher-order filters; for example, one can define

$$F'_{\mathrm{s}}(s) = F_{\mathrm{s}}(s) F_{\mathrm{s}}(s) = \frac{\mu_{\mathrm{s}}^2}{s^2 + 2\mu_{\mathrm{s}} s + \mu_{\mathrm{s}}^2}. \tag{22}$$

It is interesting to recognize that whatever is the rate of decay, the curve is always piecewise linear on the log/log scale, that is, when both frequencies and signal powers are presented as logarithmic quntities. Remember the scale-free structures: Also in fractal systems, the dependencies between variables often follow the same functional form. Perhaps this view offers yet another intuitive explanation why such *power law* behaviors are so common in nature?

Information below the cut-off frequency is seen by the system with no attenuation, but that information does not cumulate in the subsystem: Signals pass directly through, even though they can be changed. Only effects that cumulate in the covariance structures remain in the system. This gives us a new interpretation of the covariance matrix estimate adaptation formulation applied in the cybernetic models: Information is being filtered through a linear first-order filter. Information below the cut-off frequency $\omega = \gamma_s$ represents the "essence" of the subsystem (see Fig. 12), so that

$$\frac{d\hat{\mathrm{E}}\{x_s x_s^T\}}{dt} = -\gamma_s \hat{\mathrm{E}}\{x_s x_s^T\} + \gamma_s x_s x_s^T, \tag{23}$$

and

$$\frac{d\hat{\mathrm{E}}\{x_s u_s^T\}}{dt} = -\gamma_s \hat{\mathrm{E}}\{x_s u_s^T\} + \gamma_s x_s u_s^T, \tag{24}$$

where $\gamma \ll \mu$ (or, actually, $\log \gamma \ll \log \mu$). If there are subsystems in still lower frequency regions, the corresponding lower-frequency informaton is mostly captured by those dedicated subsystems; on the other hand, if there exists information in the signals at higher frequencies than μ, there will probably be another subsystem. As seen by this intermediate subsystem, the exterior ambient information is abstracted, and seen only as constant data determining the more or less fixed environment: Too fast signals are filtered so that only the mean value remains to be seen, and too slow signals are being faithfully followed. The subsystems outside the "information horizon" cannot be explicitly controlled by the current subsystem: It must be assumed (in the spirit of cybernetics) that the higher-frequency systems succesfully balance themselves — only reference values can be delivered by slower (higher-level) systems. The intermediate system is, on the other hand, subject to the reference values coming from the yet higher-level systems; as long as there is dynamics in the slow signal range, the outer environment changes, and there is need for continuing adaptation. To repeat — now data represents the resource, level in data represents the environment, and variation in data represents information; and models of that information represent memory.

In traditional cybernetic studies (Bateson, Maturana, etc.) the interactions and feedbacks in the system are very concrete sequential chains of effects or processes. Now the time scales, or frequency ranges, are selected so that the actions at one level become seemingly *instantaneous* and *simultaneous*. Rather than studying dynamic phenomena, one studies static patterns. Earlier this static nature of the patterns was explained so that it is balances that are being studied; now, however, this simplicity is reached thanks to the powerful mathematical machinery. The stationary statistical phenomena can compactly be manipulated in frequency domain, where the individual signal realizations, their initial values, etc., are abstracted away. In principle, stationarity of signals is assumed, but because it is not assumed that the system has reached the final balance, the analyses can be made more versatile. Luckily enough, because of the selected information interpretation, information being identified with signal energy, the essential signal properties are so easily transferred from time domain into frequency domain.

In the proposed framework of modeling cybernetic systems (to reach emergent models), one must always abstract the actual time axis away, concentrating on statistical patterns rather than dynamic processes. In evolution, for example, the basic unit is one generation, so that the time constants must typically be of the order of hundreds of years — and statistical learning necessarily takes thousands of years! There are phenomena taking place also faster; indeed, the interaction structure within the system is very much dependent of the selected frequency range, different time scales offering very different views into the phenomena taking place in the system. The physicists' dream concerning the *Theory of Everything,* or the theory explaining everything in terms of elementary particles, or, in this case, in terms of elementary cybernetic entities, is futile. A good model studies only the most appropriate phenomena visible at a specific level, so that when seen from different distances there are different models. No universal model exists; however, when studying cybernetic systems the model structure may still remain the same.

