1st Workshop of the Principia Cybernetica Project

Computer-Supported Cooperative Development of an Evolutionary-Systemic Philosophy

Workbook

edited by

Francis Heylighen

Workbook of the
1st Principia Cybernetica
Workshop

edited by Francis Heylighen

Free University of Brussels

PRINCIPIA CYBERNETICA

Brussels • New York

Heylighen, Francis (editor)
Workbook of the 1st Principia Cybernetica Workshop.

Preface

This workbook contains abstracts and short papers selected for presentation at the 1st Workshop of the Principia Cybernetica Project by the workshop scientific committee. It is meant to give an overview of the work that will be discussed during that workshop. As such it will allow the participants to prepare themselves for discussion by examining the links, agreements, and differences, between their ideas and those of the other participants. In a second stage it may also function as a proceedings, providing a memory of what was presented there in Brussels in July 1991.

The aim of the project, and the corresponding workshop can be summarized as follows. *Principia Cybernetica* is an attempt by a group of researchers to collaboratively build a system of cybernetic philosophy, moving towards a transdisciplinary unification of the domain of Systems Theory and Cybernetics. This philosophical system will be developed as a *network*, consisting of nodes or concepts, linked by different types of semantic relations. The network will be implemented in a computer-based environment involving *hypermedia*, *electronic mail*, and *electronic publishing*. The project naturally splits into two issues:

1) development of the philosophy itself, which is *systemic* and *evolutionary*, emphasizing the spontaneous *emergence* of higher levels of organization or control through variation and natural selection. It includes: a) a metaphysics, based on processes as ontological primitives, b) an epistemology, which understands knowledge as constructed by the subject, but undergoing selection by the environment; c) an ethics, with the continuance of the process of evolution as supreme value.

2) development of computer-based tools and methods for *collaborative theory building* (CSCW, groupware, SGML, knowledge acquisition...): many participants with different backgrounds and working in different places exchange knowledge and opinions about a common problem; their different contributions and reactions must be integrated and structured, in order to form a coherent system of concepts and values, transparently modelling the problem domain.

Both issues are united by their common framework based on cybernetical and evolutionary principles: the computer-support system is intended to amplify the spontaneous development of knowledge which forms the main theme of the philosophy.

The contributions in this book have been classified in 5 sections. The first one offers a general overview of the project, emphasizing its history, its main philosophical positions, and its method. The second one addresses the issue of foundations for cybernetics in general. The next section applies the concepts of evolution and of the emergence of multiple levels to traditional philosophical questions such as the origin

of meaning and organization. The fourth section emphasizes more in particular the development of knowledge and culture. The last section studies different ways to use computers as tools to support the further development of knowledge, in particular the knowledge system that will incorporate the Principia Cybernetica.

Brussels, May 1991   F.H.

List of Contributors

Beyls, Peter; O. Van Dammestraat 73, 9030 Gent, Belgium.

Carvallo, Marc; Dept. of Philosophy of Religion, State University of Groningen, Nieuwe Kijk in 't Jatstraat 104, 9712 SL Groningen, Nederland, E-MAIL: marccarv@hgrrug5.bitnet.

Elohim, J.L.; Antonio Sola 45, Col. Condesa, C.P. 06140, Mexico D.F.; Fax: +525-761-5023

Glanville, Ranulph; 52 Lawrence Road, Southsea Hants, PO5 1 NY, U.K., TEL. (+44) (705) 737 779.

Glück, Robert; Technical University Vienna, Institut für Prakt. Informatik, Argentinierstr. 8/180, A-1040 Vienna, Austria; TEL. (0222) 54 20 63, E-MAIL: E1802DAA@AWIUNI11.Bitnet.

Henry, Charles; 217M Butler Library, Columbia University, New York, N.Y. 10027, USA, TEL. 212-854-5477, E-MAIL:
henry@cunixf.cc.columbia.edu .

Heylighen, Francis; PO-PESP, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium; TEL. + 32-2-641 25 25, E-MAIL:
Z09302@BBRBFU01.BITNET.

Joslyn, Cliff; Systems Science, SUNY Binghamton, Box 1070, Binghamton NY 13901, USA; TEL. +1 -607 729-5348, E-MAIL:
cjoslyn@bingvaxu.cc.binghamton.edu.

Kenis, Dirk; VTBP, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium, TEL. + 32-2-641 2749.

Löfgren, Lars; Lunds Universitet, Systems Theory, Box 118, S-221 00 Lund, Sweden, ; TEL. +46-46 10 75 19 [office], + 46 - 46 12 88 27 [home], E-MAIL: lofgren@dit.lth.se.

Moreno, Alvaro; Research group IAS (Information, Autonomy, Systems), Dept. of Logic and Philosophy of Science, University of the Basque Country, Apartado 1249, 20080 Donostia-San Sebastian, Espana, E-MAIL: alvaro@fil.ehu.es,

Moritz, Elan; The Institute for Memetic Research, P.O. Box 16327, Panama City, Florida 32406, USA, E-MAIL:
moritz@well.sf.ca.us.

Pask, Gordon; 48 North Street, Clapham Old Town, London SW4 0HD, UK; TEL. + 44 - 71 - 720 18 30 (home), +44 - 71 - 738 82 41 (home).

Peschl, Markus F.; Dept. for Philosophy of Science, University of Vienna, Sensengasse 8/9, A-1090 Wien, Austria; TEL.: +43 222 42 76 01 / 41; E-MAIL: a6111daa@vm.univie.ac.at

Turchin, Valentin; Computer Science Department, City College, CUNY, Convent Avenue at 138th Street, New York NY 10031,

USA, TEL. +1 - 201-337 1761 (Home), +1-2126506178 [office], E-MAIL: TURCC@CUNYVM.BITNET.

Van de Vijver, Gertrudis; Seminarie voor Logica en Kennisleer, RUG, Lamoraal van Egmontstraat 18, B-9000 Gent ; TEL. 091 - 64 39 61 [office].

Table of Contents

The Principia Cybernetica Project

Cybernetic Foundations

Evolutionary Philosophy

Knowledge development

Computer-Support Systems

THE PRINCIPIA CYBERNETICA PROJECT

Francis Heylighen
PESP, Free University of Brussels

An Evolutionary System Modelling Evolutionary Systems: Introducing the Principia Cybernetica Project

The need for a Principia Cybernetica

It is a common observation that our present culture lacks integration: there is an enormous diversity of "systems of thought" (disciplines, theories, ideologies, religions, ...), but they are mostly incoherent, if not inconsistent, and when confronted with a situation where more than one system might apply, there is no guidance for choosing the most adequate one. Philosophy can be defined as the search for an integrating conceptual framework, that would tie together the scattered fragments of knowledge. Since the 18th century, philosophy has predominantly relied on science (rather than on religion) as the main source of the knowledge that is to be unified.

After the failure of logical positivism and the mechanistic view of science, only one approach has made a serious claim that it would be able to bring back integration: the *General Systems Theory* (von Bertalanffy, 1968; Boulding, 1956). Systems theorists have argued that however complex or diverse the world that we experience, we will always find different types of *organization* in it, and such organization can be described by principles which are independent from the specific domain at which we are looking. Many of the concepts used by system theorists came from the closely related approach of *cybernetics*: information, control, feedback, communication... In fact cybernetics and systems theory study essentially the same problem, that of organization, albeit with an emphasis on either structures and models (systems), or on functions and communications (cybernetics). In order to simplify expressions, we will from now on use the term "cybernetics" to denote the global domain of "cybernetics and general systems theory".

Though a lot of recently fashionable applications (e.g. artificial intelligence, neural networks, cyberspace, man-machine interfaces, systems therapy ...) have their roots in ideas that were proposed by cyberneticians, cybernetics itself tends to stay at a distance from the mainstream scientific developments, and is correspondingly not taken

seriously by that mainstream. Moreover, though cybernetics aims to unify science, it is in itself not unified. I wish to argue that, instead of looking down on practical applications, cyberneticians should try to understand how those applications can help them in their task of unifying science, and, first of all, unifying cybernetics. It should look upon them as *tools*, that can be used for tasks that may extend much further than the ones they were originally designed for.

A similar situation arose around the end of the last century. Mathematics proposed a great variety of very successful applications: geometry, calculus, algebra, number theory, etc. Yet there was no overall theory of mathematics: these different domains functioned mainly in parallel, each with its own axioms, rules, notations, and concepts. Though most mathematicians would agree that these subdisciplines had a "mathematical way of thinking" in common, one had to wait for the classical work of Whitehead and Russell (1910-13), the *Principia Mathematica*, before this unity could be clearly expressed. What was novel in this work was that *mathematical methods were applied to the foundations of mathematics itself*, formulating the laws of thought governing mathematical reasoning by means of mathematical axioms, theorems and proofs. This proved highly successful, and the Principia Mathematica stills forms the basis of the "modern" mathematics as it is taught in schools and universities.

Our contention is that something similar should be done with cybernetics: integrating and founding cybernetics with the help of cybernetical methods and tools. Similar to the mathematical application domains (number theory, geometry, etc.), the applications of cybernetics (neural networks, systems analysis, operations research, ...) need a general framework to integrate them. Similar to the integrating theories of mathematics at the end of the 19th century (Cantor's set theory, formal logic, ...), the integrating theories of cybernetics at the end of the 20th century (general systems theory, second-order cybernetics, ...) are not integrated themselves.

Both mathematics and cybernetics are in the first place *metadisciplines*: they do not describe concrete objects or specific parts of the world; they describe abstract structures and processes that can be used to understand and model the world. In other words they consist of models about how to build and use models: metamodels (Van Gigh, 1986). Because of this mathematics and cybernetics can be applied to themselves: a metamodel is still a model, and hence it can be modelled by other metamodels, including itself (Heylighen, 1988).

In reference to Russell and Whitehead, the enterprise we propose is called the "*Principia Cybernetica Project*" (Turchin, 1991; Heylighen, Joslyn and Turchin, 1991). The unified framework we wish to develop can be viewed as a philosophical system: that is to say a global "world view" ("Weltanschauung"), which is clearly thought out and well-

formulated, avoiding needless ambiguity, inconsistency or confusion. Starting from cybernetical concepts, it should try to integrate all the different domains of human knowledge, experience, and action. It should provide an answer to the basic questions: "Who am I? Where do I come from? Where am I going to?" Like in traditional philosophy it should contain at least an ontology or metaphysics (a theory of what exists in the world and where it comes from), an epistemology (a theory of how we can know the world around us), and an ethics or axiology (a system of goals and values that can guide us in our actions).

Evolution and Constructivism

In addition to the traditional assumptions of systems theory, based on the principle that *organization* is more basic than substance, we want to start from the principle of *evolution*: systems are not given or fixed, they are the result of a continuing process during which more and more complex forms of organization emerge. This evolution does not have a final goal, it is directed only by the trial and error process of natural selection. Different (re)combinations of systems are formed by variation, but only those combinations are retained that are stable, internally and with respect to the requirements of the environment. The stability of the organization is what turns a mere assembly into a "system". The variation process may be guided by knowledge acquired earlier, but in its most basic form it is *blind*: it does not know where it is going, or which of the variants it generates will be selected (Campbell, 1974). These principles are sufficient as a basis for a complete metaphysics, epistemology and axiology, as will be explained in a further contribution to this workshop.

Such an evolutionary philosophy is also *constructive*: it assumes that systems can only be really understood by analysing the process through which they have been assembled. The variation and selection mechanism continuously constructs new systems from previous, usually simpler, systems. These building blocks themselves have emerged from even simpler components, which are the result of combinations of yet more primitives parts, ... The properties of the system cannot be reduced to the properties of their components: they can only be understood as results of the construction process itself, of the specific way in which the components have been assembled.

In the limit, such a constructive view entails that one cannot be satisfied by a philosophy which is based on "fundamental laws of nature", "first causes" or "prime movers", that is to say on fixed foundations beyond further analysis. Whatever principle or organization is at the base of a construction process, it is itself merely the result of a previous construction and hence cannot in any way be ultimate. The only "primitives" that can be accepted in a constructive philosophy must be so simple as to be *empty of organization*. All others, including

the fundamental laws of physics, are to be viewed as the result of evolution through variation and selection, and must be analysed as such. An example of such an "empty" fundamental is the *tautological principle of natural selection*: stable systems remain, unstable systems are eliminated (Heylighen, 1990).

Examples of foundational ontologies are proposed by Newtonian mechanics, which sees hard, elementary particles moving in space according to deterministic "laws of nature" as the essence of the world, and by the traditional monotheistic religions, which see the world as created and governed by the God.

Such ontologies are not constructive: they explain the presence of properties such as organization, stability, causality, or goal-directedness, by postulating some unobservable fundamental causes (God, the laws of Nature) which by definition already have the properties to be explained. In that way nothing is really explained, the problem is merely pushed one level away, where it cannot be further analysed. Indeed, in these ontologies it is impossible to ask where God (or the Laws of Nature) came from, why He is permanent, why He is intelligent, etc., because these facts are dogmatically or axiomatically established. In that sense, such a position is not scientific, and we might even doubt to call it philosophical, since philosophical thought by definition involves continuing to ask questions.

In this sense, the constructive philosophy we propose is *anti-foundational.* Yet a constructive philosophy can be considered foundational in the sense that it takes the principle of constructive evolution itself as a foundation. This principle is different from other foundations, however, because it is empty (anything can be constructed, natural selection is a tautology), but also because it is situated at a higher, "meta" level of description. Indeed, constructivism allows us to interrelate and intertransform different foundational organizations or systems, by showing how two different foundational schemes can be reconstructed from the same, more primitive organization.

From philosophy to method

The Principia Cybernetica Project is distinguished not only by its philosophy (the "content" of the project), but also by its method (the "form" of the project). In accordance with the principle of the self-application of cybernetics, both form and content will be constructive and evolutionary, based on the development of higher levels of organization through the recombination of simpler subsystems and the selection of those assemblies that are more stable. The development of form and content, of method and theory, will hence occur in parallel, with a continous feedback from the one to the other, so that each new principle in the theory will be reflected in the method to further

develop the theory, whereas each improvement in the method will lead to the discovery of new theoretical principles.

When constructing a cybernetic philosophy the fundamental building blocks we need are *ideas*: concepts and systems of concepts. Ideas, similarly to genes, undergo a variation-and-selection type of evolution, characterized by mutations and recombinations of ideas, and by their spreading and selective reproduction or retention (see the contribution of Moritz to this workshop). The basic methodology for quickly developing a system as complex as a cybernetic philosophy would consist in supporting, directing and amplifying this natural development with the help of cybernetic technologies and methods.

