APPRAISAL

A Journal of Constructive and Post-Critical Philosophy and Interdisciplinary Studies

Vol. 1 No.1 March 1996 ISSN 1358-3336

Editor and Publisher: Dr R.T. Allen
20 Ulverscroft Road, Loughborough, LE11 3PU, England
Tel., fax and answer-phone: 01509 215438; E-mail 101625.3010@compuserve.com

CONTENTS

This issue's contributors:

Dr Csaba Pléh, having studied linguistics and psychology at Eötvös Lorand University (ELTE) in Budapest, is now an associate professor there and chairman of the Dept of General Psychology, and teaches the history of psychology, psycho-linguistics and theoretical issues in cognitive science. He has published (in Hungarian) a history of psychology, and articles on Karl Bühler, W. Wundt and L.S. Vygotsky

Professor Percy Hammond has worked in the electrical manufacturing industry and taught electrical engineering in several universities including Cambridge, Southampton, Stanford (USA) and the University of Malaya. He has written five books and a number of technical papers on various aspects of electrical engineering. He is a fellow of the Royal Academy of Engineering and of the Institutions of Electrical Engineers and Mechanical Engineers.

Continued on p.2

Dr Endre Nagy studied sociology, graduated in law and qualified as a lecturer in history. He is now Head of the Department of Social Policy at Pécs, President of the Theoretical and Historical Section of the Hungarian Sociological Association, and a member of the Editorial Board of *Polanyiana*. Among his publications are (in Hungarian) 'Introduction to the Social Thought of Michael Polanyi' (in *Logosz*, ed. I.G. Kovács, Budapest, 1990) and *Idea and Reality: Studies in the History of Sociology in Hungary* (Budapest, 1993).

Dr Gheorghe Constandache, having graduated in engineering and taken a diploma in civil engineering, then studied philosophy, which he has taught at the Polytechnic University of Bucharest since 1973. Among his books are: *Philosophy of Cognitive Sciences* (1993), *Metaphysics of Computers* (1994) and *Rationality, Language, Decision* (1995)

The Rev'd Julian Ward is Dean of Students at Elim Bible College, Nantwich, where he has been lecturing in theology and philosophy of religion since 1974, having previously worked as a research aerodynamist in the aircraft industry, before undertaking theological training.

EDITORIAL

As stated on the inside front cover, *Appraisal* will seek, in the spirit of Michael Polanyi, to encourage constructive approaches within philosophy, to issues within it, and to the philosophical dimensions of questions that arise in other disciplines and in everyday life. It is based on the traditional belief that philosophy does have something positive of its own to offer: viz. the articulation and elaboration of that global, tacit and usually vague awareness that we all have, as human beings living in the world, of man, life and the world. Although there is nowadays much less adherence to Empiricist, Analytic and Positivist positions which ignore, or explicitly deny, our possession of such knowledge, and which therefore confine philosophy to clearing up verbal confusions and to stating its incompetence to do anything else, the frequently announced collapse of 'Foundationalism'—the Cartesian and 'Critical' attempt, shared also by Empiricist and Positivist philosophers, to reconstruct upon a secure and incorrigible basis what can be salvaged from what we ordinarily but naively regard as knowledge—has too often resulted in despair or rejoicing at any other constructive task for philosophy, so that all it can do is to 'unmask' or 'deconstruct' as mere 'perspectives' or 'ideologies' all other intellectual positions, while tacitly exempting itself from the otherwise universal Relativism or Subjectivism which it proclaims.

Another tendency, allied in various ways to those just mentioned, which *Appraisal* also rejects (as do Prof. Marjorie Grene and Dr David Selbourne, whose latest books we review), is that which would confine philosophy to a professional élite engaged in a closed debate about merely technical matters. On the contrary, if the starting-point and subject-matter of philosophy is our existing and inchoate awareness of ourselves, our fellows, and the world, then we all are in a position to begin to reflect upon and to articulate that awareness. All intellectual disciplines ought, in one way or another, to address the general public at some point and to contribute to a common culture, and philosophy in particular ought to address itself to the great questions of the day and of all times.

Polanyiana

Polanyiana, the journal of the Michael Polanyi Liberal Philosophical Association, has resumed publication after a break. The current issue, Vol. 4 No. 4, contains an article (in English and Hungarian versions) on Michael Polanyi and Arthur Koestler, with many quotations from their correspondence.

The next issue will contain the studies, or extracts from them, submitted in English to the Central European University by the MPLPA for its research project on 'The Central European Tradition of Liberal Philosophy'.

The annual subscription to *Polanyiana* is $US10 for those outside Hungary. Any readers of *Appraisal* in Britain who wish to receive *Polanyiana* can send their subscriptions in £ to me and I will forward them—*R.T.Allen.*

THE COMPUTER AS AN INSPIRING AND LIMITING FACTOR IN THE CONCEPTUAL DEVELOPMENT OF PSYCHOLOGY

Csaba Pléh

ABSTRACT

The paper presents a critical discussion of the role of computers in psychology. It shows that the Neuman type of computer has been a limiting as well as an inspiring factor in present day psychology. Especially, while furthering the universalistic metaphor of human cognition, it has also helped psychologists to forget about the multiplicity of thought. Recent developments in cognitive science, especially the modular and the connectionist and parallel distributed conceptions of cognition are discussed with regard to this issue. They are presented as theories where the metaphor of a serial computer with a unitary processing mode like the General Problem Solver is replaced by a multitude of qualitatively different processes.

The paper argues that the computer that promised to be the ultimate mechanical model of man, finally turned out to be a factor that has reinforced the restatement of some basic philosophical issues such as the relationship between brain and mind in the form of the new functionalism or the issue of the relative contributions of sensory and abstract knowledge to cognition. In this framework Polanyi's conception of personal knowledge is treated as a tacit partner in the new consciousness emerging from considerations of the limits of the computational mind.

1. The machine as a model of man

It is a general feature of the modern Western image of man to treat available machines as possible analogues for man. This mental pattern extends the relevance of machines to man, and interprets man as a fountain, as a clock or as a steam-engine. This tendency is supplemented, however, by two further factors. As a second step, the idea is raised that after all man is still not an engine of this type. He cannot be interpreted as, for example, a causally determined clockwork mechanism, because he is free and an initiator of actions. Secondly, the opposite idea is raised as well, according to which man creates and interprets machines on the analogy of man. Thus, when machines are treated as the measure of man, a full circle of analogies is present: the machine was built on the analogy of man, and this analogy is re-applied to man. (See, on this general issue of machines applied to man, Crosson, 1985.)

When the inspirational and limiting role of computers on psychological thought during the last 40 years is raised, one thus has to see that there is nothing radically new and specific to the computer age in the mere facts of the search for mechanical analogies and of the cyclical disillusionment with the analogies once proved to be fruitful. The new aspect concerns the *nature* of the machines proposed this time as sources for analogy. Psychologists, among other people, consider the computer as an information-processing machine. The possibility of a description with regard to information is recognised in it, i.e. the possibility of characterising our own cognitive processes in a neutral way, irrespective of the 'vehicles of thought'. This analogy is far from innocent. It is not innocent with regard to the machine since it raises the issue of equivalence from time to time: to what extent is the behaviour of the machine identical with that of man if their performance is equivalent? (For a critical presentation of the strong versus weak equivalence issue see Pylyshyn, 1984).

Neither is the analogy innocent with regard to man. It presupposes that the working of the human mind is united and unitary as well as the working of machines. It presupposes further that this unitary nature can best be grasped in *symbol manipulation*. Finally, in the beginning, the relationships between human thought and the environment are rather neglected. The environmental information is presented to human thought in early cognitive models of psychology in a rather elaborated way as well as in the case of machines. In later years this becomes the issues of perception and interface, respectively, in a unitary model of cognition.

In the following, the limitations of this analogy will be taken up from the human side: I will limit myself to issues of the modelling of cognition, putting aside other relevant issues such as the social side of cognition and the emotion-cognition interface.