It is natural that phenomena at higher frequency ranges are observed more often than phenomena at lower frequencies. This means that — when starting from scratch — there is statistically relevant information available first only concerning the phenomena with fastest dynamics, and only this information is modeled originally. As time passes, also slower phenomena can be mapped by successive subsystems. In this sense, in complex enough environments (like ecosystems) it is natural that the structure of cybernetic systems is in a constant turmoil, because there always exist lower frequency regions where there is non-exhausted information. The newer systems typically utilize earlier ones, and hierarchies emerge also without explicit filtering of signals, because of the temporal evolution, and because the stationary state has not yet been reached.

There are also technical reasons motivating the introduction of the filtering scheme (17). Many constructivistic systems operate in discrete time — that is, information (statistics, etc.) are collected at certain time points, and it is assumed that behaviors between the sampling instants remains approximately the same. Now there is yet another intuition from control engineering available. Theory tells us that the information sampling rate in the system determines the *Nyquist frequency:* The maximum frequency in the signals must not exceed half of the sampling frequency, otherwise the high frequency components get *aliased* on the lower frequencies, creating "phantom behaviors". Too high frequencies have to be filtered away. Yet another related problem emerges when the sampled measurements are used for control: Faster and faster controls make the closed-loop system time constant shorter, thus increasing the high-frequency contents perhaps beyond the sustainable level. Finally this is the problem that is faced in "quartal capitalism", where the company performance is to be optimized without longer-term horizon, short-term benefits being maximized, trying to be faster than what is the natural dynamics in the business field; similarly, this problem applies also in politics, where the approval and popularity has to be earned in a shortsighted manner, following momentary fashions and public opinions. And even in scientific work where the "production rate" is being monitored, it may be that the overall quality suffers: Experts have no time to evaluate the innovations, and no conclusions are ever drawn. Trying to go beyond the range of available information means going beyond the edge of chaos.

But it seems that this is a completely natural and unavoidable tendency in constructivistic cybernetic systems that are subject to all-embracing optimization activities. Remember that in Sec. 2 the control actions tried to minimize the resource deviations; this is only one view of eliminating free energy from the input in an adaptive system. Time integrals of variations are reduced either through making the signal amplitudes lower, or through *minimizing the time span of the deviations.* Whereas only the first alternative (making controls more accurate) is available in natural systems, in constructivistic systems also the second one is available (making controls faster). Both approaches eliminate information from the lower-level system, so that they both are *routes to chaos.*

In an ever-evolving cybernetic environment, the closed loop behavior becomes gradually more and more pathological: As the controls become faster, the inner loops are finally no more essentially faster than the outer ones are, and the dynamics become mixed, ruining the guaranteed stability properties of the individual loops. What is the (imaginary) final outcome from this process, assuming that such optimization continues infinitely? Next, it is shown that the system can reach a *qualitatively new level,* where it is again very different theoretical tools that are needed.

4.4 Relations to optimal control

Optimality and robustness are often contradictory goals: Faster control means higher sensitivity to disturbances, and this means worsened stability properties and lower robustness against noise. Is this unavoidable, or are there ways to circumvent this? Indeed, there is some counterintuitive intuition available here.

Assume that the time constants within the cybernetic system are essentially in the same range, that is, the feedbacks have become so fast and powerful that the signals have no time to naturally converge but the whole feedforward/feedback structure constitutes a single dynamic system. This is not all; assume that the feedbacks are so powerful that the environment can be completely controlled, that is, $u(t)$ can be altered at will, so that effectively one has $u = -\Delta u$. The vector $x(t)$ is the state of the cybernetic system, starting from some non-optimal state x_0, and the goal is to bring it to optimum state (for simplicity, assume that this optimum state is $x = 0$). Because of the internal system dynamics, governed by the formula (1), the state cannot be immediately altered; the faster one wants to make the effects, the larger are the controls that are needed. One faces an optimization problem that can be formulated in the following form:

$$J = \int_{t_0}^{t_1} \left(x^T(t)Qx(t) + u^T(t)Ru(t) \right) dt + x^T(t_1)Sx(t_1). \tag{25}$$