It will require, first, a large variety of concepts or ideas, provided by a variety of sources: different contributors to the project with different scientific and cultural backgrounds. These contributions must be gathered and stored in an easy and efficient way. Therefore we must use the most advanced communication media, in particular *electronic mail*. The collected information can then be kept in store on one or more central computers ("file servers") that can be accessed from anywhere in the network of collaborators. In order to efficiently find and use the information we need a system that allows the representation of different types of combinations or associations of concepts. This can be based on a *hypermedia semantic network*, with different types of *nodes*, containing information in different formats (text, formulas, drawings, ...), connected by *links*. We further need selection criteria, for picking out new combinations of concepts, that are partly internal to the system, partly defined by the needs of the environment of people that are developing the system. Finally, we need procedures for reformulating the system of concepts, building further on the newly selected recombinations. Different ways to implement this kind of interactive structuring and restructuring of concepts in a hypermedia system will be discussed in the section on "computer support systems".

References

Boulding Ken (1956): "General Systems Theory - The Skeleton of Science", *General Systems Yearbook* 1, p. 11-17.

Campbell D.T. (1974): "Evolutionary Epistemology", in: *The Philosophy of Karl Popper*, Schilpp P.A. (ed.), (Open Court Publish., La Salle, Ill.), p. 413-463.

Heylighen F. (1988): "Formulating the Problem of Problem-Formulation", in: *Cybernetics and Systems '88*, Trappl R. (ed.), (Kluwer Academic Publishers, Dordrecht), p. 949-957.

Heylighen F. (1990): "Classical and Non-classical Representations in Physics I", *Cybernetics and Systems* 21, p. 423-444.

Heylighen F., Joslyn C. & Turchin V. (1991): "A Short Introduction to the Principia Cybernetica Project", *Journal of Ideas* 2:1, p. 26-29.

Whitehead A.N. & Russell B. (1910-1913): *Principia Mathematica* (vol. 1-3), (Cambridge University Press, Cambridge).

Turchin V. (1991): "Cybernetics and Philosophy", in: *Proc. 8th Int. Conf. of Cybernetics and Systems*, F. Geyer (ed.), (Intersystems, Salinas, CA).

Van Gigh J.P. (ed.) (1986): *Decision-making about Decision-making: meta-models and metasystems*, (Abacus Press, Cambridge).

von Bertalanfy, Ludwig (1968): *General Systems Theory*, (Braziller, New York).

Cliff Joslyn

Systems Science
State University of New York at Binghamton

General Notes about the Principia Cybernetica Project and Related Initiatives

History of PCP

The Principia Cybernetica project was conceived by Valentin Turchin, a physicist, computer scientist, and cybernetician. He had developed a cybernetic philosophy based on the concept of "metasystem transition", and wanted to further elaborate it in the form of an integrated system with a hierarchical organization, involving multiple authors.

In 1987, Turchin came into contact with Cliff Joslyn, a systems theorist and software engineer. Joslyn suggested a semantic network structure using hypertext, electronic mail, and electronic publishing technologies as strategy for the implementation of Turchin's ideas for a collaboratively developed philosophical system. Together they founded the Principia Cybernetica project and formed its first Editorial Board. They wrote a first proposal, and a "Cybernetic Manifesto" in which the fundamental philosophical positions were outlined. Joslyn began publicizing Principia Cybernetica by posting these documents on the CYBSYS-L electronic mailing list.

This generated a lot of response, including that of Francis Heylighen, a physicist, cognitive scientist, and systems theorist. Heylighen had been developing a very similar philosophy to Turchin's and had been thinking along the same lines of creating a network of people who would communicate with the help of various electronic

media. He joined Turchin and Joslyn as the third member of the editorial board in spring 1990.

Together they started to further develop their philosophical ideas, partly in the form of "nodes" and publications, through elaborate electronic mail conversations, complemented by personal meetings. They continued to attract other people to the PCP idea through several activities: a sometimes heated public debate on the CYBSYS-mailing list (winter 1990), a symposium in the context of the Int. Congress of Systems and Cybernetics (New York, June, 1990), the distribution of a leaflet, followed by the introductory issue of a newsletter, to a mailing list containing journals, associations, electronic newsgroups and individuals active in related domains, and the organization of a workshop in Brussels. This led to the compilation of a continuously expanding mailing list of people interested in collaborating in the PCP.

For the moment a specialized electronic mailing list, PRNCYB-L, is being set up to facilitate the communication among that relatively large group of people. Heylighen and Joslyn are also experimenting with the development of hypermedia systems for supporting the development and organization of PCP concepts.

Relation to similar work

The PCP will be situated in the context of general intellectual history (Talmud, Adler), and the history of systems science and cybernetics. In particular different attempts to do similar work will be mentioned, such as Krippendorf's *Dictionary of Cybernetics*, Singh's *Systems and Control Encyclopedia*, the work of Troncale and Snow in the context of the International Society for Systems Science, the *Glossary on Cybernetics and Systems Theory* developed for the American Society for Cybernetics. A brief overview of links to current development in computer systems (discussed in more depth in the section on computer support systems) will be given.

Valentin Turchin
Computer Science
City University of New York

A Tentative Sketch of the Starting Nodes of PCP

I first discuss the methodology of the construction of a system of nodes where both the contents of nodes, and their relation and organization are tightly interrelated. I propose to use the principle of step-wise formalization, on which the whole edifice of science is built. In

science we start with intuitive and often imprecise concepts and on this basis create new models of the world which are more formalized and more precise. Formalization may go in rounds, or levels, becoming more intensive and extensive. Finally we reach a stage at which we reinterpret those intuitive concepts that were taken for granted at the beginning of the construction. Thus a clock, with its hands, becomes a structure of elementary particles.

This, however, does not make unnecessary the usual notion of a clock, as well as all other simple words we use in explaining physics. This is a hierarchy of pictures of the world, where there are unbreakable ties between levels. Take "simple" notions away, and the whole edifice will crumble. This method can be referred to as the method of *step-wise formalization*.

I propose, therefore, that the nodes we are writing will be initially organized according to the usual notion of their conceptual dependency understood informally or semi-formally (the whole-part relation is also included, of course, as a reason for siblings). As the collection of nodes grows, we give more time to the work on formal semantics and the structuring of this accumulated material.

In this talk I present the result of my first attempt to sketch some basic conceptual nodes of the Principia Cybernetica Project. Needless to say, it is very imperfect, a very rough first outline of the top level of the system. But it should allow us to start a discussion—and work. The present abstract includes a list of nodes with their references up and down. For some of the nodes a brief exposition of contents is provided.

```
node: PRCYB
Principia Cybernetica
references up: none
references down:
INTRO> Introduction
FORMAT> The format of the material
MAIN> The main node
REACT> Reactions, discussions, comments
NETWORK> Network of people
```

This is the head node of the system to which you address when you start examining Principia Cybernetica. The head node includes an introduction INTRO, the FORMAT describing the organization of the material, the main node MAIN which contains subnodes which actually define our philosophical system, and the node REACT which contains reactions to our work and discussions around it.

```
node: MAIN
The main node of Principia Cybernetica
references up:
PRCYB> Principia Cybernetica
references down:
KNOW> Knowledge
WILL> Will
FUTURE> Future
```

The contents of Principia Cybernetica follows the formula:

Our knowledge + Our will = Our future.

In our thought and language we distinguish two different classes of elements about which we say that they exist: those expressing what we know, or think we know, and those expressing what we are striving for and intend to do. We unite the elements of the first class referred to as KNOWLEDGE, and the elements of the second class as WILL. They are not isolated from each other. Our goals and even our wishes depend on what we know about our environment. Yet they are not determined by it in a unique way. We clearly distinguish between the range of options we have and the actual act of choosing between them. As an American philosopher noticed, no matter how carefully you examine the schedule of trains, you will not find there an indication as to where you want to go.

We think about knowledge as a representation of the world in our mind. Representation is the term used by Schopenhauer; the world for him is *Will* and *Representation*.

Another way to describe the relation between knowledge and will is as a dichotomy between not-I and I, or between object and subject. The border between them is defined by the phrase "I can". Indeed, the content of my knowledge is independent of my will in the sense that I cannot change it by simply changing my intentions or preferences. On the contrary, I can change my intentions without any externally observable actions. I call it my will. It is the essence of my 'I'.

The origins of this approach to all that exists are cybernetical. We try to understand ourselves by building cybernetical creatures which model intelligent behavior. The model of intellect such a creature has consists of two parts: a device that collects, stores and processes information; and a decision taker—another device that keeps certain goals and makes choices in order to reach these goals, using the information from the first device. Thinking about ourselves in those terms we speak about knowledge and will.

```
node: KNOW
Knowledge
references up:
MAIN>
references down:
EPISTEM> Epistemology: what is knowledge?
METAPHYS> Metaphysics: what is the nature of things?
CYBER> Cybernetics
MATH> Mathematics
NATURAL> Natural sciences
```

The first part of knowledge is, logically, the knowledge about knowledge itself: what is knowledge? This part is known as epistemology (EPISTEM). Metaphysics, informally, should answer to the question: what is the nature of things? Attempts to understand this question in a more formal way and to give a satisfying answer produced volumes of philosophy. We treat this problem from our cybernetical positions—see METAPHYS. We divided the whole sum of exact sciences into cybernetics (including the theory of evolution), mathematics and natural sciences. It may come as a surprise that there is no place for humanities in this node. They are found in the node Will. This does not mean that humanities do not constitute knowledge—they certainly do. All the texts in Principia Cybernetica, as any texts, represent knowledge; only actions, not texts, represent will. The titles of our nodes must be understood as knowledge about human knowledge, and knowledge about human will. Humanities, as one can see from the word itself, deal with manifestation of human will.

```
node: EPISTEM
Epistemology: what is knowledge?
references up:
KNOW> Knowledge
references down:
MEANING> Meaning
TRUTH> Truth
THEORIES> Theories versus facts
SEMANT> Semantics
HEPIST> Historical review of epistemology
```

In cybernetics we say that a purposive cybernetic system S has some knowledge if the system S has a model of some part of reality as it is perceived by the system. But what is a model? The most immediate kind of a model is a device that implements the concept known in mathematics as homomorphism. After some generalization we come

to the formula: a piece of knowledge is a hierarchical (or recursive) generator of predictions.

```
node: METAPHYS
Metaphysics: what is the nature of things?
references up:
KNOW> Knowledge
references down:
KASCHO> Introduction. From Kant to Schopenhauer.
ACTION> Action
FREEDOM> Freedom
GOD> God
SEMANT> Semantics
```

A metalanguage is still a language, and a metatheory a theory. Metamathematics is a branch of mathematics. Is metaphysics a branch of physics? We argue that, in a very important sense, it is.

```
node: MEANING
Meaning
references up:
EPISTEM> Epistemology
references down: none
```

Our definition of knowledge allows us to further define meaning and truth. When we say or write something we, presumably, express our knowledge, even though it may be hypothetical. Thus to be meaningful, a proposition must conform to the same requirement as a piece of knowledge: we must know how to be able to produce predictions from it, or produce tools which will produce predictions, or produce tools to produce such tools, etc. If we can characterize the path from the statement to predictions in exact terms, the meaning of the statement is exact. If we visualize this path only vaguely, the meaning is vague. If we can see no path from a statement to predictions, this statement is meaningless.

```
node: TRUTH
Truth
references up:
EPISTEM> Epistemology
references  down:  /*  to  cybernetic  foundation  of
math. */
```

A piece of knowledge is true if the predictions made by the user of knowledge on the basis of this knowledge come true.

```
node: THEORIES
Theories versus facts
references up:
EPISTEM> Epistemology
references down: none.
```

```
node: SEMANT
Semantics
references up:
EPISTEM> Epistemology
METAPHYS> Metaphysics
```

This node starts a hierarchy defining the meaning of the most general concepts used in philosophy and science. People usually take them for granted. We want, however, to define them as precisely as possible, and to derive their necessity from the basic principles of epistemology and metaphysics. Success on this way would confirm the validity of our epistemology and metaphysics. The main principle we follow is this: our definition of meaning is tied to the concept of modeling: the homomorphism picture. Therefore we should start any attempt to formalize semantics with an analysis of various aspects of that picture and various types of such pictures.

For the time being I do not break this node into subnodes. This is left for future. At present, the following concepts have been (tentatively) analyzed and defined:
1. State, physical and mental
2. Internal and External Knowledge.
3. Causality
4. Abstraction.
5. Prediction
6. Space and Time.
7. Observation
8. Object
9. Process

By the time of the conference I plan to write the following nodes:
• knowledge about knowledge (model of a model)
• real time versus model time
• historic record (the meaning of historical statements).

```
node: KASCHO
Introduction. From Kant to Schopenhauer
references up:
METAPHYS> Metaphysics
refrences down: none.
```

```
node: ACTION
Action, the ultimate reality
references up:
METAPHYS> Metaphysics
references down: none
```

Am Anfang war die Tat (Goethe.)

Will is manifested in action. If we are looking for the ultimate undoubted reality of physics, we must turn to action, and not to the space-time picture of the world. For a picture is only a picture, while action is an irrefutable reality.

An action is a result of a free choice. The state of the world defines (or rather is defined as) the set of feasible actions for each will. The act of will is to choose one of these. We learn about action through our representations, i.e. our knowledge about the external world.

Donald H. McNeil

Developmental Systemologist
Philadelphia

The Principia Project

After more than fifty years the 'systems movement' still cannot agree upon its subject matter or promulgate a coherent theory. As yet there is not even an operational definition of 'system' acceptable to the majority of stakeholders. The idea of a Principia project to remedy this disgraceful situation has merit. Nonetheless, it matters how the project is put together, what it strives to do, and why.

In the early days of 'systems thinking', people identified systems with forms and formalisms, hence the emphasis on mathematics and hierarchies and structuralism and—more recently—the 79 isomorphies of the ISSS. An alternative movement has placed its emphasis on 'function', 'producer-product', operations research, and living systems. At the same time, some independent thinkers have centered their work on the fluzes of change and the flows of 'substance' as the essences of systems. All the while, the notion that systems could best be understood through cybernetics has attracted a well-organized following. The Principia Cybernetica Project seems now to be allied with this latter view.

It is perhaps appropriate that distinct 'schools' of systems thinking have divided themselves rather neatly into camps representing the four fundamental aspects of any system: form, function, content, and control. These four, together with the timing which inter-relates them, can be woven into the fabric of a General Theory of Systems. A partiality to one aspect, e.g. control, may skew but cannot completely subsume the other aspects. In a systemological schematic such as that in the figure below, [picture deleted] the CONTROL aspect is represented with emphasis but still in balance with the other three.

Even to make the simple drawing above presupposes a considerable amount of generalized systemological conceptualization. The Principia Cybernetica Project as currently described in its Newsletter #0 itself presupposes rather a lot about cybernetics, systems, and philosophy. It is therefore of the utmost importance here to think systemologically about the place, then meaning, and the method of the Project. In these early stages, a Principia still has a chance to reflect upon itself and to establish its mission accordingly. Let us proceed through the highlights of the Project so as to clarify its positions and their relationship to a systemological worldview.

Reference

Donald McNeil, "Systemology: The Fundamentals of a General Science of Systems", 1981. Masters Thesis, U. of Pa.