2. The computer as a support of objectivity in the cognitive revolution

In reviewing the inspiring and limiting role of computers in psychology, the broadest framework is provided by a characteristic thirst for models characteristic of the

early days of the cognitive revolution in the sixties. The essence of this shift of emphasis—indeed, it may be an exaggeration to call it a revolution—was the *(re)discovery of man as a modelling creature.* And the sciences of behaviour gradually have become sciences of modelling the model-making features of the human mind.

There were many reasons underlying the changes that are listed here without further proof (see, about them, Segal and Lachmann, 1977). The prevailing neo-behaviouristic mode of thought in experimental psychology has become liberalised. It has become tolerant both conceptually and experimentally. Technological developments called for an image of man as an information-processing system. Research on perception (the *New Look* school) has pointed to the complicated interactions beyond the local stimulus that are responsible for perception. All of these paved the way for a general shift of attitude. One important element among all these factors was the role of *formal sciences and modern linguistics* in the development of cognitive psychology. (For a survey of the appeal of formal sciences towards cognitive psychology see Lachmann, R., Lachmann, J.L., and Butterfield, E.C., 1979 and Smith, 1990.)

Mathematics, computer science and linguistics played an active liberating role in the cognitive turn in the following way. For the psychologist trained in a positivistic flavour and reluctant to talk abut the inner world, these disciplines made the talk about inner worlds (in the form of abstract models) scientifically respectable without the stigma of subjectivism. The cognitive turn proved to be a winner, not by claiming direct access to consciousness and by a revival of introspection, but by coming to the analysis of the internal world through indirect routes. The availability of conceptual approaches that provided an option to talk about inner reality, while preserving the aura of scientism, was a massive support in finding these indirect routes. Earlier, during the fifties a similar role was played by information theory. The (then) New Look has borrowed concepts like coding, recoding and capacity from the formal analysis of communicative systems, i.e. from *information theory.*

The concepts of algorithm and heuristics were also been taken over by cognitive psychology from the formal sciences. The same is true, however, for the *representation theory* so central to cognitive psychology. One of the basic solutions to this latter will be the import of logical models into psychology. Propositional calculus shows up not only in the trivially most relevant domain of psychology (in the study of language understanding and in the analysis of human inference) but it plays a central role in the analysis of memory storage as well. The analysis of human knowledge took the form of a propositional characterisation. Memory, for example, re-emerges in psychology not as an unstructured mosaic of ideas, but rather as a rich pattern of predicate-argument networks that give structure to knowledge (Collins and Quilian, 1969). The formal inspiration quite naturally shows up in the logical analysis of thought processes as well. This took the form of analysing human categorisation with graphs representing a series of hierarchically embedded decisions (see, e.g. Feigenbaum and Feldman, 1963.) Likely logical models show up in the analysis of problem solving first proposed by Herbert Simon and Alan Newell (for a summary see Newell and Simon, 1972). In this approach human problem solving protocols are analysed into *IFTHEN* types of logical reasoning chains (that are, in their turn, interpreted in the AI models as production systems).

From the point of view of psychology proper in all these cases what one witnesses is a transfer of an analytic language from another field to the study of the individual mind that allows an analysis of *the structure of performance* instead of an introspective account of the inner world. The logical type of analysis reveals what man *had to do* in order to produce performances of a certain kind. This attitude is similar in one regard to the tacit knowledge proposal of Polanyi (1962). We know more than we can tell. However, while Polanyi emphasises the soft, unformalised aspects of *skill* (cf. Crosson, 1985), what is in contemporary terminology a kind of implicit memory, the early cognitive trend starts from the idea of *formal* processes (algorithms and hidden factual knowledge) unavailable to consciousness and introspection. As Dreyfus (1990) points out, however, even Polanyi becomes partially formal when he looks for the sources of tacit unverbalised knowledge in rule-like maxims. But, as Nyíri (1992, Chapter 5.) summarised it recently, in practical knowledge, or skill, there always remains an application level that is not organised like a formally statable rule. We shall see later on that this problem is similar on the action side to the now famous symbol-grounding problem (Harnad, 1990) on the input side: cognitive categories as descriptions of the world in a representational theory of mind have somehow to be anchored to the real world through non-symbolic processes (i.e. through sensation).

After a while, this formal inspiration has become a real force in research and not merely a cover-up when wandering on the muddy terrain of thought. When the seemingly machine-like explicit systems of psychologists were turned into the basis of real machine systems, it became clear that the psychological model quite frequently was not explicit enough. Thus, the psychologist was forced to make more and more explicit his model of what man does and what types of information are in fact used in a given performance. This has become rather straightforward in the

design of question-answering or language-understanding systems. One of the exciting tasks facing the psychologist here is to make explicit all those inferential chains that make our stories coherent (see for a theoretical attempt e.g. Schank and Abelson, 1977, for some experimental studies, Pléh, 1987 and the volume edited by Graesser and Black, 1985). As we shall see later on, the application of a simplistic machine model, and traditional formal models, to psychology had to face two types of obstacles, both on the 'periphery' of the model. Perception and everyday social knowledge are not that easy to model by machine, and not easy to make explicit.

It is worth noticing here a certain duality inherent in the cognitive approach. Cognitive psychology puts *information* and *meaning* into the focus of research. Meanwhile, in its purest form, it does not care about the realising system. This focus on information would inherently drive the study of cognition towards an engagement that does not treat man as a mere object. All of this implies a kind of Platonist approach that presupposes the existence of pure forms of knowledge. At the same time, however, this Platonist view of man still tends to treat man on the analogy of a machine. The main formal support and inspiring force for the cognitive movement in psychology is a machine, the computer. This feature shows some clear affinity of the cognitive trend towards the traditional mechanical view of the world that treats man as an object. Many internal tensions of the cognitive view and many external criticisms of it are in fact based on this duality: man is treated as an information system which is characteristically a non-object, while at the same time man is treated on the analogy of an object.

3. The linear computer and the information-processing approach

The most basic analogy between the point of view of cognitive psychologists and the computer has been the trend on the part of psychologists to treat human cognition on the analogy of a linear information-processing system with a single central processor (as a Neumann-type computer). The human cognitive system, like a traditional computer, was also supposed to have a limited capacity short-term storage and a long-term background store that has a (practically) unlimited capacity, and its mode of operation is basically sequential. It is rather revealing that this image of a sequential machine is usually supplemented with the remark 'with the exception of the obviously parallel sensory processes'. With this remark, however, the issue of relationships with the outside world is usually over. These sensory processes are dealt with in models of cognition only to the extent that they provide a set of features for the description of objects.

The dominant cognitive approach for about 15 years has been the theoretical framework outlined in the seminal book *Perception and Communication* by the British experimentalist, Donald Broadbent in 1958. Essentially, this position claims that human cognition consists of coding steps corresponding to more and more abstract informational features. In 'processing' an incoming word for example, first its physical features are identified (acoustic and graphic representation). The filtering function that corresponds to selective attention would work on the output of these processes. The system is parallel and has a large capacity on the pre-attentive stages. This is supposed to be followed through a sequential coding process by a short-term store of limited capacity. Then, in semantic coding, all the information stored about the given word in the long-term memory system, would be mobilised.

The stages and the stores corresponding to them are not mere abstractions postulated for the sake of organising data. The inspirational role of the machine analogy in fact appears in this aspect. Psychologists have devoted long series of experiments for the identification of each of these stages. In the analysis of the filtering function, for example, the classic issue of whether we can indeed make several things in parallel is reformulated as a model according to which, in the early stages preceding recognition, there is indeed parallel processing through multiple channels, but this is followed by filtering in the recognition stage, recognition being a strictly sequential process. The detailed experimental questions are issues about such matters as what kind of representation is being formed at levels below full recognition, what features are and could be used in filtering, and if filtering partial or total. The corresponding experimental paradigm is the cocktail party situation (continuous listening to parallel verbal messages). The characteristic experimental situation corresponding to short-term storage is false recognition of letters following a few seconds of interposed activity.