Here, matrices Q, R, and S are positive (semi)definite; for simplicity, Q and R can be selected to be identity matrices. As shown in [10], the optimum behavior of the controlled system can be expressed in the form

$$\begin{cases} \frac{dx}{dt} = -Ax - BR^{-1}B^T v \\ \frac{dv}{dt} = A^T v - Q^{-1}x, \end{cases} \tag{26}$$

assuming that the state can be directly measured (that is, matrix D in [10] equals identity). Because $-A$ has all its eigenvalues in the non-positive half-plane, it is stable — but simultaneously $A^T = A$ must be unstable, there is positive feedback, meaning that there is inherent instability buried in the system structure! However, as the LQ theory also dictates the boundary values as

$$\begin{cases} x(t_0) = x_0 \\ v(t_1) = Sx(t_1), \end{cases} \tag{27}$$

it turns out that the solution is meaningful; because of the boundary value problem formulation, the solution of the problem in this form involves iterative simulation-based approaches. The symmetricity of the structures in Fig. 2 (and the symmetricity of A) resembles the structure in (26); indeed, this becomes still clearer when one rewrites (26) so that dynamics of v is represented in "inverse time":

$$\begin{cases} \frac{dx}{dt} = -Ax - BR^{-1}B^T v \\ -\frac{dv}{dt} = -A^T v + Q^{-1}x. \end{cases} \tag{28}$$

This can be expressed as shown in Fig. 13; the "inverted" integral denotes the "backward nature" in the feedback branch. There is a qualitative transition here, because the causality structure has been inverted. Are there natural systems where such optimal control strategy would have been implemented? Of course not — uncausal structures going backward in time are nonphysical, and unstable systems when going forward in time are equally unphysical. But in artificial systems, like in constructivistic cybernetic systems, such uncausal/unstable models could easily be implemented?

The optimality formulation can be simplified by relaxing the end time; the stationary "infinite-time" control problem can be expressed in terms of

$$J^\infty = \int_{t_0}^{\infty} \left(x^T(t)Qx(t) + u^T(t)Ru(t) \right) dt. \tag{29}$$

To implement the corresponding control one has to solve the steady-state Riccati equation

$$0 = Q - \bar{P}BR^{-1}B^T\bar{P} + A^T\bar{P} + \bar{P}A. \tag{30}$$

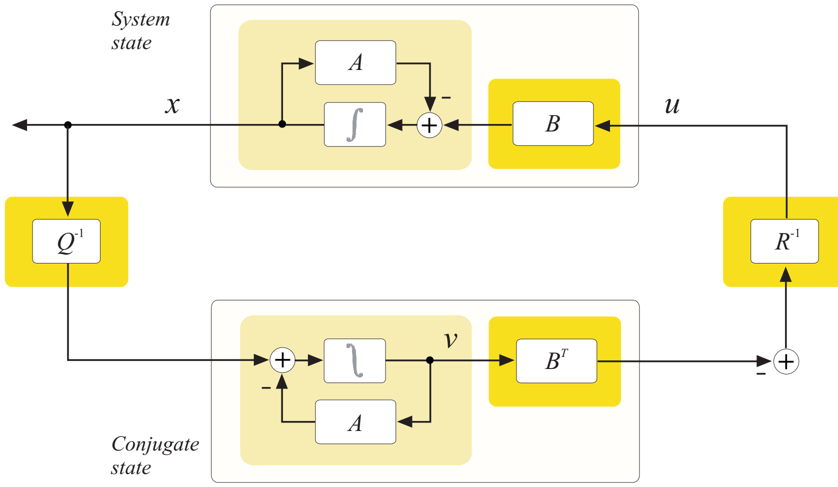

Fig. 13. Implementation of the optimally controlled system (see text)

When this is solved for $\bar{P}$, it turns out that that the boundary-value problem can be transformed into an initial value problem:

$$\begin{cases} x(t_0) = x_0 \\ v(t_0) = \bar{P}x(t_0), \end{cases} \tag{31}$$

so that the final, time-invariant state control law (skipping dynamics in the feedback branch) becomes

$$u(t) = -R^{-1}B^T\bar{P}\,x(t). \tag{32}$$

It needs to be recognized that the system model (in terms of matrices A and B) by no means needs to be optimal, as long as they describe the signal dependencies appropriately. Depending of the model, x is the interpretation of the corresponding world state, and within this framework, the presented structures still optimize the controls: The interpreted system state x_0 can be brought to optimum (zero state) applying the assumed available inputs to the system. What is interesting here is that information on both trophic levels is simultaneously manipulated and minimized; signal squares (variances) being appropriately suppressed not only in the lower-level system as is traditionally the case. Perhaps in the constructivistic systems one can get from the "hardworking idiot" metaphor back to intelligent design?