CYBERNETIC FOUNDATIONS

Gordon Pask
CICT/OOC
University of Amsterdam

The Foundations of Conversation Theory, Interaction or Actors Theory, all Cybernetic and Philosophically so (or 'Some Foundations of Principia Cybernetica')

The intention of writing this paper is to show, albeit in a blinkered and limited manner, that a philosophy of Cybernetics, encapsulated in the journal title, "Principia Cybernetica" is not only justifiable, but necessary and in this day and age, utterly essential. Have no qualms, it IS and that statement IS significant. We need only to look out at this world, lying amongst many others of similar and different kinds, to recognise this fact and, in doing so, to see that factuality is fragile, parading, like a circus-procession, or a civic, mayoral one with a lady drum-majorette in front, out of mind, into thought, from that into utterance and inscription in words and text books.

One route of demonstration is by way of argument, to see that "proof", for example, is a convention and not something sacrosanct. In order to grasp this point, it is necessary to accept the existence of certain deeply embedded disciplines, momentarily, at least. That these strange distinctions are fallible becomes evident, yet they are hard to dispel, though they must be dispelled, with the exception, perhaps, of the entrenched establishment of logicians, mathematicians, psychologists and others. In such cases one is more likely to do more harm than good, for their practitioners, logical text-book-writers, mathematicians and others are folk who have invested much in terms of effort or sheer labour, who wish to retire and, most surely, wish to do nothing which is at all novel, nothing that might disbalance the boat of their nicely equilibrial status quo.

This mode of argument can be given, for instance, but rather minimally so, by comparing and contrasting "proof by reductio ad absurdum" (which calls paradox a tautology, or a contradiction or else pushes it under the carpet, as disturbing the status balance) with the

proof form (I prefer demonstration), often known as "productio ex absurdo", which uses the interesting fact of paradox as the enticement to creative thought, new theorems, new ideas. These are minimal forms, it would not be difficult to cite one hundred or one thousand or any, possibly countable, number of more complicated, more illuminating, others.

Another method is, of course, by means of force majeur, of missiles and mortars, taken as a duly academic metaphor. In order to pull out the plug in the bathtub of science and philosophy, to empty the tub of bathwater without disposing of the baby, also.

So, what do we do?? With good reason, after much contemplation, I submit that the bathplug is called "time" and that the baby we retain is called "innovation", creativity if you prefer that term. Of course, the operation is possible and bound to succeed. It is, however, apparently destructive and I simply hate destruction or demolition of any kind. As a result I have a preference for a milder and more subtle approach, showing the multiplicity, the plurality of time and the many facets of innovation in the slightly more restricted field, still thoroughly Cybernetic and Philosophical, of Conversation Theory, Interaction of Actors Theory, and the protologic or protolanguage which they share, Lp, by name. In the sequel, this is the line pursued.

Notice, all the same, that the basic and partly enunciated theme may be expanded, like a hydrogen balloon, into the entire extravaganza. Let us be clear about this much, at least. We speak of theories, namely, C.T. (an abbreviation of "Conversation Theory") of I.A.T. (an abbreviation of "Interaction of Actors theory" and a deliberate word play or pun upon the much popularised I.A. theory-or-not, but spawned from, rather than being the ancestor of, Cybernetics itself) . There are more valid surrogates for Cybernetics under any label you elect to take up as your own, freely chosen, particularity; there have been many, like Bionics, Information Science, General and Special System Theory, heaven knows what else; you choose whichever you like, it matters not a tittle or jot. What I shall say, here summarise, remains invariant under any choice whatsoever.

Some 40 years past, in the context of the theatre, the laboratory and academia as viewed by a research assistant, it became evident, to me at any rate, that the search for a"scientific psychology "or a social science was but a fruitless endeavour if we persisted in those still prevalent habits of apeing the scientific by applying "scientific methods", like statistical techniques, to entirely inappropriate data. In place of that, our group proposed and pioneered several other frames of reference, anchored, for credibility, so far as possible upon the existing paradigms adopted by science.

What does it mean to have a "Scientific Psychology", or to have a "Science of Society"; that is, over and above the pseudo-sciences of

inappropriate data, smudged-splodged into a format which is super-ficially compatible with sciences where the real data can, for example, be treated statistically as well specified and independent event reckon-ings? Clearly, dependencies may exist. Clearly, also, such more liberal dependencies can be accounted, if they ARE so simplistic, in a manner not dissimilar to the bookkeeper's ledger, to be dealt with by accoun-tants and actuaries and the like. Lamentably, for some, or joyfully for most of us, neither psychological nor social events; call them mental events, are NOT so simple and cannot be recorded or manipulated in the ways suggested.

So what, being Cybernetists, are the scientific foundations of the mental events with which we are so often, some of us most frequently, apt to deal? What justifies the scientific flavour of the appellation "Principia", as in Newton's or Russell's "Principia"? There are, of course, many possible replies to this rhetorical question but, in this paper, I develop only one of them.

Let it be taken for granted, (failing which, you are welcome to a tedious but more-or-less irrefutable demonstration of the fact), that Cybernetics is a coherent and cohesive theoretical structure, this, in particular, being the case for the so-called New-Cybernetics. Further, let it be taken for granted that Cybernetics, a fortiori the New-Cybernetics and (not so much the System-Thinking stuff) is suffi-ciently distinct to have an identity of its own, even though it promotes interaction, itself, and may be regarded as positively engendering in-teraction between superficially disparate fields.

So it appears as a coherent and cohesive system of analogy, of metaphor but strict metaphor, designating analogies in which the similarities and the differences are well specified. One asks, quite naturally, why this should be deemed a "science" with pretensions to having firm "principles", rather, for example, than an art or a philosophy or the logical aspect of a theology.

On this score, of being definitively "scientific", I am not so deeply convinced as I am, dogmatically so, of the plain fact that Cybernetics most surely DOES have PRINCIPLES which, for all their global breadth, maintain integrity. Perhaps that is because I am not so convinced about any significant differences between, say, science and art, believing that they must coexist together if either one or the other is to make sense. However, it can be strongly argued that the principles of Cybernetics resemble those of such disciplines as physics, biology, cosmology, chemistry, molecular biology, microphysics, archeology, social anthropology and geology. If the ossature of these disciplines is deemed to be scientific, then, presumably, Cybernetics is scientific.

Upon these slightly tenuous grounds, let us survey some of the structural similarities at hand. The list is by no means exhaustive at this moment, and it is evolving.

(1) In real, rather than school-science, we seek appropriate hypotheses and data, entities over which practitioners may agree or agree to dis-agree and know why they do so. Admittedly, in school-science, many of us were TOLD that the testable hypotheses emerged from great the-ories and that some even greater theory will be revealed, but not until next year. Also, most of us were TOLD that scientific data are repeat-able, objective, causally mapped in progression as objects and, later, unmentionable until next term, events.

Admittedly, if we were fools enough to accept these half true falsities as anything other than the infrastructure of an elaborate, even if cost effective examination process, then we may still entertain deeply ingrained concepts of an unduly naive picture of science. But real and mature science is not, at all, like that. Quite obviously, SOME, not ALL, data of reaction kinetics are inappropriate to psycho-social-mental events. The search for "hard" scientific data branches in different directions, appropriate to the field of enquiry, plural in any field of enquiry. Thereby, hypotheses are posed, formulated, tested by appropriate data, inductively verified or deductively falsified and theoretical structures erected.

Cybernetics admits, maybe preaches, all this. It also asserts that the kind and the truth functional modality of the logics underpinning science are varied, like the appropriateness of the domains which their logics generate. Fundamentally, they are logics of many-sorted coher-ence, of many-sorted distinction, of self reference and other reference, all of them are dynamic. The nowadays standard Aristotelian view, is not denied. It is reified, locally, in those regions of a manifold where there are valid metric space type representations. In this respect, at least, the structure of Cybernetics resembles the structure of science.

(2) It may be demonstrated, with passable elegance, that Cybernetics shares, with science, certain skeletal principles. In many places, at least, these skeletal attributes lie in one to one, isomorphic, correspon-dence. In other places, the correspondence is, more likely, homomor-phic and in others, it may only be expressed by the category theoretic relations of functors between categories, or their topological equiva-lents. However, so far as I know, there is no basic dissonance. Amongst the principles involved are conservation, complementarity, duality, exchange relations, parity, symmetry and symmetry breaking, uncertainty, indeterminacy and the obvious mathematical or meta-mathematical properties of distinction, of knottedness, of singularity in contrast to continuity, of various types of demonstration, some

being proof theoretic and others not so. To these it is necessary to add a few others such as the void, the not void, the self and the other. Science, itself, might benefit by their proper inclusion within its orbit.

Thus, in the classical sciences, we commonly revere the conservation of mass and of energy, under the elegant equivalence of $E = mc^2$, c being the limiting velocity of light, E = Energy and m = Mass, we have, in Cybernetics, several conservations such as application (of procedures, complementary to products), of procedures acting upon procedures (to produce and incidentally, reproduce them), of meaningful information transfer and of distinction. For sure, they are not so neatly related. But that is hardly surprising, once you keep in mind the scope of Cybernetics which is so much more encompassing than that of classical science, for instance, adumbrating scientists and the theories they develop.

The foregoings notions are intended to illustrate an evolutionary trend in Cybernetics with which I am, personally, very familiar. The train of thought could be extended further backwards and forwards (though "backwards" and "forwards" are terms up for question in that self-same framework). However, using these terms in the common language sense, (stripped, that is, of particular formalities), I shall try to go by interpolation and by extrapolation in each direction especially into what is, often brashly, called the future and in serious Cybernetic Discussion, open to serious discussion.

Lars Löfgren

Department of Information Theory
University of Lund, Sweden

Foundational Issues Addressed by Cybernetics

With its wholistic aims and its understandings of self-reference, cybernetics addresses issues which have proved foundational not only for sciences and information technologies but also for cybernetics itself.

Within the hard sciences, like physics, foundational issues concern justification problems. Which in classical physics are resolved by observation and measurement–with a belief in a "detachable observer". In quantum mechanics, which includes the measurement process, the justification problem becomes severe, requiring a more abstract form of justification in terms of knowledge of observability versus definability.

In cybernetic studies of knowledge of knowledge processes, the insight is gained that such knowledge must be relativized to language,

and that the "detachable observer" be changed into a thesis of a "non-detachable language". This enforces the autological predicament—to conceive of language in language—for which a resolution in terms of a complementaristic conception of language has been proposed. The complementarity may be conceived from various views. One is as a tension between describability and interpretability within a language. Another, in terms of degrees of partiality of self-reference within a language (where the impossibility of a complete self-reference is synonymous with the "non-detachability of language"). In cases where an object language has a metalanguage, the complementarity of the object language is describable in the metalanguage (but not in the object language). The complementarity is then said to be transcendable, and the self-reference problem, that of describing a language in the language itself, is "unfolded" (a characteristic cybernetic justification of self-reference).

In particular, we discuss the Bohr-Pauli dialogue on a detachable observer, and suggest complementaristic linguistic models for the self-referential measurement problem in quantum mechanics.

We also suggest such linguistic models for the foundational problems of probability theory, namely of how to conceive of models for probability—which, as has been observed in particular for Kolmogorov's axiomatic approach, are not describable within the theory. We attach to Josephson's view, concerning strategies of science towards form versus meaning, that "the technique of statistical averaging is especially irrelevant in the context of meaning, since its influence in general is to transform the meaningful into the meaningless".

The problem of induction, foundational for most sciences, obtains a natural explanation in the complementaristic conception of language. We suggest that quests for inductive inferences of general laws from particular observations are, and will forever be, in vain. Instead, an inductive inference is conceived as a linguistic (mostly unconscious) process which utilizes not only particular observations but also properties of the language which are beyond describability (hidden) in the language itself. Thus, to be able to describe induction, as it occurs in a language, we must have access to a metalanguage in which the object language is describable. In reality, languages are themselves not produced from descriptions, but are evolved.

The foundational problem of describing evolution, in biology as well as in epistemology, is again conceived in terms of the complementaristic conception of language—this time genetic language. In particular, we are able to give a metamathematical argument for the higher force of an evolutionary process than that of a planning process based on inductively generated descriptions in a scientific language.

The impossible task of aiming at a complete description, in some language, of the biological process of evolution is as we know replaced by aiming at less ambitious goals. To a certain extent such goals can be analyzed in terms of goals on higher levels. But the impossibility of reaching a complete description, enforces a goal hierarchy with ultimate goals that are exempt from scientific analysis—like ethical goals.

We analyze Moore's Principia Ethica and his concept of "naturalistic fallacy". In particular we illuminate fallacies in trying to base ethics on evolution.

Reference

Löfgren, Lars (1991): "Complementarity in Language; toward a general understanding," in Carvallo, M (ed.) Nature, Cognition and System II, Theory and Decision Library, Dordrecht-London-Boston: Kluwer, 73-104.

Ranulph Glanville
University of Amsterdam and Portsmouth Polytechnic

Excavation and Underpinning
Foundation and Building

In a number of earlier publications, I have examined both the nature of fundamentals (in a belief system or thesis), and some of the fundamental concepts of cybernetic systems (such as control, communication, variety, responsibility, distinction, recursion and re-entry), especially in the light of, and as generating, second order / the new / the cybernetics of Cybernetics.

In this paper I shall systematically consider the intension and extension of other fundamental concepts from (especially Ashby's) early writings in Cybernetics, both to consider of what they are made, and upon what they rest, and to see how this casts them in a new light, particularly in view of the insights we have gained in and through second order / the new / the cybernetics of Cybernetics.

Thus, to use the architectural metaphor implicit in the title (the firmnesse of Webb's original translation into English of Vitruvius's classic definition of architecture—firmness, comodotie and delight), I examine the

> **Excavation and Underpinning**
> **Foundation and Building**

of cybernetics.

Gertrudis Van de Vijver
Senior Research Assistant NFSR
University of Ghent

Error: Epistemological Options in Cybernetics

Error plays an important role in the ascription of teleological properties and capabilities to systems. It is possible, on the basis of the meaning and the place of error, to trace out the history of purposiveness. We aim at doing this in going through the different cybernetic stages—the cybernetics of the first, the second and the third order—and through the theories which were inspired by cybernetics—connectionism and neo-connectionism—.

In first order cybernetics, and in most A.I. views, error is to be interpreted in terms of the dysfunctioning of systems. Goal-directed behavior is always to be interpreted on the basis of a 'goal-deficiency' model. Difficulties of the missing goal-object, problems of circularity between the goal and the relevant behavioral properties, arise in this context.

In an attempt to model certain properties of complex purposive systems, it became clear that the possibility to behave in an erroneous way had to be build in. The possibility of error is in this case linked with the possibility of building up a representation in an inductive way. It is also brought in connection with a relation of under-determination existing between a behavior (an idea, a theory) and certain conditions preceding it.

How will the possibility to behave erroneously in this case be evaluated ? How shall we make the possibility of going through a history, a history that is characterized by an under-determination, into an integral part of an artificial system ? What are the epistemological consequences of this ? We are confronted here with specific epistemological difficulties which have to do with the knowability of autonomous or self-organizing systems.