This general metaphor of the Neumann-type computer in psychology, beside its positive impact upon research, has introduced or reinforced *several limitations* in the psychological research on cognition.

1. The overwhelming nature of sequential processing

This is the most controversial issue today. We shall return to it later on, in connection with the controversy around parallel and sequential processing.

2. The idea of a limited capacity.

Our resources are limited in a unified manner. That is to say, all mental computations use the same central processor; they have to

share a common capacity. With respect to limitations, all cognitive processes can be reduced to a common denominator. This idea is still with us even in the avant-garde propositions of the seventies that made processing contextand task-dependent in contrast to the mechanical automatic image of the original model. These new proposals drew a distinction between automatic and controlled processes, the latter ones requiring effort as first proposed by Kahnemann and Tversky (for a review see Kahnemann, 1973). This new approach still preserved the basic idea that we have one limited cognitive resource that has to be divided among different tasks. To take one experimental example: while listening to sentences, the recognition of irrelevant noises is slowed down if they are presented around the end of the sentence. The reason for this is that if our resources are busy with the analysis of the sentence (this is the case around the end of the clause as opposed to its beginning) no more resources are left to analyse the irrelevant material (Fodor, Bever and Garrett, 1974). Thus, in the common denominator of the shared limited capacity, even the processing of a buzzing tone and the meaning of a sentence have something in common.

In the eighties, both the idea of parallel processing (e.g. Rumelhart and McClelland, 1986) and that of modularity (Fodor, 1983) challenged the concept of limited capacity as the common denominator in cognition. The capacity limitations, as Allan Allport (1980) very aptly put it, are not due to a neutral central processor, but are due to the competition of different tasks for the same executive system, i.e. for the control of speech.

3. The idea of stores with a fixed order and parameters

The linear-sequential metaphor has proved to be of limited validity in memory research as well. First, it has become obvious that in the classical view, a correspondence was suggested between the representational form and the temporal sequencing of stores. Left-to-right sequencing in this way corresponded to representations going from the physical code towards the more abstract ones (Table 1).

Table 1
Storage systems and forms of representation in traditional cogitive psychology

Very short	Short	Long term
physical features	name codes	semantic codes

This conception has clearly relied on the machine analogy of operative and background storage. Three considerations questioned this general image. As an opposition to the bottom-up linear sequencing metaphor of human cognition, models emphasising top-down processes, the anticipatory role of higher order organisation, showed up, most notably in the form of different schema theories. These theories are, in a way, structural formulations of the New Look ideas on perception (for a clear exposition see Rumelhart, 1980). It has also become evident that the temporal fate of different codes and representations is a flexible process. The context and the task have a determining influence over the retention of different codes. The physical code, as in the famous matching experiments of Posner (identity judgements of AA vs. Aa after 0 to 2000 ms delays) that seemed to disappear after 1 sec., may be present for longer times provided that physically identical pairs are much more frequent in the set (Posner, 1986). According to the levels of processing theory proposed by Craik and Lockhart (1972), flexible coding processes should be postulated rather than boxes with fixed parameters and a fixed sequence to account for the dominance and availability of different representations.

The working memory conception of Baddeley (1976) also emphasises flexibility in the memory system. Although it presupposes stores, it treats their interrelationships in a functional rather then in a preassigned way.

It has also become clear that in models of recognition long term stored knowledge cannot be assigned to such a late phase in the process as the original model has suggested. If we want to account for, say, the recall of words after a few seconds (the so called short term memory performance), we have to postulate a temporary activation of the knowledge store (the long term store) that is prior to short term storage. Thus, in a way contrasting with the initial model, the workings of the 'long term box' somehow precede the workings of the 'short term box'.

4. Machine and human parsimony

It is a rather peculiar way to map machine thought onto human information processing when psychologists start from machine parsimony. Furthermore, its starting point is the early machine model that considers memory to be the most expensive ingredient. Thus, all data should be stored only once. This conception had a dual effect on models of semantic memory interpreted for psychology. The paradigmatic example is the famous knowledge representation and retrieval system proposed by Collins and Quilian (1969). Due to the hierarchical relations among nodes (canary-birdanimal), predicates valid over the extension of the superordinate node are only stored by that node (e.g. the *canary* node has no separate predicate about flying; only the *bird* one has). Experimental studies to verify this model compared decision times for specific predicates (*The canary is yellow*) with decision times for generic predicates (*The canary can fly*). Early studies supported the theory: specific predicates had faster reaction

times (RTs) than generic ones. It did not take much time to realise that the real situation was more complicated. There are subordinate concepts that have fast RTs for certain generic predicates. Also, not all subordinates have the same RTs for a generic predicate (compare *Swans can fly* with *Eagles can fly*). It seemed to be safer to propose a dual parsimony principle for humans. In the human processor, redundancy is avoided, but decision speed plays a role as well. Frequently used items of information have multiple representations in the network (Smith, Shoben, and Ripps, 1974).

This modification is a good example for the general workings of the machine analogy towards psychology. We simplify the human model by using ideas derived from the machine. Later on, contradictions are observed between the simplifying model and experimental results. Therefore, the oversimplifying model is refined. However, both cases are fruitful in this process. Simplification and revision are two steps in the endless process of approximating reality, and flirting with the machine and disillusionment from the machine are examples of this dyadic process. In the reversed engineering conception of cognitive research, the ideal is to start from the performance (the behaviour) of the system and find actual computational models that are increasingly more and more equivalent in their performance to the human mind. The approach, as fruitful as it is, certainly has to face some basic boundary limitations: e.g. could we ever try to include experience into the reversed engineering model? is modelling the physical structure (the body) a realistic aim? (See about these matters Harnad, 1993). But even at levels where reversed engineering works, it seems to be that the actual computational models it uses tend to show this circularity of being inspired by actual machines and then rejecting the given machine model for a new one. Polanyi (1967) in his antireductionist stand concerning life implies in this respect that although we have to look for mechanical (machine) models what we shall actually find will be mechanisms that are irreducible boundary conditions as in the case of life interpreted in a physico-chemical framework. Similarly, when he discusses the logic of tacit inference (Polanyi, 1969) he claims that when we are using a machine model for the functioning of the body we are using only some aspects of biological functioning: the function itself cannot be understood starting from the machine.

4. Uniformity and multiplicity in thought

Some of the most basic dividing factors in contemporary cognitive psychology could be related to the way psychologists deal with the initial simplified linear machine as a model of cognition. Two characteristic and central dividing factors will be selected here: the issues of *uniformity* and *sequentiality* in thought. In both cases the rival alternatives, while questioning the modelling value of machines with a traditional architecture, are themselves flirting with a new machine metaphor or implementing new architectures. Thus, they preserve the fruitful relationship between the study of man and machine in a productive way: they turn to new types of machines to support human imagination.

With regard to the uniformity-multiplicity issue, from the beginning of cognitive psychology onwards, one can witness two different attitudes. The idea of uniformity was always dominant from the beginning of the empiricist movemen, but from time to time there were strong voices of dissent claiming for a more quality-oriented view of the human mind postulating several different types of mental processes. The first modern multiplicity-view was presented by classic faculty psychology (see papers in the volume edited by Smith, 1990). In modern times this attitude was revived by Sir Frederic Bartlett (1932). His multiplicity centred approach was aimed to demonstrate, for the behaviourists as well as for the classical associationists, that the human mind was too complicated to be accounted for by simple principles of connection formation. Similar ideas were proposed in early cognitive psychology by Neisser (1963) in a largely forgotten paper where he emphasised that in psychoanalytic theory, in the experimental study of cognition, and in the study of human abilities as well, one can differentiate between ordered and unordered, and analytic and holistic, cognitive processes as basically different qualities.