4.5 Agents vs. populations

Because of the principal component nature of the information manipulation in a cybernetic system, information is maximally relayed from u to x; this is what theory says. However, this information preservation is not complete: As the dimension of x is assumedly lower than that of u, typically variation (information) in some direction in the data space is ignored. But it turns out that cybernetic systems are truly marvellous: Nature can do better, exceeding the theoretical constraints; in principle, *all* information in data can be exploited by the system.

The vector x has been interpreted (in ecological, etc., applications) so that x_i stands for the size of the population representing the population i. The key to understanding the above claim is to recognize that whereas the number of variables within the vector x, or the number of distinct populations is less than the dimension of u, the total number of *individuals* within the populations still by far exceeds the degrees of freedom in u. Indeed, this information/entropy analysis gives us tools for understanding the diversity of individuals within a population, and why such nonhomogeneity has evolutionary advantage and thus never disappears even in stationary environmental conditions.

Assume that the dimension of u is m and that of x is n. After the n most significant principal components have been captured in the subspace spanned by the populations characterized by the variables x_i, the remaining $m - n$ dimensions in data (assuming that the data is non-redundant) are visible within each population, and *subpopulations* with differing properties can emerge within the main populations, the intra-population behaviors following the same principles that also govern the inter-population behaviors. Remember that typically there are no clear "knees" in the eigenvalues when covariance matrices are analyzed; this means that it is difficult to say what is the natural dimension of the data. Nature has the same problem when it implicitly operates on that same data, dimension selection meaning selecting a specific number of variables, or populations, in the model — this means that there is no such thing as the absolutely "correct" number of populations (or species) to be chosen. There is fractal continuity of properties within an ecosystem, and the distinction between populations and subpopulations becomes blurred; however, because of the biological reasons and ecological definitions, clear boundaries between species exist. As a matter of fact, another factor increasing diversity within an ecosystem is the observation that, when looking at the formulae in Sec. 2, it seems that the higher the populations of identical agents are (x_i) the lower is the available activity level (v_i) of the individuals; in this sense, it is clear that subpopulations try to differentiate themselves — in terms of total number of individuals, it is better to have two (or more) nonidentical subpopulations than only one, because this reduces competition.

The current framework makes it possible to analyze the property distributions further, and it is also possible to make predictions. For example, it can be assumed that in the final state, the total number of subpopulations (within each population separately) equals the number of remaining degrees of freedom in data; and the spectrum of properties in different species is the same. This means that within the populations there are latent substructures — if the environmental conditions suddenly change, so that some less significant principal components suddenly become more prominent, there already exists capability of reacting to such changes in the ecosystem, and the adaptation to the new conditions can be remarkably fast.

4.6 Views of thew world

The very concrete view of the emergent (sub)system structure being dictated by the available information deserves to be studied still closer. The cybernetic system always optimizes its representation of the visible information; or, as the system is living in its own narrow world, it *thinks* it is behaving optimally. This optimality is determined in terms of the properties of the measurement data, dictated by the statistical signal properties, so that the *analysis of the final system properties can be reduced to analysis of statistical signal properties.* There are many possibilities to utilizing this observation. In concrete terms, parallel systems forage different information when their feature extraction strategies differ from each other.

Information in signals cannot be increased at will, but it can be reduced appropriately. A good example was presented above, where dynamic filters where applied to suppress certain frequency bands. The approach presented in Sec. 4.3 was specially interesting, because it was through applying linear dynamic structures that the features were defined, so that the there are again powerful tools for analysis of the closed-loop system available. It needs to be commented here that it is typically assumed that linear structures are uninteresting when representing complex phenomena, because several linear layers do not increase the expressional power of the system; also, in a cybernetic model one linear layer already exhausts available correlations-related information. However, multiple layers still *can be* interesting in the cybernetic case: Because it is also nature that constructs non-optimal redundancy, one layer on top of another explaining the same variation, the layer-like linear models can reflect the cybernetic realm.

But truly unanticipated phenomena in the signals can be revealed when nonlinearities are employed. There are many alternatives to how features can be extracted, or how some information can be ignored, applying nonlinear functions; for example, there are examples in [8] where *sparse coding* of data is reached when "cut" function is applied, eliminating negative signal values altogether. Another possibility is to implement *topology* among the structures using nonlinearities: For example, localaized sensor networks were studied in [19] by explicitly cutting away the far-apart agent connections. This kind of location-based representations are typical in nature, where local populations that are located far from each other, for example, do not interact.