EVOLUTIONARY PHILOSOPHY

Valentin Turchin
Computer Science,
City University of New York

Metasystem Transition as the Quantum of Evolution

Consider a system *S* of any kind. Suppose that there is a way to make some number of copies from it, possibly with variations. Suppose that these systems are united into a new system *S'* which has the systems of the *S* type as its subsystems, and includes also an additional mechanism which controls the behavior and production of the *S*-subsystems. Then we call *S'* a metasystem with respect to *S*, and the creation of *S'* a metasystem transition. As a result of consecutive metasystem transitions a multilevel structure of control arises, which allows complicated forms of behavior.

In my book [1], I show that the major steps in evolution, both biological, and cultural, are nothing else but metasystem transitions of a large scale. The concept of metasystem transition allows us to introduce a kind of objective quantitative measure of evolution and distinguish between evolution in the positive direction, progress, and what we consider an evolution in the negative direction, regress. In the present paper I outline the main ideas of this book, and concentrate, in particular, on one of the aspects of biological evolution: the appearance of human thinking.

When we speak of cybernetic systems, we can describe them either in terms of their *structures*, or phenomenologically, in terms of their *functioning*, their behavior. We cannot claim at the present time that we know the structure of the human brain well enough to explain thinking as the functioning of that structure. However we can observe evolutionizing systems and make conclusions about their internal structure from a phenomenological description of how they function.

From the functional point of view the metasystem transition is the case where some activity *A*, which is characteristic of the top control system of a system *S*, becomes itself controlled as a metasystem transition from *S* to *S'* takes place. Thus the functional aspect of metasystem transitions can be represented by formulas of this kind:

control of A = A'

When a phenomenological description of activities of some systems fits this formula we have all reasons to believe that this is a result of a metasystem transition in the physical structure of the systems. Here is the sequence of metasystem transitions which led, starting from the appearance of organs of motion, to the appearance of human thought and human society:

control of position = movement
control of movement = irritability (simple reflex)
control of irritability = (complex) reflex
control of reflex = associating (conditional reflex)

control of associating = human thinking
control of human thinking = culture

In [1], I show how the most characteristic features of human thinking: creation of tools, imagination, planning, overcoming the instincts, understanding of the funny and the beautiful, creation of language, self-knowledge, can all be understood as control of associating and its direct consequences. Then the principle of metasystem transition is used for an analysis of cultural evolution, and first of all, the development of science. We see that in the history of science, as well as in the history of biological evolution, the major steps forward are done through metasystem transitions. Looking even farther, we can try to guess (and at the same time influence) the more remote stages of the evolution of the mankind.

Reference

[1] Turchin V. (1977): "The Phenomenon of Science" (Columbia University Press, New York)

Cliff Joslyn
Systems Science
State University of New York at Binghamton

Control Theory and Meta-System Theory

This paper, as part of PRINCIPIA CYBERNETICA, is intended to be integrated into the structure of that project. Therefore, we note potential links to the following nodes:

Action, Behavior, Constraint, Constructivism, Control, Control System, Dreaming, Dynamic Equilibrium, Emergence, Equilibrium, Evolution, Feedback, Freedom, Goal, Hallucination, Hierarchy, Imagination, Intention, Knowledge, Life, Memory, Purpose, Selection, Self-Organization, Stability, Thought, Variation, Will

The 1970's produced (at least) two great cybernetic meta-theorists: Valentin Turchin and William Powers. In *The Phenomenon of Science* [TUV77] and *Behavior: The Control of Perception* [POW73], respectively, they provide grand biological and psychological theories resting on common principles: that evolved organisms are hierarchically organized belief-desire control systems; that these cybernetic systems are involved in cyclical modeling relations with their environments; that blind variation and selective retention is a universal mechanism of

both biological and non-biological evolution; and that a consequence of these views is that human freedom is necessary for social evolution.

While Turchin and Powers differ on the nature of control, and particularly the origin of control systems, they share the great majority of a theoretical core. Much of their theories are not unique in Cybernetics and Systems, or in general. Indeed, the key aspects of their theories (e.g. the use of "hierarchy", "control" and "purpose") are central to all of Cybernetics and Systems (e.g. [ASR52, ASR56]). But in their work, these ideas have been developed in conjunction, and have been successfully extended to produce elegant, consistent, general theories of living systems in the context of Cybernetics and Systems theory.

In this paper we will examine Powers' Control Theory [POW73, POW89]. We will do so from the perspectives of: Turchin's work, as expressed in the works of the PRINCIPIA CYBERNETICA project to date—with which we assume the reader is familiar [HEF90f, HEFJOC90, JOC88e, TUV77, TUV81, TUV87a, TUV90 ,TUV91a, TUV91b, TUVJOC90]—and which we will call (for want of a better term) "meta-system theory"; and the wider theories of evolving systems as developed by the Cybernetics and Systems disciplines.

Powers' Control Theory

Powers' central thought is simple: all living systems are hierarchically organized negative feedback control systems, where "feedback control system" is essentially the same concept as that used in Control Engineering for the design of regulatory mechanisms [MAO70, WIN48]. Thus, as in classical Cybernetics, the simplest regulatory mechanisms, like thermostats, are prime examples. However, Powers' intent is to claim that this theory of *machines* has *universal* applicability to organisms. Thus control theory is an attempt to return Cybernetics and biology to each other, as Cybernetics has lost biology for engineering; and even the most sophisticated forms of theoretical biology [EIMSCP79, NIGPRI89, VAFMAH74] have lost all concept of fundamental control mechanisms as being the essence of life.

The following is an extremely terse outline of Control Theory.

Definition of Control

Control of an entity requires *constraint*, that is a *selection* or *reduction in variety* of the possible states q of that entity. Typically, the exercise of control will reduce the variation of possible states q to one, thus determining the final state q^* of the system.

But constraint is not sufficient for control. Constraint and determination results from a variety of situations, many of which are not control, the primary of which are *stable equilibria*. For example, supply-demand interaction in markets stabilizes prices, but Adam Smith's

"invisible hand" does *not* "control" the market. Rather, control requires a constant and on-going interaction of the controller with the controlled entity, such that continued constraint results in sufficient stability around a state q^* (or another kind of attractor) despite perturbations and disturbances. Thus systems maintained at an *unstable equilibrium*, such as an inverted pendulum or balanced broom, are exemplars of control systems. Thus we arrive at the definition of control as offered by Rick Marken:

> A controlled event is a physical variable (or a function of several variables) that remains stable in the face of factors that should produce variability. [MAR88]

Control as a Phenomenon

Since any dynamical system which is being maintained at a state (or attractor) which is out of equilibrium is under control, control theory legitimately encompasses a great swath of current interesting work in Cybernetics and Systems—in particular: most of the so-called "self-organizing systems" theories, "far-from-equilibrium" physics [PRINIG72], synergetics [HAH78], and those biological theories which focus on "metabolic" definitions of life [SC67].

Powers' Control Systems

The classical feedback control system is described by Powers as a "stimulus-response", or S-R feedback controller. The topology of the system is a throughput device with two inputs and one output, and an internal loop. A feedback control system of Powers' design has the topology of a whole *closed* loop with *two* inputs, the environmental disturbances and the reference level, and *no* outputs.

Powers uses the following terminology:

Physical Quantity: That aspect of the environment whose variation is eliminated in the face of disturbances, as in our definition.

Disturbance: Environmentally induced fluctuations of the physical quantity.

Output Function: The action, or behavior of the control system.

Error: Signal internal to the control system which directs behavior.

Comparator: Determines whether the perceived variable matches the reference level, and generates an error signal if it does not.

Perceived Variable: The "appearance" of the physical quantity to the control system. In neural organisms this is a "sensation" or "perception".

Reference Level: Similar to the role of the set point in an S-R controller. This signal represents the controlled state of the perceived variable, or that state of the perceived variable which produces no error.

Input Function: How the physical quantity is transduced in the control system into the perceived variable.

Powers' argues that his view is superior to the S-R model in that the closed environmental loop of his model is always implied in an S-R model. S-R models purport to be control systems, but lack the controlled quantity, the entity whose variation is eliminated despite environmental disturbance, and is in the environment.

Constructivist Epistemology

Powers adopts a revolutionary view of control, in the context of constructive epistemology, through two steps. First, we note that the controlled quantity is in the environment, and assert as false the traditional control theoretic idea that the output (behavior, action) of the system is the quantity under control. The variation of action is rather large, on the same order as the variation of the disturbance, and of opposite magnitude, in order to cancel out the effect of the disturbance on the controlled quantity.

Second, we note the necessity that the input function mediates the appearance of the environment to the comparator. We have to say that for the control system, aspects of the environment only exist to the extent that corresponding input functions exist, and we can effectively say that perceptions *are* the environment for the control system. Assuming that the input functions are "good"—in the sense of providing a relatively strong homomorphic mapping or model of the environment, albeit of selected aspects—then when the variation of the physical quantity is eliminated, then so is the variation of the perceived variable. Thus, control of the physical quantity results in effective control of the perceived variable.

Since output is not controlled, and even the physical quantity is not controlled, since the physical quantity only exists for the organism in virtue of mediation through perception, we arrive at the revolutionary idea that it is the *perception* that is in fact controlled; that, for the organism, it is the *input* which is in fact the controlled quantity. Thus, the novel title of Powers' first book [POW73]: it is not, as the received behaviorist tradition would have us believe, that perceptual stimuli allow an organism to correctly control its behavior, but rather the organism's behavior which allows it to correctly control its perceptual stimuli.

Hierarchical Control

Control systems are hierarchically nested when the output function of a "higher" control system does not affect the physical quantity, but rather serves to set the reference level of a "lower" one. The higher level system has the lower level as its environment, and, speaking

loosely, controls the lower level system. The multi-level control system has the same topology as the single level: a closed loop with two inputs. The lower level system must necessarily act at a faster temporal scale than the higher level.

Powers identifies nine levels of hierarchy in human control systems, which can be outlined according to the following schema:

"Systems concepts" require	*control of*	*principles;*
which require	*control of*	*programs;*
which require	*control of*	*relationships;*
which require	*control of*	*sequences;*
which require	*control of*	*transitions;*
which require	*control of*	*configurations;*
which require	*control of*	*sensations;*
which require	*control of*	*intensities.*

A "systems concept" is a unifying conceptual and ideological system of thought, such as a religion, or the "scientific method". The hierarchy extends downward towards more specific *perceptual* categories, since in control theory it is *perceptions* that are controlled, not actions.

Learning and Organization

Learning and change is provided by introducing a meta-control system which stands aside the entire control hierarchy. While the perceptual control hierarchy acts in real time, and is the result of learning, this second, "organizational" layer acts on the perceptual layer over a longer time than the behavior, and affects changes on the perceptual control system—in short, learning. This "organizational system" is genetically innate, and unchanging itself.

The entire perceptual control system is stimulated by and acts on the environment, but the environment also makes physiological affects on the "intrinsic", or physiological state of the organism. Genetically determined structures produce intrinsic perceptual signals, such as hunger, thirst, lust, and pain; but also emotions such as satiation, satisfaction, joy, anxiety, etc. Either positive or negative signals could require learning, to increase or decrease the intrinsic perception respectively. This is mediated through an intrinsic error signal. Output of the organizational system is directed at the perceptual control system, and produces change in it. This change can be either random, blind variation, or somewhat directed (meta-learning). In either case, a null intrinsic error level results in selection and retention of the new configuration.

Memory and Imagination

Powers asserts that memory is uniformly distributed not only among all levels of the control hierarchy, but also in each control system at each level. Thus, a sensational-level control system might "memorize" a color; while a program-level control system might "memorize" a Bach etude. The simpler model of the control system is now modified so that memory is addressed from the output of an upper-level control system, and provides the immediate reference signal to the control system. Perceptions in turn are stored in that memory.

The final additions are two switches on both the reference signal and the perceived variable. Each switch can be either off or on. When the perception switch is on, perception proceeds normally; when it is off, it is the memory signal which is transferred to higher levels. When the memory switch is on, action proceeds normally; when it is off, action is disabled, but a signal of the memory is transferred to the perceptual signal, if it is prepared to receive it. There are four cases:

		Input	
		On	*Off*
Output	*On*	Control	Automatic Actions
	Off	Observation	Imagination

At various levels, "imagination" can be sleep, dreams, hallucinations, or thought.

Challenges from Control Theory to Meta-System Theory and vice versa

In considering control theory from the perspective of meta-system theory and PRINCIPIA CYBERNETICA, and vice versa, we are first very pleased with the opportunity to examine a full-fledged cybernetic and evolutionary theory which is very similar in spirit, but not in detail, to meta-system theory. The interaction of these two schools of thought, and their independence from each other, must continue, to the advantage of both.

For example, consider the great similarity, yet also the great conceptual differences, between Powers' perceptual control hierarchy and Turchin's evolutionary control hierarchy:

Culture is	*control of*	*thought;*
which is	*control of*	*associating;*
which is	*control of*	*complex reflex;*
which is	*control of*	*simple reflex;*
which is	*control of*	*movement;*
which is	*control of*	*position.*

Aside from the fact that Turchin's hierarchy is in terms of actions, while Powers' is in terms of perceptions, there is clearly a similar intent behind each one: to provide a consistent and elegant cybernetic treatment of organisms from the perspective of control hierarchies. Undoubtedly both require significant revision, but the overall program remains clear.

More specifically, we will consider some points of comparison:

Control Specifics: A great advantage that control theory holds for meta-system theory is that it greatly specifies and clarifies what is meant by the definition of the meta-system transition: that "the top control system of a system becomes itself controlled" [TUV91a]. Powers gives this an operational definition, and allows meta-system theorists the opportunity to match their more abstract, philosophical theory more closely to the phenomena as revealed by our specialist colleagues.

Evolution: Both meta-system theory and control theory adhere to Campbell's [CAD74] view of "blind variation and selective retention" as a universal mechanism for all kinds of evolution, including genetic, learning, and social development. But an advantage of meta-system theory is that it is *primarily* interested in these evolutionary steps, and intends to explain the evolution of all emergent levels. Thus, it asks the questions: what are the "essences" of physical phenomena, life, genetics, sex, multi-cellular organisms, social organization, and intelligence? Although in his later works [POW89], Powers is expanding to consider social organization and the origins of life as control phenomena, this is being done in a somewhat piecemeal manner. Although ultimately the conceptual unification of control theory should succeed, in terms very similar to meta-system theory, meta-system theory pursues these subjects at its most basic task.

The Uniquely Human: What explains humans as unique animal forms? Hunger and lust are output functions of the organizational system, and provide inputs directly to the relationships, or perhaps programs level. Which level is unique to humans? How can its origin be explained in evolutionary terms? How can linguistic ability, humor, and aesthetics be explained in control theory terms? These are all addressed directly by meta-system theory, but are underdeveloped in control theory.