At the birth of modern cognitivism proper, the opposition between uniform and multiple views of mental functioning took the form of an opposition between associative and structure-bound processes. Chomsky and his followers opposed learning-theory principles and language-acqui-sition, the latter being characterised by self-organisation and rule-formation rather than association and environmental determination (Chomsky, 1959; Miller and Chomsky, 1963). This was the beginning of a straightforward dual view of mental functioning: there is a mechanistic and a structure-dependent mode of functioning, the latter being characteristic of complex forms of behaviour like language. From the late sixties on, however, this view of dual organisation was gradually replaced by an overextending unifying trend within the same structuralist group. Specifically, a conception took form suggesting that all interesting processes should be dealt with within the framework of structuredependent processes and hypothesis-testing rather than association. This conception replaced the duality of learning versus rules, association versus structure, with an overwhelming structuralism (Chomsky, 1968, Fodor, 1968).

This view is not shared by the

entire cognitive trend. It has become gradually clear that the cognitive trend is radically divided in this respect, and its division might very well be interpreted as a division of opinion with regard to the explanatory power of the unitary symbol-manipulating machine. The idea that human cognition follows basically the same principles everywhere has taken definite form. Its most explicit version is the theory proposed by Herbert Simon and his associates. According to this view, the very same general problem solving principles show up in every human process. Human cognition applies so-called production systems for repeated problems. Their structure is the same in every problem area. They look for an input of a certain kind and from there on they arrive at a result as a consequence of a chain of reasoning. The neutral language of the human mind is the language of logical calculus. The train of thought is the same, be it a question of playing chess, understanding a sentence, or even typewriting. At the same time, our entire background knowledge system participates in all cognitive processes (Simon, 1979; for a new synthesis along these lines, see the volume of the late Allan Newell, 1990).

Parallel with this process, the conception inspired by generative linguistics—itself being a unifying conception for a while—gradually became the most straightforward opponent of unitary theories. This took the form of the modular conception of mind based on the model of modern linguistics. Modularity, of course, has several sources and interpretations, from the electronic idea of modules to the concept of neurological modules as put forward, for example, by Szentágothai (1975). One of the basic features of this proposal is the emphasis on a unitary kind of organisation in the cerebral cortex. The units in the cortex serving different functions basically have the samc kind of organisation, and that makes their genetic determination feasible. They function as a highly interrelated structure, with much more internal connections than external ones and they communicate to other modules only the net result of their computations. The cognitive modularistic conception takes over the idea of encapsulated computational units from the neurological doctrine. It proposes, however, modules of a much larger scale. And at the same time, it puts the emphasis not on the identity of structure, but on the qualitative differences in function.

This conception was most clearly articulated by Jerry Fodor (1983). He started from a critical analysis of a specific shortcoming of the classical 'unified' information processing positions, namely that they are in trouble with regard to the specific processes of perception. They just suppose that the analysis gives a description of the stimuli in terms of features but the features and the process of their extraction are not clearly articulated. For an information-processing paradigm there are no quantitative differences between the treatment of data from vision and from audition, for example. However, states the modularist claim, no matter how similar the processes are for an algorithm of concept-formation, if it has to do with a series of decisions based on visual features, when it has to classify objects into the categories of triangle and square, and when it has to classify certain stimuli as salted and sweet, these are only apparent similarities (as any phenomenologist should have known for long). Patterns can be regarded as equivalent to each other only to the extent that they are treated as descriptions. In this case, however, nothing was said about the way the descriptions themselves were obtained.

Phrased in the idiom of the machine world, the conception of modularity brings into focus an important aspect left out of traditional machines: the relationship between representations and the outside world. Fodor (1983) and Pylyshyn (1984) define as'input systems' those supposed neural components that perform the separate tasks of coding incoming information independently of each other in a prototypically genetically determined order, and interact with each other only at the level of their outputs. These micro-machines are encapsulated, and also impenetrable in the sense that knowledge, in the traditional meaning of the term, cannot influence their workings: i.e. knowledge, expectation and other contextual factors, have an influence only on the results of these computations and never on their inside. This conception is generalised by Fodor into an overall modular view of the human mind. Not only perceptual processes, but most of our cognition in general, is organised according to the principle of input systems: most of the human mind consists of encapsulated, task-specific modules that fulfil their tasks with remarkable speed and in a reflex-like automatic way.

From the point of view of machines, this conception suggests that several task-specific small processors co-exist in our mind, and the results of their computations are made available to a symbol-manipulating system which is like a General Problem Solver. Thus, a little room is left by Fodor for experience-bound general cognition. But this is done unwillingly and the domain of this factor is gradually narrowed.

Concerning language, for example, the modular thesis claims that context or frequency has no effect on the immediate mechanisms of word-recognition. These factors have only a *post hoc* effect modulating the ease of word use. In a similar vein, there is no interaction between the lexical, syntactic and semantic components of understanding, all of which operate as self-contained systems. Interactions appear only on the level of their outputs, on the level of the results of their computations. The two diverging views here are the en-

tirely interactionist view (this isthe one supported by the general cognition idea) and a modular view that postulates totally autonomous subprocesses. For a presentation of the views and the supporting empirical evidence see the volume edited by Garfield (1987).

Historically—and Fodor is very conscious of taking up here a classical position—this conception is a revival of the eighteenth-century faculty psychology of Franz Gall. (Concerning this aspect of Fodor, see Pléh, 1985.) And this one, in its turn, leads us towards classical theories of the multiple action views of the human mind, including not only Reid and the Scottish school, Kant and Leibniz, but even the scholastic teaching about faculties. (See Smith, 1990.)

5. Polanyi on multiplicity

In this broad context, the multiplicity view proposed by Polanyi (1962) for human nature is rather parallel to the modular view. When he proposes different approaches to the world, among them a rational-discursive (propositional) and a more experiencebound, sensation-oriented approach, he stands on the same side with the multiplicity people and with the claim for qualitatively different cognitive mechanisms. Polanyi, inspired by, among other things, the organisational ideas of Gestalt psychology (see Ujlaki, 1992) in his anti-mechanistic and anti-elementaristic world-view, quite naturally stood for a multiple view of mental functioning in several respects. He postulated—without considering the machine-minded cognitive trends, but seriously considering and challenging the neo-positivistic implications of cybernetics (e.g. Polanyi, 1969)—that beside explicit processes which can be reconstructed in a logical way, one has to suppose several types of hidden processes like the ones underlying skill, the empathic understanding of movement, intuition, etc. (Polanyi, 1968).

Dreyfus (1990) clearly presents Polanyi as a precursor for a dynamic view of cognition. According to this view, in contrast to the entirely Socratic representational theory of cognition, in which all our mental acts should be composed of operations clearly statable in terms of explicit rules, real skilful cognition is always based on tacit knowledge and non-focal consciousness. Not only do we always have non-stated sides about what we know, but the important cases of knowing how (as opposed to knowing what) are based on maxims rather than rules. Polanyi is on the side of a view on cognition where knowledge and the process of knowing are more than what is statable in the form of a propositional calculus. We should add that his emphasis on non-explicit thought is not irrational, but includes large sub-symbolic components (perception) and tacit inferential components (practical inference).

6. Innateness versus experience

The modular approach is put forward as an opposite of and a challenge to the New Look view of perception that has played such a great role in the formation of the entire cognitive movement. The essential component of this conception was that perception depends on several factors beyond the local stimulus: our expectations, frequency, the actual context, and the motivational significance of the signs, all contribute to perception (Bruner, 1957). The modular conception, on the other hand, treats perception as impenetrable from knowledge (Pylyshyn, 1984). In this sense, it denies continuity between different levels of cognition. In this respect, the computational theory of vision proposed by the late David Marr (1982) shows some parallels with the modular conception concerning the structure of the perception of form and space. Marr also proposes a solution with algorithmic steps that are independent of experience with the given individual pattern and context. Note that, in this process of reconstructing the process of human vision, he also relies on machines in two respects. First, the general framework of the computational theory proposes that the first step is to clarify the logical structure of the task to be solved by any system capable of vision independently of the properties and constraints of the system. This is followed by a description of actual algorithms and by the implementation of these algorithms in humans and in machines. On the other hand, Marr also relies on machines in everyday research. He is constantly willing to learn from artificial modelling (sometimes from its failures). Vision is understood if we are able to reproduce it.