The structural nonlinearities can also be utilized to extend the dynamic structure in cybernetic models. Normally it is static mappings from u to x that are being modeled here, so that there is assumed to be no connection between, say, $u(k)$ and $u(k+1)$. In nature, the successive inputs are, however, typically not independent of each other: There exist correlations between resources in successive years, for example. And assuming that predators are longer-living than prey, they see longer periods of time, and it is the total amount of prey over the whole life span that determines the population rather than the situation during one single year. Indeed, as presented in [2], this kind of dynamic states can be captured in terms of time series samples. If one augments one single signal by including also the past values in the data vector, one can start modeling dynamics,

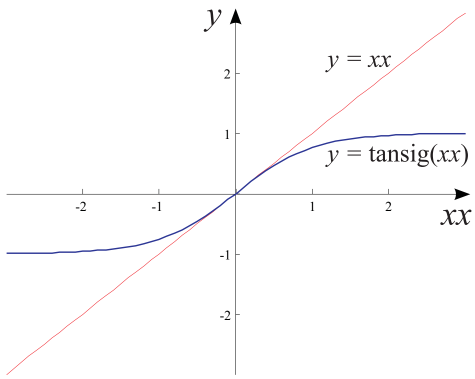

Fig. 14. Nonlinear cumulation of information

so that static principal component analysis changes into *subspace identification* (see [12]):

$$u_i'(t) = \begin{pmatrix} u_i(t) \\ u_i(t-1) \\ \vdots \\ u_i(t-d) \end{pmatrix}. \tag{33}$$

It has been assumed that information is interpreted directly in terms of variations (that is, the "atoms of information" are coded in the form $\mathrm{E}\{u_i^2\}$, for example) or covariations (in the form $\mathrm{E}\{u_i u_j\}$). However, the nonlinearities can also be applied here. Whereas the storages for information were in the "linear" case the covariance matrices, now one can define, for example,

$$\frac{d\hat{\mathrm{E}}\{f(xu^T)\}}{dt} = -\gamma \hat{\mathrm{E}}\{f(xu^T)\} + \gamma f(xu^T), \tag{34}$$

the nonlinearities in the matrices being calculated elementwise. This means that *not all information has the same weight.* For example, in Fig. 14 a nonlinearity is presented where the "high ends" have relatively lower emphasis, the nonlinearity being mathematically defined as

$$f(x_i u_j) = \frac{2}{1 + \exp(-2x_i u_j)} - 1. \tag{35}$$

The results after model convergence are shown in Fig. 15. The data in the experiments was first whitened, that is, $\mathrm{E}\{uu^T\} = I$, so that second-order statistical properties in the data vanish; whereas the standard cybernetic model will not have any preferences what comes to directions in the data space, the model with the nonlinear covariance calculations seems to have consistent behavior. Indeed, when data is concentrated at a certain distance from the center, it has proportionally higher weight, and the principal component axes will be tilted accordingly. Without any concrete proofs, it is claimed that rather than implementing principal subspace analysis (PSA), actually the model carries out *independent subspace analysis* (ICA or ISA) here (see [9]).

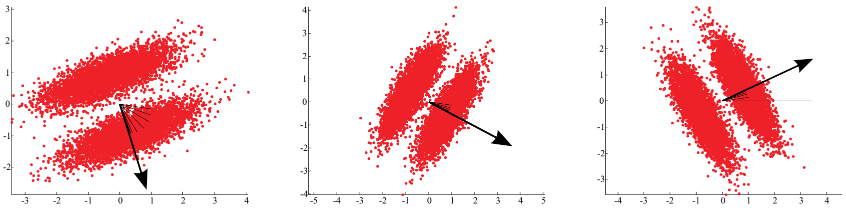

Fig. 15. Resulting behaviors of "nonlinearized" cybernetic models: Latent variable axes turn towards *independent components*

This implementation of "cybernetic ICA" is conceptually so simple that it may offer fresh intuitions for further developments of higher-order statistical analyses. Can it be assumed that the cybernetic combination of feedforward and feedback structures always results in lower-level properties being inherited on the higher level, that is, properties defined in terms of scalar functions f being generalized into high-dimensional functionalities?

5 Conclusion

Cybernetics is interdisciplinary, the same principles being applicable in very different disciplines. And this is not all; the interdisciplinary nature extends not only over the spectrum of models for natural systems, but also over the spectrum of modeling paradigms and philosophies. It is not only analysis but also synthesis of systems that can be understood within the cybernetic framework; and it is not only humans that are faced with the cybernetic complexity, but also Nature itself is. Modeling of natural systems is not only describing behaviors, but models — when implemented in an appropriate way — can also tell something fundamental about the underlying system. The human's subjective world is model of the environment — but also the objective world consists of models. It seems that building "mechanical brains" in the sense of deep AI may sometime be possible: Man can understand the mechanisms of nature, and these mechanisms can also be explicated and implemented in technical environments (see Fig. 16).