Imagination: Meta-system theory would agree with control theory that imagination is central to the selection of goal states, and describes these as acts of Will. But meta-system theory asserts that imagination is unique to humans, and this ability to control associations of mental representations is the essence of intelligence. But in Powers' model both imagination and memory are inherent at all levels of the perceptual control hierarchy. Perhaps there is empirical evidence to support one view over the other, or a conceptual unification of the two.

Meta-System Transitions and Ultra-Meta-System Transitions: Clearly, from the perspective of meta-system theory, each level of the control hierarchy indicates a meta-system transition. But meta-system theory also involves ultra-meta-system transitions, which are incorporations of entire meta-system hierarchies in another at a qualitatively higher dimension, allowing unlimited replication of the now lower level meta-systems. It seems that Powers' organizational system is an ultra-meta-system transition, yet applied to only a single meta-system hierarchy. Further, especially the transition to thought and rationality (the program and principle levels?) should allow these kinds of ultra-meta-system transitions, which are evidenced in human linguistic systems and their unlimited ability to generalize [TUV77].

References

[ASR52]   Ashby, Ross: (1952) *Design for a Brain*, Wiley, New York

[ASR56]   Ashby, Ross: (1956) *An Introduction to Cybernetics*, Methuen, London

[CAD74]   Campbell, Donald T.: (1974) "'Downward Causation' in Hierarchically-Organized Biological Systems", in *Studies in the Philosophy of Biology*, eds. F.J. Ayala and T. Dobzhansky, U. California Press, Berkeley

[EIMSCP79] Eigen, M, and Schuster, P: (1979) *The Hypercycle*, Springer-Verlag, Heidelberg

[HAH78]   Haken, Herman: (1978) *Synergetics*, Springer-Verlag, Heidelberg

[HEF90f]   Heylighen, Francis: (1991) "Cognitive Levels of Evolution", in: *Proc. 1990 Int. Congress of Systems and Cybernetics*, ed. F. Geyer, Intersytems, Salinas, CA

[HEFJOC90] Heylighen, Francis; Joslyn, Cliff; and Turchin, Valentin: (1991) "A Short Introduction to the PRINCIPIA CYBERNETICA Project", *Journal of Ideas*, v. 2:1, pp. 26-29

[JOC88e]   Joslyn, Cliff: (1988) "Review of the Works of Valentin Turchin", *Systems Research*, v. 4:4, pp. 298-300,

[MAR88]   Marken, Richard S: (1988) "The Nature of Behavior: Control as Fact and Theory", *Behavioral Science*, v. 33, pp. 196-206

[MAO70]   Mayr, O: (1970) *Origins of Feedback Control*, MIT Press, Cambridge MA

[NIGPRI89] Nicolis, G, and Prigogine, Ilya: (1989) *Exploring Complexity*, WH Freeman, New York

[POW73]   Powers, WT: (1973) *Behavior, the Control of Perception*, Aldine, Chicago

[POW89]   Powers, WT ed.: (1989) *Living Control Systems*, CSG Press, Gravel Switch, Kentucky

[PRINIG72] Prigogine, Ilya, and Nicolis, Gregoire: (1972) "Thermodynamics of Evolution", *Physics Today*, v. 25, pp. 23-28

[SC67]   Schrodinger E.: (1967) *What is Life?*, Cambridge U., Cambridge

[TUV77]   Turchin, Valentin: (1977) *The Phenomenon of Science*, Columbia U., New York

[TUV81]   Turchin, Valentin: (1981) *The Inertia of Fear and the Scientific Worldview*, Columbia U. Press, New York

[TUV87a]   Turchin, Valentin: (1987) "A Constructive Interpretation of the Full Set Theory", *J. of Symbolic Logic*, v. 52:1

[TUV90]   Turchin, Valentin: (1990) "Cybernetics and Philosophy", in: *Proc. 8th Int. Congress of Systems and Cybernetics*, ed. F. Geyer, Intersytems, Salinas, CA, NOTE: Draft

[TUV91a]   Turchin, Valentin: (1991) "Metasystem Transition as the Quantum of Evolution", in: *Workbook of the 1st PRINCIPIA CYBERNETICA Workshop*, Heylighen F. (ed.), Principia Cybernetica, Brussels-New York

[TUV91b]   Turchin, Valentin: (1991) "A Tentative First Sketch of the Starting Nodes of PCP", in: *Workbook of the 1st PRINCIPIA CYBERNETICA Workshop*, Heylighen F. (ed.), Principia Cybernetica, Brussels-New York

[TUVJOC90] Turchin, Valentin, and Joslyn, Cliff: (1990) "The Cybernetic Manifesto", *Kybernetes*, v. 19:2-3

[VAFMAH74] Varela, FG, and Maturana, HR et. al.: (1974) "Autopoiesis: The Origin of Living Systems, its Characterization, and a Model", *Biosystems*, v. 5, pp. 187-196

[WIN48]   Wiener, Norbert: (1948) *Cybernetics*, MIT Press, Cambridge

Francis Heylighen
Free University of Brussels

Evolutionary Foundations for Metaphysics, Epistemology and Ethics

A process metaphysics

Philosophies traditionally start with an ontology or metaphysics: a theory of being in itself, of the essence of things, of the fundamental principles. In a traditional systemic philosophy "organization" might be seen as the fundamental principle of being, rather than God, matter, or the laws of nature. However it still begs the question where this organization comes from. In a constructive systemic philosophy, on the other hand, the essence is the *process* through which this organization is created.

There have been several attempts at building a process metaphysics, by philosophers such as Whitehead (1926) and Teilhard de Chardin (1959). However, these early process philosophies are characterized by vagueness and mysticism, and they tend to see evolution as goal-directed, guided by some supraphysical force, rather than as the blind variation and selection process that we postulate. They are thus not constructivist in the radical sense as defined in my first paper in this book.

The ontology we propose would start from elementary actions or processes, rather than from static objects or particles. These processes are the primitive elements, the building blocks of our vision of the universe, and therefore remain undefined. In fact they can be modelled in such a way that they are in themselves completely empty (Heylighen, 1990). Relatively stable "systems" are automatically constructed by such processes through the mechanism of blind recombination and selective retention of stable combinations (Heylighen, 1991b).

This leads to a *self-organizing* evolution of the universe as a whole. It is characterized by the spontaneous emergence of more complex organizations (cf. Simon, 1962) during evolution: from space-time and elementary particles, to atoms, molecules, crystals, dissipative structures, cells, plants, animals, humans, society, culture... In this hierarchy of system types (Boulding, 1954), cybernetic models typically start from about the level of thermostats or dissipative structures. Yet a constructive systemic approach can also be used at a much lower level, for example to reconstruct the elementary structures of space-time (Heylighen, 1990), or the fundamentals of set theory (Turchin, 1987). A reconstruction of the most important stages in this

global evolution should allow us to answer the questions: "Where do I come from? Who am I?"

Processes of emergence are the "quanta" of evolution: discontinuous transitions which do not change just the state of a system but its organization itself. They lead to the creation of a new system with a new identity, obeying different laws and possessing different properties (Heylighen, 1991a). In such a system, the behaviour of the whole is constrained by the parts (a "reductionistic" view), but the behaviour of the parts is at the same time constrained by the whole (a "holistic" view) (Campbell, 1974a).

Perhaps the most important type of emergence is the "*meta-system transition*" (Turchin, 1977). Examples of metasystem transitions are the emergence of life, of multicellular organisms, of the capacity of organisms to learn, of human intelligence... A metasystem transition is characterized by an increase of the variety of possible actions (freedom) at the object level (usually through the assembly of a multiplicity of object systems), together with the emergence of a situation-dependent control at the metalevel, which coordinates, and chooses from, the variety of actions available at the level below (Heylighen, 1991a).

A constructive epistemology

Evolution can be likened to a problem-solving process searching through trial and error for an answer to the question: how to build a system that will survive in a maximum variety of situations? Knowledge is one of the results of that search: a mechanism that makes systems more efficient in surviving different circumstances, by short-cutting the purely blind variation and selection they have to do (Campbell, 1974b). The appearance of knowledge in the hierarchy of metasystems corresponds roughly with the emergence of life. Knowledge functions as a *vicarious selector* (Campbell, 1974b) which selects possible actions of the system in function of the system's goal (ultimately survival) and the situation of the environment. By eliminating dangerous or inadequate actions before they are executed the vicarious selector foregoes the selection by the environment, and thus increase the chances for survival of the system. Vicarious selectors are organized in a hierarchy of control levels (Campbell, 1974b), in accordance with our metaphysics based on metasystem transitions.

A vicarious selector can be seen as the most basic form of a *model*: an abstract system representing processes in the environment. A model is necessarily simpler than the environment it represents, and this enables it to run faster than, i.e. *anticipate*, the processes in the environment (Heylighen, 1990). It is this anticipation of interactions between the system and its environment, with their possibly negative

effects, that allows the system to compensate perturbations before they have had the opportunity to damage the system.

Models are not static reflections or homomorphic images of the environment, but dynamic constructions achieved through trial-and-error by the individual, the species or the society. This construction of models is similar to the continuous construction of systems by variation and selection that takes places everywhere in the universe. What models represent is not the *structure* of the environment but its *action,* insofar as it has an influence on the system. They are both *subjective,* in the sense of being constructed by the subject for its own purposes, and *objective,* in the sense of being naturally selected by the environment: models which do not recursively generate adequate predictions are likely to be later eliminated. There is no "absolutely true" model of reality: there are many different models which each may be adequate in solving particular problems, but no model is capable to solve all problems.

The most efficient way to choose or to construct a model which is adequate for the given problem is by reasoning on a metacognitive level, where a class of possible models can be analysed and compared. This requires a metasystem transition with respect to the variety of individual models.

An evolutionary ethics

The evolutionary philosophy can also be used for developing an ethics or system of values. The basic purpose here would be the continuation of the process of evolution, avoiding evolutionary "dead ends". Natural selection entails survival and development (growth, reproduction, adaptation...) as the essential value. However, the idea of an evolutionary ethics has not been very popular until now, and we will therefore go into a little more detail about this aspect of our philosophical system. Evolutionary ethics got a bad reputation because its association with the "naturalistic fallacy": the mistaken belief that human goals and values are determined by, or can be deduced from, natural evolution (Campbell, 1978). Values cannot be derived from facts about nature: ultimately we are free in choosing our own goals (Turchin, 1991).

However, we must take into account the principle of natural selection, which implies that if our goals are incompatible with the conditions necessary for survival, then we will be eliminated from the natural scene. Of course, there is no natural law or absolute moral principle which forbids you to commit suicide, but you must be aware that this means that the world will continue without you, and that it will quickly forget that you ever have been there. If we wish to evade this alternative, this means that we will have to do everything for maximising survival.

A second fallacy to avoid is the naive extrapolation of past evolution into the present or future. The mechanisms of survival and adaptation that were developed during evolution contain a lot of wisdom—about *past* situations (Campbell, 1978). They are not necessarily adequate for present circumstances. This must be emphasized especially in view of the creativity of evolution: the emergence of new levels of complexity, governed by novel laws.

For example, biological evolution, based on the survival of the genes, has favoured *egoism*: maximizing one's own profit, with a disregard for others (unless those others carry one's own genes: close family). In a human society, on the other hand, we need moral principles that promote cooperation, curbing too strong selfishness. Once the social interactions have sufficiently developed the appearance of such moral principles (e.g. "thou shalt not steal") becomes advantageous, and hence will be reinforced by natural selection, even though it runs counter to previous "egoistic" selection mechanisms (Campbell, 1978). The development of human society is an example of a metasystem transition, which creates a new system evolving through a mechanism which is no longer genetical but cultural (Turchin, 1977).

One of the implications of that transition concerns the interpretation of survival. Although the death of individual organisms may be useful for the renewal of the gene pool, making it easier for the genes to adapt to changing circumstances, it is no longer necessary for cultural evolution. In biological evolution survival means essentially survival of the genes, not so much survival of the individuals (Dawkins, 1976). With the exception of species extinction, we may say that genes are effectively immortal: it does not matter that an individual dies, as long as his genes persist in his off-spring. In socio-cultural evolution, the role of genes is played by cognitive systems ("memes", Dawkins, 1976), embodied in individual brains or social organizations, or stored in books, computers and other knowledge media. However, most of the knowledge acquired by an individual still disappears at biological death. Only a tiny part of that knowledge is stored outside the brain or transmitted to other individuals. Further evolution would be much more efficient if all knowledge acquired through experience could be maintained, in order to make place only for more adequate knowledge.

This requires an effective *immortality* of the cognitive systems defining individual and collective minds: what would survive is not the material substrate (body or brain), but its cybernetic organization. This may be called "cybernetic immortality" (Turchin, 1991). We could conceive its realization by means of very advanced man-machine systems, where the border between the organic (brain) and the artificially organic or electronic media (computer) becomes irrelevant.

The death of a biological component of the system would no longer imply the death of the whole system.

Cybernetic immortality can be conceived as an ultimate goal or value, capable to motivate long-term human action. It is in this respect similar to *metaphysical immortality* (Turchin, 1991): the survival of the "soul" in heaven promised by the traditional religions in order to motivate individuals to obey their ethical teachings (Campbell, 1979), and to *creative immortality* (Turchin, 1991): the driving force behind artists, authors or scientists, who hope to survive in the works they leave to posterity.

Another basic value that can be derived from the concept of survival is "self-actualization" (Maslow, 1970): the desire to actualize the human potential, that is to say to maximally develop the knowledge, intelligence and wisdom which may help us to secure survival for all future contingencies (Heylighen, 1990). Self-actualization may be defined as an optimal, conscious use of the variety of actions we are capable to execute.

However, if that variety becomes too great, as seems to be the case in our present, extremely complex society, a new control level is needed (Heylighen, 1991b). This may be realized by a new metasystem transition, similar to the one mentioned in the section on epistemology, leading to a yet higher level of evolution. A more detailed understanding of this next transition may help us to answer the question "Where are we going to?".

The main remaining problem of an evolutionary ethics is how to reconcile the goals of survival on the different levels: the level of the individual (personal freedom), the society (integration of individuals), and the planet (survival of the world ecology as a whole). It is an open question whether the "cybernetically immortal" cognitive system that would emerge after the next metasystem transition would be embodied most effectively in an individual being ("metabeing", Heylighen, 1991b), or in a society of individuals ("superbeing", Turchin, 1991). It is clear that the different levels have very complicated interactions in their effect on selection (Campbell, 1979), and hence we need a careful cybernetic analysis of their mutual relations.

References

Boulding Ken (1956): "General Systems Theory - The Skeleton of Science", General Systems Yearbook 1, p. 11-17.

Campbell D.T. (1974a): "'Downward causation' in Hierarchically Organized Biological Systems", in: Studies in the Philosophy of Biology, F.J. Ayala & T. Dobzhansky (ed.), (Macmillan Press), p. 179-186

Campbell D.T. (1974b): "Evolutionary Epistemology", in: The Philosophy of Karl Popper, Schilpp P.A. (ed.), (Open Court Publish., La Salle, Ill.), p. 413-463.

Campbell D.T. (1979): "Comments on the sociobiology of ethics and moralizing", Behavioral Science 24, p. 37-45.