From a historical point of view, all of these debates and divergent opinions remind us of the debates that have been with us at least since the middle of the eighteenth century regarding the experiential and innate theories of the perception of space. There is one new feature, however, most clearly seen in the work of Marr. The idea of formal analysis brings in an abstract level into the structure of research endeavour. The first task of the scientist is to make a conceptual analysis of the problems, which is the 'computational level' of the logics of research. There is a one-to-many relationship between this level and that of the algorithms actually used by humans. In principle, several realisations could be made available for the very same computational theory. That is the reason why the work of Marr has become a model in cognitive science. This is a new pattern, slightly different from, for example, that proposed by early researchers into artificial intelligence, where the starting point was the programme rather than the structure of the task. It is a conscious realisation of the thought-provoking possibility of machines on an abstract level.

A similar conceptual model char-

acterises modern linguistics. There too, the need for an abstract characterisation of the system and the products is emphasised, preceding considerations of the actual implementation of this model in the performance of human speech. In language this duality is between the abstract theory describing a generative grammar of the given language as a theory of competence that is to be followed by theories of how competence is used in actual performance (Fodor and Garrett, 1966). This duality, originally proposed by Chomsky (1965) some thirty years ago, sometimes even takes the form of a triple system resembling the three levels of approach proposed by Marr. The grammatical theory, the theory of competence, is to be followed by a theory about the algorithms representing that competence (this is sometimes referred to as a theory of competence mechanisms, sometimes as an abstract theory of performance, e.g. by Fodor and Garrett, 1966), and this is followed in its turn by an account of the actual processing and production of sentences, in a theory of performance (Watt, 1970). However, and this is a very important similarity, as Miller and Chomsky (1963) have emphasised a long time ago in connection with language, the theories of performance (say, the implementation level) have to embody a grammar, a theory of competence. Or, to put it in another way, a theory about the requirements that any implementation theory has to fulfil (Pylyshyn, 1972). In all of this, in the case of language for a longer time, and in the case of vision quite recently, the machine, rather then giving with its physical constraints a limitation to the fantasy of the researcher, stimulated the need for an explicit abstract theory, which is the last formal kind of inspiration taken over from issues of computation.

7. Sequential and parallel

During the last decade another approach has appeared. It is a rival both of traditional architecture and of modularity. In respect of its psychological content, this is a radical neo-associationist way of thought (cf. Pléh, 1991). Similar to classical architectures (for the characterisation of psychological models as 'architectures', see Fodor and Pylyshyn, 1988) it advocates a unified and unitary model of cognition. The key to this unity, however, is not the language of logics and rules, but the world of networks based on mere contiguity. The unity of cognition is provided by a theory of representation where only facilitating and inhibiting connections exist between the nodes interpreted as abstract neurons. The apparent complexity of mental life would be explained by the fact that nodes and sub-networks coexist in different levels (e.g. in the case of word recognition, on the level of features, letters, and words) and the connecting lines between them may produce rather different activation patterns: there is top-down as well as bottom-up and collateral facilitation (reading *pr* facilitating the processing of *o*) as well as inhibition (reading *pr* inhibits the activation e.g. of *a* and *u* in English). Knowledge, however, remains, in all of these additions, a mere activation of a partial network (McClelland, 1988), rather than the application of rules. In the case of reading, for example, the classical approach would interpret effects such as an easy reading of *pro*as the result of phonetic rules, while the connectionists stand for a system where only connections between an enormous number of individual units are postulated. This is the basic content of the new approach with regard to representation that is aptly referred to as connectionism. (The starting volume for the connectionist movement is Hinton and Anderson, 1981. For an exposition of the development of the movement and its relationship to other approaches see the volumes edited by Pfeiffer et al., 1989 and by Brink and Haden, 1989.)

This is supplemented by the fundamental idea of an overwhelming parallel processing (Feldman and Ballard, 1982). This is exposed most explicitly in the conception of Rumelhart and McClelland (1986; McClelland and Rumelhart, 1986). This feature gives the name of the most active group within the more general connectionist movement: *PDP, Parallel Distributed Processing*. The parallel processing appears in this framework not as a particular issue (e.g. the extent of parallel processing before and after attention as it was treated in classical cognitive psychology from the Broadbent model onwards), but as the basic feature of all cognition. In the familiar case of the recognition of written words, for example, starting from the letters, all possible word candidates (i.e. all words that are consistent e.g. with the first letter) would be activated in parallel, and the most probable one would gain victory due to multiple activation (both from each of its letters and from the activation of all partially compatible words as well) and also because the rejected candidates would get inhibitory input as well (e.g. from the non-corresponding letters).

This new and ambitious project has generated several penetrating discussions (see the volumes edited by Pinker and Mehler, 1988, and by Pfeifer et al., 1989). Two of these critical aspects are relevant in the present context. First, while the theories of parallel processing try to get rid of the metaphors inherent in traditional machine architecture, at the same time they try to prove their theories by constructing new machines. In their research strategy they rely on three foundations. They try to use data from psychology, the construction of machine networks, and neurology. The connections among neurology, the computational networks and mental processes is assumed to be rather strict. In this respect, they take up the lead of McCulloch and Pitts (1943). In the original proposal the binary neuron was made responsi-

ble for the unitary binary principles of logics in mental life. This neuronal structure explained the most basic scaffolding of the mind. In a similar reasoning, in the new conception, the most general argument in favour of overwhelming parallel processing is found in the famous so-called *100 steps rule* (Feldman and Ballard, 1982). Essentially, this rule again claims a fundamental isomorphism between neuronal and mental organisation. It relies on the insight that the neuronal parameters of individual neurons are in the order of magnitude of milli-seconds, while usual human recognition and reaction times are in the order of hundreds of milliseconds. The relationship of their magnitudes does not allow more than 100 steps to be arranged in a sequential manner in the modelling of any given cognitive task. Since most of the usual conceptions imply much more elementary steps, one has to suppose parallel processes.

Following this lead, on the one hand, massively parallel machine architectures are constructed (see for this Fahlhman, 1988, and the full volume edited by Kowalik, 1988). From the psychological point of view, in the theoretical models, the abstract neurons and their connections represent directly the coding steps involved in given cognitive tasks (McClelland 1988). Thus, the now traditional flirting between machines and psychology is strongly reinforced here. Simply, machines with a classical architecture are replaced by machines with new architectures as the inspiring forces.

Another important aspect is the sincere negative approach of connectionism to the issue of rules in mental life. Rumelhart and McClelland (1986, 1987; see also McClelland, 1988) are especially clear concerning language that they treat rules not as internal, inherent laws of the mental system but as external characterisations of the products of the mental system. The task, they claim, is exactly to produce mechanical models that imitate, and thereby eliminate from the proposed mechanism, the level of rules. One does not need too much fantasy here to realise that these proposals have lead to the unearthing of old animosities in linguistics and psycho-linguistics (see the Pinker-Mehler, 1988 volume, and especially Fodor and Pylyshyn, 1988, and Pinker and Prince, 1988, which, by the way, figures in that volume as well). It is of interest on the conceptual level, that the storms raised around and by the connectionist camp have revived and combined two debates that have been with us for decades. Are rules inherent in the machines or are they present only in the mind of the designer, and—transplanting the machine model to man—are rules indeed in the human mind or are they only to be found in the researcher's model about behaviour? The classical issue of the mental reality of rule systems, as proposed by e.g. generative grammar early on (Chomsky, 1968; Fodor, 1968) not only gets a fresh start in this new debate but it becomes a parallel epistemological issue of the rule governed nature of machines and man.