It may be that the above discussions go too deep in mathematical nuances. However, the beauty of Nature is buried in the details. And now there exist powerful frameworks for keeping the details in control — multivariate statistics, linear system theory, and specially control engineering are the basics of tomorrow's intellectual *avant garde*. Philosophy and logics is the basis of mathematics, and mathematics is the basis of applied mathematics and engineering; however, it can be claimed that applied mathematics gives substance to future philosophies. Without concrete down-to-earth application domains, and without the mathematical tools and control engineering intuitions, the abstract claims about entropy etc. would have no substance.

The diversity of nature can be explained in the cybernetic framework, as soon as the *kernel of variation* exists. The role of the divine designer can be reduced to the question: Why is there *something* instead of *nothing*?

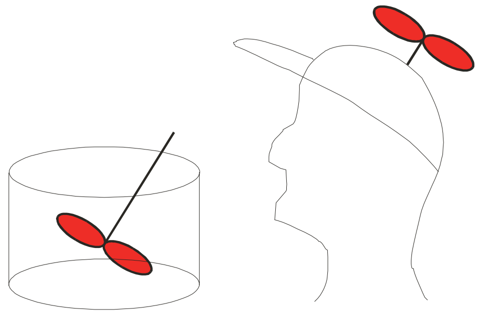

Fig. 16. In some dissipative systems ("ideal mixers") adding energy results in higher entropy, whereas in some other ("idea mixers") structures emerge and entropy goes down

References

1. K.J. Åström and B. Wittenmark: *Adaptive Control.* Addison–Wesley, Reading, MA (1989).
2. K.J. Åström and B. Wittenmark: *Computer-Controlled Systems — Theory and Design.* Prentice Hall, Englewood Cliffs, NJ (2nd edition 1990).
3. G. Bateson: *Steps to an Ecology of Mind.* Paladin Books (1973).
4. A. Kleidon, R.D. Lorenz (eds.): *Non-Equilibrium Thermodynamics and the Production of Entropy: Life, Earth, and Beyond.* Springer–Verlag, Berlin (2004).
5. J. Horgan: *The End of Science: Facing the Limits of Science in the Twilight of the Scientific Age.* Broadway Books (1996).
6. H. Hyötyniemi: *Self-Organizing Artificial Neural Networks in Dynamic Systems Modeling and Control.* Helsinki University of Technology, Control Engineering Laboratory, Report 97 (November 1994).
7. H. Hyötyniemi: Cybernetics — Towards a Unified Model? Submitted to the *Finnish Artificial Intelligence Conference (STeP'04),* Vantaa, Finland (September 2004).
8. H. Hyötyniemi: Hebbian and Anti-Hebbian Learning: System Theoretic Approach. Submitted to *Neural Networks* (2004).
9. A. Hyvärinen, J. Karhunen, and E. Oja: *Independent Component Analysis.* John Wiley & Sons, New York, NY (2001).
10. H. Kwakernaak and R. Sivan: *Linear Optimal Control Systems.* Wiley (1972).
11. S.-K. Lin: Correlation of Entropy with Similarity and Symmetry. *Journal of Chemical Information and Computer Sciences,* Vol. 36, pp. 367–376 (1996).
12. P. van Overschee and B. de Moor: *Subspace Identification for Linear Systems: Theory — Implementation — Applications.* Kluwer Academic Publischers, Boston, MA (1996).
13. R. Rosen: *Essays on Life Itself.* Columbia University Press (1998).
14. E. Schrödinger: *What Is Life?* Macmillan (1947).
15. R. Swenson, R.: Autocatakinetics, Evolution, and the Law of Maximum Entropy Production: A Principled Foundation Toward the Study of Human Ecology. *Advances in Human Ecology,* Vol. 6, pp. 1–46 (1997).
16. P. Turchin: *Complex Population Dynamics: A Theoretical/Empirical Synthesis.* Princeton University Press (2003).
17. E.O. Wilson: *Consilience: The Unity of Knowledge.* Abacus (1999).
18. S. Wolfram: *A New Kind of Science.* Wolfram Media (2002).

19. *Additional material on cybernetics will be available in public domain in near future at* `http://www.control.hut.fi/cybernetics`.