Dawkins R. (1976): The Selfish Gene, (Oxford University Press, New York).

Heylighen F. (1990): "A Structural Language for the Foundations of Physics", *International Journal of General Systems* 18, p. 93-112.

Heylighen F. (1991a): "Modelling Emergence", *World Futures: the Journal of General Evolution*, (Special Issue on "Creative Evolution", G. Kampis, ed.)   (in press)

Heylighen F. (1991b): "Cognitive Levels of Evolution: from pre-rational to meta-rational", in: *Proceedings 8th Int. Conf. on Cybernetics and Systems (Vol. II)*, F. Geyer (ed.), (Intersystems, Salinas, California).
(in press)

Maslow A. (1970): Motivation and Personality (2nd ed.), (Harper & Row, New York).

Simon H.A. (1962): "The Architecture of Complexity", Proceedings of the American Philosophical Society 106, p. 467-482.

Teilhard De Chardin (1959): The Phenomenon of Man, (Harper & Row, New York).

Turchin V. (1987): "Constructive Interpretation of Full Set Theory", J. of Symbolic Logic 52:1 , p. 172-201.

Turchin V. (1991): "Cybernetics and Philosophy", in: Proc. 8th Int. Conf. of Cybernetics and Systems, F. Geyer (ed.), (Intersystems, Salinas, CA).

Turchin, V. (1977): The Phenomenon of Science, (Columbia University Press, New York).

Whitehead A.N. (1929): Process and Reality: an essay in cosmology, (Cambridge University Press, Cambridge).

Marc E. Carvallo
State University of Groningen,
The Netherlands

Self-organization, Evolution, and Religion:
Some Notes on Erich Jantsch's Theory of Religion

In his theory, Erich Jantsch uses both terms 'religion' and 'religio'. The former is usually valued pejoratively as ideological, institutionalized or at least as belonging to the structural-functional order that not only basicly but also 'surplus-ly represses life, dissipation and creativity of thee fluctuational order. This type of religion is characterized as the established traditional, western, monotheistic, and dualistically engined (comp. Jantsch 1980: 73, 177, 181, 241, 249, 257, 264). Only very seldom religion is valued positively, viz. that of cultures which are predicated upon paradigms essentially different from that of the above established one, as one might come across the buddhism, the mysticism, etc. (comp. Jantsch 1950: 303). This type of religion is an expression of what he calls 'religio' which generally means 'linking backward to the origin', 'restoring the broken symmetry' etc. (comp. Jantsch 1950: 216 ff., 264, 300-311). 'Religio' defined as such is the very evolution in Jantsch's vision. Consequently, the second type of religion is one of the vortices, splashes or ripples of the stream which is called 'religio' or evolution.

Evolution in Jantsch's vision is principally non-darwinistic and is characterized by :
1. non-dualism, coherence and self-consistency;
2. indeterminism and openness;
3. dissipative self-organization.

In this paper we will try to critically assess these principles and look for their possible theoretical and practical viability. Which condition should be fulfilled by the principle of self-consistency in order to also be 'ante-hoc' or 'a-priori' valid? What is the exact nature of the relationship between the future and the past according to Jantsch? Is it symmetric as has been propounded by e.g. Spinoza, Hegel or Kierkegaard or is it asymetric as was asserted by some modern panentheists e.g. Hartshorne (1973)? And from our present position of being the world of symmetry-breaks: to what or whom are we exactly or ultimately linking backward? Or does the endless 'religio' constitute the very ultimacy and infinity of ours? These are some profound questions surreptitiously hidden between the lines of Jantsch's evolutionary vision.

References

Hartshorne, C. (1973), *The Logic of Perfection*, La Salle, Illinois: Open Court Publ. Co.

Jantsch, E. (1980), *The Self-Organizing Universe*, Oxford/New York: Pergamon.

Alvaro Moreno, Arantza Etxeberria & Jon Umerez
Dpt. of Logic and Philosophy of Science
University of the Basque Country, Spain

Biological Information: The Causal Roots of Meaning

From an epistemological point of view, Life involves two kinds of processes that are, until now, irreducible one to another: the process of materially causing an effect and the one of representing or controlling another process. Semiotics draws a borderline between "semiotic" and "pre-semiotic" phenomena distinguishing between a "natural" meaning (that a sign must possess in respect to its referent by reason of a causal relationship between them) and a "non-natural" one (established by mediation of an interpreter and being the binary sign/referent relationship arbitrary without it). We would like to argue that even if this dichotomy is widespread in science nowadays, biological systems present phenomena of a mixed nature, where some relationships among components, even if being intrinsically causal, will not be established without the concurrence of a third instance that regulates them and could, therefore, be consider an interpreter of them.

Our hypothesis will be the following: *Natural meaning in Biological Systems has to do with cause/effect relations at a certain level of organization that revert in "emergent" configurations at a higher level that, on the one hand, fulfill some functional action and, on the other, are unpredictable from the lower level*. In this way, a relationship that is diadic (cause/effect) becomes triadic if we take into account the functional interpretation of it that occurs in the higher level; inversely also accomplishes a regulating action (boundary condition or constraint) over the processes taking place in the lower level.

We intend to discuss the following points to argue in favor of this hypothesis:

1) Causality and Determinism; even if often treated as synonyms there are important epistemological differences between them and most of the biological phenomena are causal not being deterministic.

2) Causality between different levels of biological organization will be characterized as forms of emergence. We will take into account three forms of observation of emergent phenomena:

a) Epistemological; in the sense of a deviation of the behavior of a system in respect to a model of it. Consequences are novelty production and unpredictability.

b) Ontological; from a bottom-up perspective there appears a great simplicity in the upper level, in contrast to the variety of the lower one, which has a regulating or controlling effect on it. Selection of equally viable alternatives can also be taken as a case of ontological emergence.

c) Methodological; phenomena studied in a) and b) revert in problems for classical tools of system description. The main difficulties are the modelling of the variability of relevant components in biological processes and the necessity of an-always-changing dynamics that stems from it.

3) From these points we can describe two types of information in biological systems (*information1* and *information2*).

a) *Information1* is characterized by self-referentiality.
More specifically, it is a form of organization characterized by the construction (starting from the lower dynamical level) of a network formed by components constituted in sequences of metastable structures which produce and inverse transformation discrete to continous by the effect of the dynamical components of the network. The result of this network is an overdetermination on the dynamical organization of lower level. Some degrees of freedom dicrease at this level and the action of the upper metastable level creates new functional components if necessary for the maintenance and reproduction of the network as a whole. The upper metastable structures (discrete) can be characterized as a "self-descriptive" information within the system (for example, genetic information).

b) *Information2* grasps the notion of "knowledge" and its referent is external to the system. The way of providing *information2* of a qualitative and semantic content is not to make it a constituent of models or descriptions independently constructed in priviledged spaces (minds, brains, computers or libraries). Instead it is important to collect its semantic content from the active/causal role it plays within the system itself, in an intrasystemic way. *Information2* requires a more complex type of network than *information1*: it involves a transformation of external physical patterns into sequences of metastable units. The action of these latter is functionally evaluable by a loop that ensures the reproductive identity of the system (the network described for *information1*) and its causal action consists in the establishment of a functional correlation towards the environment through some specific control action on the network.

References

Cariani, P. (1990) "Adaptivity and Emergence in Organisms and Devices" . To be published in World Futures: the Journal of General Evolution, (Special Issue on "Creative Evolution" G. Kampis, Ed.) .

Csanyi V. (1989) Evolution in Biological and Social Systems. A General Theory. Duke University Press.

Eco, U. (1976): A Theory of Semiotics; Bloomington, Indiana University Press.

Fernandez, J; Moreno, A. & Etxeberria, A. (1990): "Life as Emergence: the Roots of a New Paradigm in Theoretical Biology". To be Published in World Futures: the Journal of General Evolution, (Special Issue on "Creative Evolution", G. Kampis, Ed.).

Heylighen, F. (1989) "Causality as Distinction Conservation: a Theory of Predictability, Reversibility and Time Order". Cybernetics and Systems, 20, pp 361-384.

Kampis, G. (1990). Self-Modifying Systems in Biology and Cognitive Science. Pergamon Press.

Klee, R.L. (1984): "Micro-Determinism and Concepts of Emergence". Philosophy of Science, 51, pp. 44-63.

Moreno, A.; Fernandez, J. & Etxeberria, A . (1990): "Biological Computation and the Emergence of Cognition"; Presented in the "Symbols and Dynamics Workshop" in Storrs (Connecticut). Submitted to Systémique.

Polanyi, M. (1968). "Life's Irreducible Structure". Science.160, pp.1308-1312.

Pattee, H.H. (1982) "Cell Psychology. An Evolutionary Approach to the Symbol-Matter Problem". Cognition and Brain Theory.5(4), pp.325-341.

Rosen, R. (1985)."Organisms as Causal Systems Which Are Not Mechanisms. An Essay into the Nature of Complexity". in R. Rosen. Ed. Theoretical Biology and Complexity. Academic Press.165-203.

Sercarz E.E.; Celada, F.; Mitchinson, A.A. & Tada, T. (1988): The Semiotics of Cellular Communication in the Immune System. Springer.

KNOWLEDGE DEVELOPMENT

Markus F. Peschl
Dept. for Philosophy of Science,
University of Vienna

The Emergence of Symbols in Subsymbolic Neural Representation Systems

Knowledge representation is one of the central problems in the investigation of cognitive processes, cognitive science, AI and cognitive modeling. In the traditional approach of orthodox (i.e. symbol manipulating) AI symbols are assumed to be the ultimate or atomic representation structures. As will be discussed in this paper, it turns out that, if we are assuming a more epistemological perspective, the assumptions being made in this traditional approach are not adequate for achieving a deeper understanding of cognitive processes. Orthodox AI and cognitive science are mainly interested in technical and computer science issues; the naive understanding of (natural) language and its generality as a representation system is not reflected—this will be done, however, in this paper in order to show the basic problems of this approach.

In traditional Artificial Intelligence and cognitive science the central problem of knowledge representation is very much reduced to technical issues and symbol manipulation. This paper discusses some problems arising, if neither epistemological nor neuroscientific issues are considered in the field of cognitive science and of investigating knowledge representing and knowledge processing systems. An alternative approach is presented: *computational neuroepistemology*; it tries to consequently and interdisciplinarily integrate epistemological, neuroscientific, second order cybernetics as well as computer science (Parallel Distributed Processing) issues. Some methodological issues of this approach will be presented: it is based on the assumption that (scientific as well as common sense) knowledge develops in a cybernetical feedback process of speculation, construction, empirical investigation and verification; computer science plays the important role of integrating these two poles by applying its simulation techniques (i.e. neural computing).

Both natural language and formal symbols are assumed to be one of the most important representation structures in natural as well as in artificial cognitive systems—I am trying to differentiate between various levels of representation in an epistemological investigation considering both traditional (i.e. symbol manipulation) and neurally inspired simulation methods of AI and cognitive science. The pros and cons are discussed; it turns out that a symbolic representation system can be *integrated* and *embedded* in the more general neural representation system if we are considering constructivist and second order cybernetics concepts (of language, knowledge, etc.; Maturana, von Glasersfeld, von Foerster, Varela,...).

The neural representation system as well as language are understood as a system of references; generally spoken, a certain pattern refers to another pattern or state by an artificially generated and constructed relation. It turns out that both have a constructivist character which means that knowledge and language are the result of a process of construction both being realized in neural processes. Language is understood as one special and very complex form of behavior which is generated, as all other behavior, by the nervous system. What we are calling symbols (in our language, in computers, etc.) are emergent properties of the more general neural representation and reference system.

Symbols and language have to be understood as a highly specialized system of references following rules which we describe as the grammar of a language—it is important to see, however, that the grammar of a language is only one possible way of describing laguage on a very superficial level—computational neuroepistemology suggests a bottom-up approach being determined by the neural dynamics rather than by artificial "systems of explanation". The implications of such a view on the development of a model of cognition will be discussed in detail. A model of cognition being based on these assumptions is presented.

Charles Henry

Butler Library
Columbia University, New York

Non-Verbal Aspects of Language and Knowledge Structuring

Any cybernetic system that relies on the semantic relationship of words as part of its structuring needs to confront an inherently paradoxical aspect of language. This paradox, which concerns the relationship between verbal and non-verbal components of language, involves the structuring of knowledge as well, and represents a kind of 'covert' act of intellection that has recently become the focus of cognitive studies. The premise underlying this paradox can be summarized as:

1) language engenders *images* in the mind, whether the language is written or spoken

2) words or phrases that are unrelated etymologically and have no syntactic, phonetic, or semantic correlation can nonetheles produce identical images

3) the images thus produced as mental representations can in fact contradict or oppose the apparent (linguistic) meaning of the text

4) in some cases the meaning of a word or phrase can only be understood by the recognition and analysis of these images

5) the end result of this process can be the acquisition of new knowledge

6) elucidation of the non-verbal aspects of language in comparison to the linguistic models of language sheds new light on the mind/brain question

The purpose of this paper is twofold: to elucidate the non-verbal aspect of language from concrete examples separated by milennia in order to underscore that this aspect of language is universal, not limited to a particular language, historical period, linguistic structure, or theme. Secondly, to show the implication this phenomenon holds for cybernetic design, touching upon contemporary research in cognitive science, artificial intelligence, and the physiological properties of the brain. Linguistic examples are drawn from the ancient Egyptian Book of the Dead, the Hebrew Genesis, and a contemporary poem by P. Neruda to underscore the universality of the premise.

Elan Moritz

The Institute for Memetic Research
Florida

Memetics: Introduction and Implication to the Evolution of Knowledge

Memes are information clusters whose patterns and meanings provide selective advantage for their replication and spread. In the context of human society, memes can be regarded as *units of cultural transfer*. Examples of simple memes are hair and clothing fashions, slogans, certain religious beliefs, popular music tunes, certain graphic designs [e.g., the "peace" symbol, "male" and "female" symbols, the multiple orbits symbol for "atomic" related equipment or hazard]. The attributes that characterize memes are their preferential copying [with a high degree of fidelity to the original version], by many individuals, as compared to other informational entities. More complicated meme-like constructs are also possible. These may be collections of simple memes. Examples of 'meta-memes' are scientific theories, religions, movies, musical symphonies, etc. In the first part of this paper we present the basis for and recent progress of a quantitative science of memes [memetics] that combines a decriptive calculus for memes, principles of population dynamics, information theoretic measures with physics based least action principles. In the second part of the

paper we discuss the implications of memetics for the evolution of knowledge.