8. Rules or connections: Some basic criticism of connectionist models

The most interesting criticisms of connectionist theory all concentrate on the issue of structure versus elements. On the contemporary scene this takes the form of discussions about whether there are rule-governed actions not reducible to simple connections or the formation of habits. The generativist-structuralist critics point out that their traditional rejection of associative and stochastic principles for the explanation of language (Chomsky, 1957; Miller and Chomsky, 1963) is still valid concerning the neo-connectionist models, and that these latter ones are unable to deal with rule-like regularities in a clear way.

Table 2

The juxtaposition of connectionist and classical cognitive achitecture according to Fodor and Pylyshyn (1988)

Connectionists	Classical View
nodes	descriptions
only causal relations (history of excitation)	rich relationships (language of thought)
excitation paths	rewriting rules
units independent of structure	entities dependent on structure (constituents)

Fodor and Pylyshyn in their criticism of connectionism, summarised in Table 2, clearly show that the basic limitation of connectionist models is their lack of structure. One can characterise this feature in several ways:

> Models based on patterns of (co)excitation cannot differentiate between two concepts being active simultaneously and them being in a given relation (like IS, PART OF etc.). This was supposed to be solved by labelling the graphs in 'classical' network models.

A connectionist representation has no clear syntax (lack of structure).

The associationism of connectivist models situates the human mind at the mercy of the arbitrary and unsystematic nature of the world: it allows any connections whatsoever. It is worth remembering that Max Wertheimer (1922), in one of his theoretical papers on Gestalt psychology, has also criticised what he called the 'mere existential relationships' responsible for association in classical associationism. No essential or meaningful relationship had to exist between the elements to be associated. He proposed that this was only true in extreme situations,

in certain limited cases. On the whole, it had to be replaced by meaningful and top-down organisation in mental life. As Fodor and Pylyshyn put it today: 'All it (i.e. the connectionist model) can do is to build an internal model of redundancies in experience by altering the probabilities of transition among mental states' (Fodor and Pylyshyn, 1988, pp. 49-50).

In order to account for the systematic nature of mental phenomena one has to go beyond this, to postulate structure-sensitive mental processes. The basic flow of connectionism is parallel to the flow of all associationism from the classical associationist accounts of the mind through Hull (1943) to Hebb (1949) and Osgood (1963): in order to reconstruct semantic coherence of thought, associations, being insensitive to structure, have a hard time. Why not to give up their explanatory power and replace them with internal organisation, or structure?

In a more permissive formulation, Lachter and Bever (1988) also conclude that connectionist accounts are associative in nature. However, being parallel and allowing for associations at different levels, they provide for an enormous amount of elementary habits. But habits never become rules. However, 'It is equally obvious that some actions are habits' (Lachter and Bever, 1988, pp. 243-244). Connectionism can be claimed to be a description of that (lower) level of behaviour. Habits and rules should still be differentiated, as it was proposed a good 30 years ago by Chomsky (1959), and the associationistic account should be reserved for habits.

9. Simple and complicated in a new light

It was mainly the pitfalls of machine modelling and artificial intelligence (or, to be more kind, their difficulties) that directed attention towards the fact that, earlier on, human specificities have been viewed in a short-sighted way (see about this Dehn and Schank, 1982). Classical psychology has developed a rather definite sensualistic conception concerning the relationships between simple and complicated. The more something is independent of any stimulus, the more complex and higher the process is postulated to be. According to this image, working from the outside world inward, the achievement becomes more complex and more human. Thus, thought is assumed to be more complex and more human than perception. However, machine intelligence has shown us that processes that were supposed to be of the highest and most human type were rather easy to model even if not necessarily in a strongly equivalent way. Logical calculus and mathematical proof are classical examples of this effect. And psychology traditionally treats them as higher order functions specific to humans.

As we have alluded to it several times before, perceptual achievements are rather difficult to model on machines. Just think of the modest results in computer vision or the automatic understanding of speech. This factor concerning machines had a twofold inspiring force with regard to psychology as well. First, psychologists have realised that the world of transducers, i.e. the sub-systems assuring our connection to the outside world, are in fact very complex and complicated. Their contribution to the representation of the outside world cannot be dealt with as a trivial achievement (by just stating that they somehow provide a description of the stimulus), since exactly the birth of this description becomes one of the key issues.

Besides becoming the driving force behind the general popularity of modular models, the complicated nature of perception has led to the realisation that seemingly low-level systems may be specifically human as well. All that we witness, especially in connection with modularity, the computational theory of vision and the complexity of 'low level vision' (Julesz, 1984) means that the relationship between simple and complex is certainly not trivial in human beings. Anything human can be as complicated as the most abstract levels of complexity. It is of non-trivial interest here that, besides the perceptual world, in fact the mundane reality, our everyday knowledge and its use in understanding the world proved to be the most challenging for AI models (Schank and Abelson, 1976; Schank, 1982). We may have to conclude that even in the study of cognition one has to overcome human vanity. We may have to admit that what seems to be specifically human (e.g. the effect of language on the perception of the world) is specifically human, not because of its being complicated, but because of its intersubjective, communicative origin.

The apparent non-success of the machine model in modelling non-algorithmic and non-propositional knowledge, and the resulting reconsideration of what is specifically human and of what is simple and complex, have some natural resonances in Polanyi's work. The whole idea of tacit knowledge being central to our knowledge of the world, the emphasis on skill and perception, are all consonant with the view that cognition is based on non-conscious schemata. Historically, it is rather telling that, not unlike Polanyi, Frederic Bartlett (1932) who first formulated the schema-based view of cognition in psychology, also started from the analysis of skill and bodily functions and extended this analysis into an overall conception of knowledge and memory.

10. Body and mind in the light of machines: functionalism rediscovered

Relationships between man and machines processing information placed the issue of the relationships between body and mental life in a new light, more precisely, in a

rediscovered new light. In the terminology of Kripke (1972), the contemporary question can be phrased as follows: Can one postulate a rigid designation between the mental and the physiology of the human brain (is there an identity there that is valid over all possible worlds?)? Or—in a more traditional idiom—is there a strict identity between human cognition and brain processes, given that machines are also able to show evidence of intellectual achievements? The answer to this question has basically led to a renewal of the classical Aristotelian type of functionalism (see Fodor, 1981, and the volume edited by Block, 1980). What used to be soul as the form of bodily functions appears now as soul being a program, a set of instructions to run the system, the software running on the hardware, if you like it.

The following line of thought led to the renewal of functionalism in the traditional sense as regards relationships between man and machine. Is the 'machine thought' characterising the information-processing approach indeed a viable route to the understanding of the human mind? Specifically, is it a sensible approach to take experimental data— mainly reaction-time data—as the point of departure and to postulate on the basis of them a stupid machine that only takes into account the factors under the control of the experimenter and would produce similar results following algorithmic steps? When this kind of model was found, it was assumed that an explanation was created for the procedure under study. Functional models in an information-theoretic sense claim to be explanations while in fact they are only new descriptions. Arguments external with regard to the functional model are in fact needed to show that the subjects had indeed followed the proposed strategy. When looking for explanations, one has to leave the traditional framework of the cognitive laboratory and use biological functions and culture. The information-processing model certainly presents important constraints: it clearly shows the kind of complexity the explanatory factors (e.g. psycho-physiology) have to cover. It does not show the explanatory solution itself, however.

Different types of *explanatory biological models* showed up, basically corresponding to the two usual types of biological reductionism in psychology. One takes a long step and looks for evolutionary bases of behaviour, while the other takes a shorter step and looks for the immediate physiological mechanisms of behaviour. One of them is the evolutionary theory of human cognition. Interestingly enough, on the contemporary scene there are some attempts to combine the two approaches, and have not 'merely' a Darwinian explanation of behaviour over the range of millions of years, but try to extend the selectionist paradigm to neuro-physiology proper over the range of a lifetime. Evolutionary theories of cognition are proposed in the sense that Darwinian type of selection is postulated in the individual somatic life to account for the genesis and stabilisation of neural circuits responsible for cognition and perception (Changeux, 1983; Edelman, 1987).

The other kind of biological interpretation, a direct physiological interpretation that treats cognitive mechanisms as reducible to discrete physiological events in the brain, have also undergone important changes lately. The traditional silent assumption of all psycho-physiology, especially electro-physiology, had always been a kind of event reductionism when we were looking for neural correlates (?), equivalents (?), or bases of given cognitive events that result e.g. in pressing a RT button. However, in the idiom of the last decade, this type of physiologising has been supplemented, or—in the eyes of some—even replaced, by a new type of neural talk. Cognitive mechanisms are expressed as computational solutions that function over networks composed of abstract neurons. This approach is most characteristic of the connectionist group (see McClelland, 1988, and Pfeifer *et al.*, 1989 for a clear expression of this) where it is overtly claimed that this new kind of thought speaks a neural language rather than a *lingua mentis* of Fodor (1975). Thus, the traditional paradigm of a symbol-manipulating machine is replaced by a sub-symbolic brain language. It is of course a matter of debate whether this brain talk is really about the neurons, or about a sub-symbolic level interpreted in an abstract and not necessarily neural way (see about these solutions Smolensky, 1988 and Clark, 1989). Are not the abstract neurons simplified to the extent that it becomes unlicensed to talk about the brain at all? The details are not important for us at this place. It is relevant to note, however, that this kind of new brain-talk—both the evolutionary and the network version—has become so prevalent during the last years that a new materialist approach has evolved in the philosophical interpretation of these developments that sometimes refers to itself as neuro-philosophy (Churchland, 1986; about the interpretation of this new eliminativism, see the volume edited by Lycan, 1990).

This kind of direct biological anchorage is, in a way, counterbalanced by a functionalism that is biological in principle but mental in practice. There are several versions of functionalism on the contemporary cognitive scene. One is a type of machine functionalism that basically claims that type-identity is untenable: a mental event is not necessarily always identical to the same brain event. A better image for analysing body-mind relations would be to see a similarity to the relationships between functional states of a machine and the corresponding (variable) physical processes (Putnam, 1960; Fodor, 1968). Another type of functionalism is nearer to the classical biologically

minded one: it starts from the apparent teleology of mental life. (See about the different varieties of functionalism, Block, 1980, Lycan, 1990, Putnam, 1989).

Not ignoring the important divergences in the interpretation of functionalism, one could summarise this line of thought in the following way. Mentality is indeed a form of organisation of neural processes. It has to be characterised functionally as a certain type of information-processing. But that requires us to look later on for the actual systems that accomplish this performance and how they in fact do this. Beside the computational theory of Marr (1982), already referred to several times, the functionalism put forward by Jerry Fodor (1981) also belongs to this camp. Beside proposing a physiological metaphor, this approach also reverses the traditional line of influence between physiology and psychology and questions the reductionism inherent in some of the physiologising models. In particular, this approach assumes a three-level analysis of cognition. The computational analysis (or, the grammar, to that effect) characterises the task which has to be done in order to understand sentences, recognise shapes and so on. The algorithmic characterisation that (among other facts) relies on data from experimental psychology, provides us with a functional analysis of the *how*: how do we proceed to solve the given task phrased in terms of information processing stages? Finally, this is implemented by psycho-physiological research and computer science (for machines by the later and for humans by the former).

A simpler, although misleadingly simplifying, metaphor to characterise this is the application of the software-hardware opposition to the study of mind-brain relationships. The psychologist is presented as the scientist of natural softwares who is looking for a functional description in terms of symbol-manipulation. Doing this, he presents the physiologist with a task: show how the given hardware is able to solve the given processing? It also raises the issue whether our machine is an entirely flexible one with no or very little system implemented (the traditional empiricist view that is partly echoed in contemporary connectionist models) or a machine that has a built-in basic processing mode that is part of its hardware (binary logic, universal grammar, the language of thought, or what you like, as part of our nervous system.) This corresponds, of course, to the traditional rationalist-innatist position. We are able to use other modes of information-processing, other softwares, but the latter would all be secondary processing modes (see Block, 1980).

11. Eliminativism and the varieties of functionalism

The different interpretations of functionalism, of course, touch upon many other issues. First, in a historical sense, the diverging interpretations make sense historically as well. Functionalism has its direct antecedents and sources in mentalistic trends like Chomsky's philosophy of language and the neutralising attitude of cognitive psychology and cognitive science that tries to make as few substantial commitments as possible. Moreover, an equally likely and traditionally even more naturally given antecedent is the Darwinian functionalist psychology and biology. Some of the controversies are due to the fact that many functionalists (like Fodor, 1968, 1981, Block, 1980) are 'structuralists' in a traditional sense, while others (e.g. Dennett, 1991) are inheritors of a Darwinian functionalism with an eye on the origin and function of intentions and teleology in the mind rather than just mental structure interpreted in itself.

Another background feature of the controversies is that while most of the functionalists combine a psychological or computational and a philosophical training, most of the advocates of a reductionist-eliminativist way of thought are people who are socialised in the practice of physiology and the neurosciences. Because of these differences in tradition and in the content of the respective theories, the two approaches present themselves as divergent ways of thought regarding progress in science as well. Basically, in the eyes of most functionalists, the proper way to study cognition is from psychology (linguistics, cognitive science etc., the softer disciplines) towards physiology. It is the psychologist who gives the task to be explained to the physiologist. The reductionist camp, on the other hand, presents a rejuvenated version of the traditional belief that the proper way to study cognition is from physiology (neuroscience etc.) to psychology, and it is the study of brain that should have the keys and the strategic lead in understanding cognition. The physiologist presents tasks and solutions to the psychologist, and to that matter, even to the mathematician as the debate between Changeux and Connes (1989) shows where the neurobiologist claims for a neural Darwinist solution even the origin and development of mathematics.

Regarding the issue of research style and how to proceed in science, the new functionalism and the (ideal) computationist approach broke with three traditional lines of thought. They got rid off an alternative inherent in old-fashioned AI research that builds strictly from the bottom up. They have broken with physiologising in a bottom-up way, and (most of all) with any fetishistic usage of reaction-time data gained from traditional experiments in cognitive psychology. With the computational level of analysis, they present an abstract characterisation of the performance (the functions) as the primary starting point for a study of cognition. Logical analysis has to proceed empirical data-gathering. The starting point should be neither a given machine nor a given programme,

rather the logical analysis. With respect to the research affinity between machine thought and human cognition, functionalism, in the general sense and the need for an abstract analysis (computational models), has proved to be the most promising kind of machine inspiration in psychology and in cognitive studies in general. In our present belief, this kind of functional characterisation indeed gives a neutral description of knowledge that is independent of the presently prevalent metaphors and models. In the dilemma of either machine or information, the abstract, Platonistic inspiration regains its deserved position (that, by the way, was there from the beginning in the work of people like Frege). We are not entitled to believe, however, that after a while this neutral model won't prove to be just another metaphor.

With regard to Polanyi, this kind of anti-reductionism is in accord with the idea so forcefully projected by Polanyi (1962, p.31, see also Dreyfus, 1990) that it is a primary task to try to reveal the structure of the silent, unconscious skill of the expert when we aim to understand a cognitive function rather than to start with some preformed conception about that structure.

One specific historical remark is relevant here. The kind of functionalism proposed today was prefigured in terms of analytic philosophy by Ryle (1949) and by the Hungarian psychologist Schiller (1947) half a century ago in a behaviour-centred context. Although they not talking about functional models of the mental (and therefore, Ryle is usually criticised as a logical behaviourist), their functionalism sounds very modern if we trade behaviour for cognition. To phrase it in everyday terms, their view corresponded to a naive materialism that is Darwinistic on the one hand, and despises direct minute physiological reductionism on the other:

(i) mental phenomena are realised of necessity by some material system;

(ii) this is not necessarily the human nervous system, however;

(iii) the very same function could be realised by different material systems (think of humans and computers both doing simple calculations);

(iv) thus, as a research strategy, the function has to be first clarified and then it is the task of the biologists to find out what part of the brain could fulfil this function and how it might work.