With respect to the objectives of the Principia Cybernetica Project [PCP], and the interest of developing computer based linking of knowledge, several mappings of mental memes, or ideas, to physical representations, are discussed. In complete analogy to the biological gene – genetic engineering metaphor, it is possible to utilize the PCP framework to construct new knowledge using meme mutation, combination, and spread. If we categorize PCP participants as humans [H] and machines [M; e.g. computers, books, videotapes, or any non-biological information capture/manipulation devices], then new knowledge can emerge by one or more of the following interactions: H–H, H–M–H, H–M, M–M. Estimates of quantity of new knowledge [though not necessarily correct knowledge] generation and spread can be obtained given empiricaly available memetic relationships. Results of simulations using a Zipf Law [inverse frequency] meme spread activation are presented. Suggested resource [e.g., time, energy, memory, space] cost metrics for PCP interactions are described. Results of meme spread and new meme generation simulations are interpreted in terms of the suggested resource cost metrics.

References

Moritz E. (1990): "Memetic Science: I - General Introduction", Journal of Ideas 1, p. 1.

Moritz E. (1990): "Replicator Based Knowledge Representation and Spread Dynamics", in: Proc. IEEE International Conference on Systems, Man & Cybernetics [Nov 4-7, 1990 Los Angeles, California], p. 256-259

J.L. Elohim

Seccion de Graduados e Investigacion
ESIME, Instituto Politécnico Nacional
Mexico

Culture, Cybernetically Interpreted, is a Cybernetic Reflection of Nature Altered by Culture

Apparently, hominids first and human beings later, since those very early times, after their emergence as "quite cleverer animals", while they were engaged in "searching out" how to survive, were increasingly obliged to reflect into their minds diverse aspects of that "reality" where they were located, most often without realizing that they were

also constituent parts of such reality. Gradually these beings had to learn how to get better reflections of everything in their surroundings.

Nowadays, many of us are certain that *decision-making*—aiming to organize consciously what we assume our respective performances should be in the future, at least for surviving (physically, emotionally or intellectually)—must be supported indispensably by our thoughts constituted by relatively suitable images, i.e. *reflections* of particular aspects of the reality; aspects that we are capable to perceive and whose images in our minds are judged useful for such purpose.

When analyzing the kind of thoughts expressed during mankind's history, it seems proper to assert that a lot of them constitute the main source of information that has made possible to build a certain artificial world, not another. This "world" inserted into the natural one makes up the "human" civilization.

But this artificial world has never been fully conceived in advance; it has never been designed as a whole nor the whole set of available reflections of the reality has been implemented at the same time. On the contrary, civilization has always been a set of facts and events that have come out of chaotic combinations of quite dissimilar processes:

- progression, retrogression and deterioration
- evolution and involution
- innovation and tradition
- peace and war
- cultural changes and social bonds

Yet it cannot be denied that civilization is the outcome of an increasingly *improved comprehension* of the evolution of real phenomena which comprises both the natural ones and the others invented by men, which are in fact the essence of the artificial world already mentioned.

Recently, during the last two decades, a generalized cybernetics has emerged and developed as an alternative guide that, no doubt, has greatly improved such comprehension, and that has gradually allowed us to consciously conduct the dynamics of phenomena belonging to diverse realities: inanimated, living or artificial.

In accordance with this cybernetics, I would claim that every object of the natural world is in fact a subject which can be seen as a relatively well-structured system that "exists" by itself, and has a place in space, while its "performance" is a function of a certain degree of autonomy in an environment which is under the influence of many other subjects. This is a relative autonomy that arises as the outcome of diverse cybernetic relations among the subject's elements. These relations "help" the subject to organize its "performance" by itself and create suitable conditions allowing the subject to learn how to take into account effects of the expected performance on many other subjects, while finding out how to perform "freely".

Stones, amoebas, plants, animals, ... which emerge as effects of particular involutionary and evolutionary natural processes, are clear evidences of the infinite number of possibilities that arise from the manifestation of these cybernetic relations. The natural emergence of adaptive, prospective, intuitive, ... processes, awareness, consciousness etc., are also evidences of such kind of possibilities.

Human beings, which emerge as well from natural processes, have apparently reached the highest level of autonomy by means of their thinking, which offers them the possibility of getting a proper cybernetic understanding of everything that moves in time and in space. Such understanding is knowledge that quite circumstantially becomes the source of cultural actions. These actions become as well sources of specific technological, economic, political, educational, ... actions which in accordance to the way they have been developed can be considered as systems that manifest themselves first as intellectual possibilities and become later societal phenomena.

I would claim through this paper that any kind of intellectual systems, and human culture in general, are necessarily particular reflections, relatively faithful (though sometimes distorted on purpose), of cybernetic possibilities intrinsic to the dynamics of Nature, which is increasingly altered by an artificial world that so far is built rather unconsciously by men.

COMPUTER-SUPPORT SYSTEMS

Cliff Joslyn

Systems Science
State University of New York at Binghamton

Software Support for Principia Cybernetica Development

Principia Cybernetica [1] is a project to develop a collaborative, consensually based, constructive, philosophical system. Essential to such a project are computerized tools to aid in system construction. Such tools and technologies as hypertext, hypermedia, electronic mail, and textual markup, would allow the construction and publication of structured, non-linear, multi-dimensional semantic systems and documents by a collaborative group of spatially separated contributors in a hybrid natural and formal language environment.

We will consider the purposes (ends) and the architecture of a possible computer system (means) through which these goals could be approached.

Desiderata

- The constructive development of a system of philosophy.
- The collaborative development of a system of philosophy.
- To support the development of consensually-held views among a number of researchers.
- The dynamic development of a system of philosophy.
- To support the process of semantic analysis.
- To fully represent the semantic relations among the components of the system.
- Easy movement between natural language, formal language, and mathematical notation.
- To publish the whole or portions of the system through traditional means.

Structures

- Strict and loose hierarchical structures.
- Semantic categories for content development (e.g. part-whole relations).
- Syntactic categories for formal development (e.g. bibliographical references, time-flow of argumentation).
- Various publishing models (e.g. dictionaries, encyclopedias, reference books, compilations, arguments, textbooks).

Methods and Technologies

- Hypertext and Hypermedia systems.
- The necessity of ASCII source files and hypertext markup languages.
- The role of the SGML standard for general textual markup and electronic publishing.
- The role of the `HyTime` standard for hypertext markup.
- Commercial SGML and `HyTime` products to support PRINCIPIA CYBERNETICA.
- Electronic mail, electronic journals, and the `PRNCYB-L` mailing list.

References

[1] Heylighen, Francis; Joslyn, Cliff; and Turchin, Valentin: (1991) "A Short Introduction to the PRINCIPIA CYBERNETICA Project", *Journal of Ideas*, v.2:1, pp. 26-29

Dirk Kenis
Policy Consulting Services
Free University of Brussels (VUB)

MacPolicy: Delphi and Group Decision Support Ideas for Computer Supported Cooperative Working

The first part of this paper will provide a state of the art in *Group Decision Support Systems* (GDSS) research and related fields. The second part will give an overview of the results of preliminary research. The third part will describe the aims of the ongoing research at the V.U.B. on Computer Supported Cooperative Working.

Group Decision Support Systems : State of the art

With the boom of networking and the necessity for sophisticated well-structured communication and discussion software, offering more than just passive data transfer like in traditional mailing systems, we expect the GDSS ideas to become implemented more and more in enhanced network communication software, that enables collaboration and interactive simulation.

An important shift in recent GDSS theory is making it clear that the traditional narrow approach of developing network applications for GDSS-rooms to ameliorate decision-making sessions directed by 'animators' is replaced by a more general interest in what is described best as *Computer Supported Cooperative Working*. Hence, we witness also the recent emergence of GDSS related fields:

- Computer Conferencing (especially Computer scientists);
- Shared Information Technology (especially Management Information scientists);
- Computer-aided (or -facilitated) Group Decision Making ;
- Computer Supported Cooperative Working (especially Knowledge Scientists, Cybernetics);
- Computer Based Group Communications and Computer Mediated Communication (especially Communication scientists).

Computer Supported Cooperative Working projects at the V.U.B.

Long term (2-3 years) interdisciplinary research programmes, of which one in collaboration with a software house, with in total 3 full-time social scientists and 4 full-time computer scientists will start mid '91 at the Free University of Brussels, in order to find out ways to enhance group-work through structured (network) communication.

In a first phase we will further develop and experiment with a GDSS based on the principles of a—in human sciences—succesful research method, the (Policy) Delphi method. Our GDSS, with HyperCard as interface on top of a powerful database, will be enhanced with several operational research techniques (*Multiple Criteria Decision Aids*) and, after substantial testing in real world settings, will be rewritten through an object-oriented approach as an independent software application. Through Technology Assessment and Analogy Methods applied to similar communication technological innovations, combined with group-dynamic experiences in our experimental GDSS-setting, we will try to find out the conditions to improve 'Computer Supported Cooperative Work'.

On the basis of these findings we want to develop a flexible *meta-tool*, a hypermedia environment where 'groupwork' applications can be easily created or adjusted. This 'flexible meta-tool' will enable us to develop network-applications in the field of interactive simulation (i.e. preparing answers for opposing questions on an important meeting through interactive gaming), or with expert systems enriched instruments for group planning (i.e. aids to construct scenarios in a network session), etc.

Francis Heylighen
PESP, Free University of Brussels

Structuring Knowledge in a Network of Concepts

The basic evolutionary-systemic and constructive principles that have been discussed in my two previous contributions to this volume can be directly applied to the design of a computer support system that would help Principia Cybernetica collaborators to develop a coherent system of philosophical thought. In fact the same type of support system might be applied to any complex problem domains where on the basis of a lot of ill-structured, ambiguous and sometimes inconsistent data a more or less simple and reliable model is to be built. The problem we are speaking about is one of *applied epistemology*. A good epistemology, offering a concrete and general theory of how knowledge develops during individual or cultural evolution, should

also be useful as a guide when a new model is practically to be developed.

Network representations of knowledge

I start from the assumption that a lot of knowledge is already available, in literature and in the heads of different (potential) contributors to the project, but that that knowledge must be integrated into a coherent and transparent model. The knowledge will be assumed to be written down in the form of "chunks", containing text, formulas, drawings, sound, ..., whatever media are most appropriate to express the underlying ideas. I further suppose these chunks to be split up into distinct "ideas" or "concepts", such that one chunk should define not more than one concept.

Of course, these different concepts will be related and one chunk will in general contain references to several other chunks. For example, the chunk denoting the concept "dog" might contain the following sentence: *a dog is a carnivorous mammal, with a protruding snout*. This means that the concept dog has associations with a least the concepts mammal, carnivorous and snout. If these concepts are also available as chunks, then we might create a link from the dog chunk to the mammal chunk and so on. Computer applications that allow such an easy representation and manipulation of chunks connected by links are called *hypermedia* systems. The chunk with its text and graphics can be shown in a window on the screen, and it suffices to click on one of the links to show the next chunk to which the link is pointing (Heylighen, 1991).

Hypermedia system are useful for storing a large amount of complex, interrelated information (e.g. an encyclopedia) in a easy to handle way. However, there is an inherent ambiguity involved, since it is not a priori clear what a link is supposed to mean: any kind of association, as well causal, as logical, as intuitive as spatial, ..., might be represented by a link. Therefore we need a better structured system if we want our networks of concepts to support us more efficiently. By introducing different *types* of chunks (nodes) and links we may turn our hypermedia system into a *semantic network*: the different types of links will determine (part of) the meaning of the concept to which they are attached. The problem with semantic networks for knowledge representation is still that of ambiguity: there is an unlimited number of link and node types that may seem appropriate, and their interrelationships will in general be very unclear. In order to limit the set of types, we need an unambiguous, fundamental interpretation of what concepts and links in our network really stand for. I will now propose such an interpretation with the corresponding types, and show how it can be applied to the structuring of knowledge.

Distinction and entailment types

A concept (node) is supposed to represent a *distinction*: a way to separate phenomena denoted by the concept (belonging to its class or extension), from phenomena that do not belong to its extension. Defining a concept means proposing a procedure for explicitly carrying out that distinction. Definition will be assumed to be a *bootstrapping* operation: a concept is always defined in terms of other concepts, that are themselves defined in terms of other concepts, and so on. In general there is no primitive level of meaningful concepts in terms of which all other concepts can be defined. This is in accordance with my constructive philosophy, stating that any foundations of a conceptual system must be empty of meaning in order to be acceptable as basis for a complete philosophical explanation (Heylighen, 1990b).

One way to define a concept is by listing the set of concepts that it entails together with the set of concepts entailed by it. By *entailment* I mean an "if...then" relation, which is more general than the logical (material) implication. For example, *if* a phenomenon is a dog, *then* it is also a mammal: dog → mammal. It means that a phenomenon denoted by the first concept cannot be present or actual, without a phenomenon denoted by the second one being (simultaneously) or becoming (afterwards) actual.
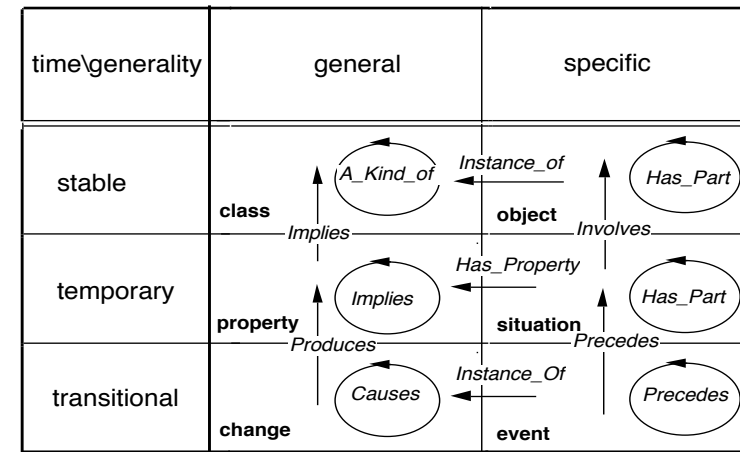
In order to derive fundamental types of distinctions (concepts, nodes) and links (entailments), we will posit two basic dimensions of distinction: stability (or time) and generality, with the corresponding values of instantaneous - temporary - stable, and of specific - general. The combination of these 3 x 2 values leads to 6 types of distinction (see table).

TIME\GENERALITY	general	specific
stable	class	object
temporary	property	situation
instantaneous	change	event

For example, an object is a distinction that is stable (it is not supposed to appear or disappear while we are considering it), and specific (it is concrete, there is only of it). A property is a distinction that is general (several phenomena may be denoted by it, it represents a common feature), and temporary (it may appear or disappear, but normally it remains present during a finite time interval). An event is instantaneous (it appears and disappears within one moment), and specific (it

does not denote a class of similar phenomena, but a particular instance).

With these node types we can now derive the corresponding link types by considering all possible combinations of two node types. There is one constraint, however: we assume that a more invariant (stable or general) distinction can never entail a less invariant one. Otherwise, the second would be present each type the first one is, contradicting the hypothesis that it is less invariant than the first one. For example, a class cannot entail an object, a situation cannot entail an event. Yet it is possible that concepts with the same type of invariance, (e.g. two objects) might be connected by an entailment relation. All remaining possible combinations can now be summarized by the following scheme (the straight arrows represent entailment from one type to another (more invariant) one, the circular arrows entailment from a concept of a type to a concept of the same type):

time\generality	general	specific
stable	A_Kind_of *(class)* / Implies	Instance_of / Has_Part *(object)* / Involves
temporary	Implies *(property)* / Has_Property / Produces	Has_Part *(situation)* / Precedes
transitional	Causes *(change)* / Instance_Of	Precedes *(event)*

For example when an object A entails a class B, A → B, then A is an Instance_of B. When an object A always entails the presence of another object B, then B must belong to or be a part of A. When a change A entails another change B, then A and B "covary" and hence A can be interpreted as the cause of B. When an event A entails a situation B, then A must be simultaneous with or preceding B in time.

The advantage of this scheme is that most of the intuitive and often used semantic categories (objects, classes, causality, whole-part relations, temporal precedence, etc.) can be directly constructed from it, in a simple and uniform format. Complementarily, given some of those everyday categories, we can use the scheme to reduce them to

simple entailment links between nodes of specific types. In fact the types themselves can be represented as nodes, and each node of a particular type will have an entailment link to that 'type'-node. This allows us to reduce a complicated set of semantic categories to an extremely simple formal strcuture.

Knowledge structuring

Given that structure, consisting of a list of nodes and entailment links between them, we can now start to formally analyse the network. Define the input and output sets of a node:

Input: $I(x) = \{ y \mid y \to x \}$ = "extension" of concept x
Output : $O(x) = \{ y \mid x \to y \}$ = "intension" of concept x

The meaning (definition, distinction) of x can be interpreted as determined by the disjunction of its input elements, and the conjunction of its output elements. Our previous remark about definitions can now be reformulated as the following *bootstrapping axiom* (Heylighen, 1990ab):

two nodes are distinct if and only if their input and output sets are distinct:
$x \neq y \Leftrightarrow I(x) \neq I(y), O(x) \neq O(y)$

However, such a complete definition assumes that all concepts allowing to distinguish between x and y are present in the network. In practice, the network of concepts we are building by writing down our knowledge in the form of connected chunks, will be *incomplete* in some respects, *redundant* in other respects. Instead of using the axiom as a static description of how a complete network should be structured, we can use it as a procedure to find ways to make the network more adequate, by adding missing concepts, or by deleting redundant ones. We can distinguish the following two main techniques (cf. Heylighen, 1991; Bakker, 1987; Stokman & de Vries, 1988):

Node identification

When input and output sets of two nodes x and y are identic or similar, the computer support system may propose the user to either identify (merge) the two nodes, and replace them by one single node, or to add new nodes or links that would more clearly differentiate between x and y. An algorithm may test the identity or inclusion of the input and output sets, and according to the results, propose the following possibilities to the user:

1) $I(x) = I(y)$:
    a) $O(x) = O(y) \Rightarrow$ Identify (or distinguish) x and y

    b) $O(x) \subset O(y) \Rightarrow$ Identify x and y, or distinguish I (x) from I (y)

2) $I(x) \subset I(y)$:
    a) $O(x) = O(y) \Rightarrow$ Identify x and y, or distinguish O(x) from O(y)
    b) $O(x) \subset O(y) \Rightarrow$ Identify x and y
    c) $O(y) \subset O(x) \Rightarrow$ Connect x to y, $x \to y$

Node integration

When a cluster of nodes have a common set of "external" input or output nodes (that is to say nodes that do not belong to the cluster), then from the point of view of those external nodes, the nodes inside the cluster are indistinguishable. Hence the nodes, though not strictly indistinguishable according to the bootstrapping axiom, behave indistinguishably from a certain viewpoint.

From that point of view, the cluster may be called *closed* (Heylighen, 1990a) and it might therefore be replaced by a single "integrated" node. The integrated node "summarizes" the cluster nodes on a more abstract level, and may hence simplify the conceptual model. Similar to the case of node identification, the external indistinguishability of clustered nodes may be spurious, and this should prompt the user to add additional distinguishing links and nodes.

There are different types of closure, with different meanings and formal properties, depending upon which sets of external input or output nodes are common among the cluster, for example: transitive closure, equivalence, cyclical closure, ... If the closure is only approximative (the cluster nodes have several external neighbours in common, but these do not form a complete set of any specific type), then this method is similar to the one called "conceptual clustering" in machine learning, where the boundaries between clustered and non-clustered nodes become fuzzy, and depend on the treshold chosen for the number of common neighbours.

In conclusion, the present set of concepts and techniques, when implemented on a computer through a suitable intuitive interface, should enable an individual or group of users to elicit and structure their knowledge about a domain under the form of a network of concepts connected by entailment links, and support them to minimize the redundancy, complexity and incompleteness of their model.

The introduction of new nodes and links by the user corresponds to a form of variation by recombination of concepts. The recognition of a closed cluster of nodes by the system corresponds to the selection of a distinction that is more stable or invariant than the distinctions

between the internal concepts of the cluster (Heylighen, 1990a), with closure as fundamental selection criterion. The elicitation and structuring of concepts in this manner hence follows the general evolutionary mechanism that was postulated in my previous papers about evolutionary philosophy.

References

Bakker R.R. (1987): *Knowledge Graphs: representation and structuring of scientific knowledge,* (Ph.D. Thesis, Dep. of Applied Mathematics, University of Twente, Netherlands).

Heylighen F. (1990): "A Structural Language for the Foundations of Physics", *International Journal of General Systems* 18, p. 93-112 .

Heylighen F. (1990): "Relational Closure: a mathematical concept for distinction-making and complexity analysis", in: *Cybernetics and Systems '90*, R. Trappl (ed.), (World Science, Singapore), p. 335-342.

Heylighen F. (1991): "Design of a Hypermedia Interface Translating between Associative and Formal Representations", *International Journal of Man-Machine Studies.* (in press)

Stokman F.N. & de Vries P.H. (1988): "Structuring Knowledge in a Graph", in: *Human-Computer Interaction, Psychonomic Aspects*, G.C. van der Veer & G.J. Mulder (eds.), (Springer, Berlin).

Robert Glück
Institut für Prakt. Informatik
University of Technology Vienna

Metasystem Transition in the Machine and its Application to Knowledge Systems

We propose to incorporate the notion of metasystem transition (MST) in knowledge systems and to provide tools capable of performing MST with regard to a knowledge system, an idea first stressed in [5]. In particular, we propose to investigate MST for systems supporting the Principia Cybernetica Project, a project dealing with cybernetic philosophy in which the very concept of MST plays a fundamental role. This is supported by very promising applications of the concept of MST, in the form of the Futamura Projections (FMP), to compiler-construction and -generation, central fields of computer science [1]. Why not apply and take advantage of the benefits of the MST for knowledge systems implemented on a machine?

What follows is a review of the principle of MST and the formulation of two potential application to knowledge systems.

• MST is a transition from one system S to a metasystem S^*. It expands the hierarchy of systems by adding a new level of control. In principle, adding a new level of control can be done repeatedly. In case the same system S is added as metasystem S^*, it is called self-application of S.
• A knowledge system (in general) comprises two parts: a knowledge base which is structured according to some knowledge representation scheme (e.g. semantic network) and a mechanism maintaining and organizing the knowledge base. More importantly this mechanism (inference engine) is capable of answering questions by examining (reasoning about) the knowledge base. Such engines may e.g. incorporate fuzzy or non-monotonic logic (of course not restricted to those).

The application of MST to knowledge systems can be described formally: Let *infer* be an inference engine, k the knowledge base and q a question. This will be formalized as follows (the notation is the same as in [2]).

<infer (k, q)> "Run the inference engine *infer* to answer the question q by examining the knowledge base k."

Definition: A program *alpha* is a program specializer (e.g. partial evaluator, supercompiler) iff for all programs p and arbitrary values x, y and the metavariable Y the following characteristic equations holds:

(1) *<p (x,y)> = <<alpha ↓ <p (x,Y)>> (y)> = <p-x (y)>*

Formula (1) represents the first MST: the knowledge base and the inference engine become objects under the control of *alpha*. Note that *p-x* is a program that is fixed to the value x. In addition note that the expression *<p (x,y)>* is metacoded (↓ = arrow down) [4]. The operation inverse to ↓ will be denoted by ↑ (arrow up). It is obvious that p can be substituted by *infer* in formula (1).

(2) *<infer (k,q)> = <<alpha ↓ <infer (k,Q)>> (q)> = <infer-k (q)>*
(3) *infer-k ::= <alpha ↓ <infer (k,Q)>>*

The program *infer-k* represents a program which is capable of answering questions about k without the need for examining and interpreting the knowledge base. All actions needed for interpreting the knowledge base have been removed so that q can be answered more efficiently.

The knowledge base may be large and change more frequently than the inference engine. In this case it may take some time for *alpha* to analyze *infer* and k in (3). Doing one more MST we get the following formula by applying *alpha* to the right side in (3), where *gen* is a program that is constructed by the second MST according to the semantics implemented by the inference engine (4)

$<alpha \downarrow <infer\ (k,Q)> = <<alpha \downarrow <alpha \downarrow <infer\ (\uparrow\ K,Q)>>>\ (k)> = <gen\ (k)>$

(5) $gen ::= <alpha \downarrow <alpha \downarrow <infer\ (\uparrow\ K,Q)>>>$

Consequently from (2) and (4)
(6) $infer\text{-}k ::= <gen\ (k)>$

Potential benefits:

- The knowledge base can be structured in a flexible and modular way, without tailoring it for a specific inference engine. To utilize the knowledge base for new purposes or to analyse it from a new angle, the new inference engine *infer'* is supplied. Reorganizing the knowledge base, adapting and filtering relevant information will be done automatically in the process of constructing *infer'-k* by the MST in (3) or (6).
- The knowledge representation scheme can be chosen independently of the internal data structure and without having the needs of a specific inference engine in mind.
- The creator of the knowledge base is free to choose a scheme appropriate for the representation of his subject without any penalties regarding efficiency and implementation.

The author suggests to investigate the application of Supercompilation to knowledge systems, because deep structural transformations can be performed by driving and generalization [5] and because it is intrinsically more powerful than partial evaluation (which already has been successful used for compiler generation by metasystem transition). Conclusion: MST as very general and fundamental concept may not only be beneficiary in classical fields of computer science but also in the knowledge systems. What kind of *k* and *infer* are practical, their properties with regard to MST and how *alpha* may be constructed for knowledge systems in particular if it is suited for the Principia Cybernetica Project may be a subject for further discussion at the workshop.

References

[1] Bjorner D., Ershov A. P., Jones N. D. (ed.), *Partial Evaluation and Mixed Computation*. North-Holland: Amsterdam 1988.
[2] Glück R., Towards multiple self-application. In: *Proceedings of the Symposium on Partial Evaluation and Semantics Based Program Manipulation*. (New Haven, Connecticut). ACM Press (to appear) 1991.

[3] Glück R., Turchin V.F., Application of metasystem transition to function inversion and transformation. In: *Proceedings of the ISSAC '90*. (Tokyo, Japan). 286-287, ACM Press 1990.
[4] Turchin V. F., The language Refal - the theory of compilation and metasystem analysis. Courant Institute of Mathematical Sciences. *Courant Computer Science Report* No. 20, 1980.
[5] Turchin V. F., The concept of a supercompiler. In: *ACM TOPLAS*, 8(3): 292- 325, 1986.

Peter Beyls
Computer Artist
St. Lucas Institute, Brussels

Conceptual Exploration through Intimate Machine Interaction: a statement *

The objective of my work is *conceptual navigation*. Pragmatic considerations lead to the design of computerized vehicles allowing elegance and optimal flexibility while playing with ideas. The general approach is cognitive rather than procedural or mechanistic. We conceive and develop machine partners which assist the artist in the process of exploration and discovery. Digital media may encourage *intimate machine interaction*, i.e. the interactive evaluation of the behavioural potential of a given idea. In addition, the artist learns about the true nature of his intentions through visual feedback.

Consider the development of virtual workspaces of which the artist is both inventor and explorer. The central material component is knowledge, rather than information. This implies that we are interested in the *meaning* of things rather than their visual appearance. The automatic generation of intricate pictorial complexities as such is of no concern. However, the study of levels of autonomy in the creative process is important since we aim to design computational environments that accomodate mental models of creative behaviour. Computers allow for manipulation of ideas on the symbolic level. Arbitrary concepts like conflict resolution, adaptation or responsibility are formalized and activated in a simulated, *virtual world*. The activity in this world manifests itself in pictures. These pictures are visual representations that emerge from the inherent abstract activity and careful selection of physical attributes imposed by the artist. The pictures document themselves.

In summary, the sharing of responsibilities between man and machine—while aiming to create in a common effort—is the heart of the

matter. The initial spark for many incarnations of activity and interactivity is borrowed from examples in nature or it may be a product of human imagination. In either case, our objective remains the interpretation rather than the understanding of the internal dynamics of the cognitive process. The idea is to create a context for the exploration of the psychology of humans as well as the psychology of machines. The final works are side effects of the very activity of navigating in unknown conceptual territories.

* this statement about his artwork was prepared by Peter Beyls on April 20, 1991, for an exhibition in Antwerpen.

Elan Moritz
Institute for Memetic Research
Florida

The Case for Imperfect Machines

In almost all attempts to develop computer based expert system, artificial intelligence, or even common-sense reasoning systems, certain demanding and rigourous performance goals are set. Typically performance goals aim at perfection and repeatability, where perfection is interpreted to mean performance at the level of an 'above average' or 'expert' human practitioner. It is argued in this paper that such performance expectations for machines [a general term for any man-made electronic, mechanical, biological, or other devices] may be too demanding. If one considers the length of time for 'nature' to evolve human beings operating at human skill levels, it becomes apparent that it is not realistic to demand similar performance of machines designed in an infinitesimal fraction of the time to 'design' man. While machines have been built that can perform some calculational functions extremely fast and reliably, no machine has yet been designed that remotely approximates the multiple attributes and abilities of even the 'simplest' man.

Many of the characteristics of Man are due to the existence of 'Man' as part of a socio-cultural network and the training of new members of 'Man' by that socio-cultural network. Due to the significant numbers of individuals in this socio-cultural network, the training of new members of 'Man' takes on some random aspects which lead to opportunities of significant departures from the mean. While most individuals are familiar with unusually successful members of 'Man', the instances one is likely to encounter are imperfect instances of 'Man'. If we restrict our attention to cognitive imperfections, the types of imperfections we are likely to encounter

may range from simple things such as immaturity, incomplete knowledge, slow-learning, mild retardation, to more severe pathologic conditions such as schizophrenic delusions, manic-depressive conditions, madness [incorrect combination of basic assumptions].

While no one sets out as an end goal to build imperfect machines, it is argued in the paper that one needs to anticipate imperfect machines, especially when operating in knowledge based domains. It may be the case that to achieve 'useful' knowledge-based machines one needs to start with a collection of self-organizing imperfect machines that are allowed to 'mature'. The nature of the 'maturation' process may be as simple as a Hebbian process, or perhaps more complex. We explore the negative as well as positive ramifications of design of imperfect machines as compared to attempts of designing perfect machines.