This kind of functionalism was born far away from actual computer science, most of all in the realms of philosophical discussions. However, for the computer-minded cognitivist this is quite good. It represents the practice where someone could be a dedicated software person without paying any attention to the hardware. The psychologist can also be happy with this kind of functionalism. One can stay objectivist but still keep the independence of the trade. Furthermore, the psychologist can also entertain illusions that it is actually he who will show the physiologist or the UFO expert what kind of material systems to look for, namely ones that are in principle capable of realising the given functions described by psychology.

* * *

Thus the overview of the way psychology has been flirting with computers during the last 40 years brings up the following morals. Psychologists are inspired and disillusioned in a cyclical way by analogies with machines. At the end of each cycle, however, it comes out that the technological metaphor finally leads back to the basic philosophical questions. Classical issues become relevant again: such questions as, Is the human mind unitary or divided? Is man a conglomeration of accidental habits or is the mind to be characterised rather by given rules that are at least partially *a priori*? What is the correct interpretation of the relationship between body and mind? We end up in a peculiar intellectual situation where the most technically minded excursions in contemporary psychology lead back to Plato, Aristotle, Descartes and the classical issues of epistemology and psychology in general. Thus the machine was unable to put through an entirely technological metaphor. Polanyi would be glad to see that machine-minded thought turns against itself towards a revitalisation of classical issues in the humanities.

Department of General Psychology
Loránd Eötvös University (ELTE)
Budapest

Reprinted, with permission, from ***Polanyiana****, Vol. 2 No. 4 (1992) & Vol. 3 No.1 (1993) (joint issue).*

References:

Allport, D.A. (1980): 'Patterns and actions: Cognitive mechanisms are content specific'. In: Claxton, G.L. (ed.): *Cognitive Psychology: New Directions.* London. Routledge and Kegan Paul, pp.112-153.

Anderson, J.R. and Bower, G.H. (1973): *Human Associative Memory.* Washington: Hemisphere.

Baddeley, A.D. (1976): *The Psychology of Memory.* New York: Wiley.

Bartlett, F. (1932): *Remembering.* Cambridge: Cambridge University Press.

Block, N. (ed., 1980): *Readings in the Philosophy of Psychology,* Vol. 1. Cambridge (Mass) Harvard University Press.

Brink, J.R., and Haden, C.R. (eds., 1989): *The Computer and the Brain: Perspectives on Human and Artificial intelligence.* Amsterdam. North-Holland.

Broadbent, D. (1958): *Perception and Communication.*Oxford: Pergamon Press.

Bruner, J. (1957). 'On perceptual readiness'. *Psychological Review,* 64, pp.132-52.

Changeux J.-P. (1983). *L'homme neuronal*. Paris: Fayard.

Changeux J.-P. and Connes, A. (1989). *Matière a pensée*. Paris: Jacob.

Chomsky, N. (1957). *Syntactic Structures*. Hague: Mouton.

Chomsky, N. (1959). 'A review of Skinner's *Verbal Behavior*'. *Language*, 35, pp.26-58.

Chomsky, N. (1965). *Aspects of a Theory of Syntax*. Cambridge, Mass: MIT Press.

Chomsky, N. (1968). *Language and Mind*. New York: Harcourt.

Churchland, P. (1986). *Neurophilosophy*. Cambridge, Mass.: MIT Press.

Clark, A. (1989). *Microcognition*. Cambridge: MIT Press.

Collins, A. M. and Quilian, M. R. (1969). 'Retrieval time from semantic memory'. *J. Verbal Learn . Verb. Beh.*, 8, pp.240-247.

Craik, E I. M. and Lockhart, R. S. (1972). 'Levels of processing: A framework for memory research'. *J. Verb. Learn. Verb. Behav.*, 11, pp.672-684.

Crosson, E J. (1985). 'Psyche and the computer'. In Koch, S. and Leary, D. E, eds. *A Century of Psychology as Science*. New York: McGraw Hill, pp.437-451.

Danchin, A. (1980). 'Comment peut-on parler de l'automate cérébrale aujourd'hui?' *Revue philosophique*, 105, pp287-304.

Dehn, N., and Schank, R. (1982). 'Artificial and human intelligence'. In Sternberg, R.J., ed. *Handbook of Human Intelligence*. Cambridge: Cambridge Univ. Press, pp.352-391.

Dennett, D. C. (1991). *Consciousness Explained*. Boston: Little Brown.

Dreyfus, H. (1979). *What Computers Can't Do: A Critique of Artificial Intelligence*. 2nd. ed., San Francisco: Freeman.

Dreyfus, H. (1990). 'Socratic and Platonic sources of cognitivism'. In Smith, J. C., ed., (1990). *Historical Foundations of Cognitive Science*. Dodrecht: Kluwer, pp.1-17.

Edelman, G. M. (1987). *Neural Darwinism*. New York: Basic Books.

Fahlmann, S. E. (1988). 'Parallel processing in artificial intelligence'. In Kowalik, J, S., ed., *Parallel Computation and Computers for Artificial Intelligence*. Boston: Kluwer, pp.3-19.

Feigenbaum, E. A., and Feldman, J., eds. (1963). *Computer and Thought*. New York: McGraw Hill.

Feldman, J. A. and Ballard, D. H. (1982). 'Connectionist models and their properties'. *Cognitive Science*, 6, pp.205-254.

Fodor, J. A, (1968). *Psychological Explanation*. New York: Random House.

Fodor, J. A. (1975). *The Language of Thought*. Cambridge, Mass: Harvard University Press.

Fodor, J. A. (1981). 'The mind-body problem'. *Scientific American*, 244 (4), pp.124-132.

Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge, Mass: MIT Press.

Fodor, J. A., Bever, T. G., and Garren, M. F. (1974). *The Psychology of Language*. New York: McGraw-Hill.

Fodor, J., and Garrett, M. (1966): 'Some reflections on competence and performance'. In Lyons, J. and Wales, R., eds. *Psycholinguistic Papers*. Edinburgh: Edinburgh University Press, pp.135-154.

Fodor, J. A. and Pylyshyn, Z.W. (1988). 'Connectionism and cognitive architecture'. *Cognition*, 28, pp.3-71.

Gardner. H. (1985). *The Mind's New Science: A History of the Cognitive Revolution*. New York: Basic Books.

Garfield, J., ed. (1987). *Modularity in Knowledge Representation and Natural Language Processing*. Cambridge, Mass: MIT Press.

Graesser, A., and Black, J., eds. *The Psychology of Questions*. Hillsdale: Erlbaum.

Harnad, S. (1990). 'The symbol grounding problem'. *Physica D*, 42, pp.335-346.

Harnad, S. (1993). 'Foundations of cognition'. Course given at Eötvös University, Budapest, pp.26-30 April, 1993.

Hebb, D. O. (1949). *The Organization of Behavior*. New York: Wiley.

Hinton, G. and Anderson, S., eds. (1981). *Parallel Models of Associative Memory*. Hillsdale, N. J.: Erlbaum.

Hofstadter, D. (1987). 'Cognition, subcognition'. *Le Débat*, 47, pp.26-44.

Hull, C. L. (1943). *Principles of Behavior*. New York: Appleton.

Julesz B. (1984). 'Towards an axiomatic theory of pre-attentive vision'. In Edelman, G.M., Cowan, W. M., and Gall, W. E., eds. *Dynamic Aspects of Neocortical Functions*. New York, Basic Books, pp.585-612.

Kahnemann, D. (2973). *Attention and Effort*. Englewood Cliffs: Prentice Hall.

Kowalik, J. S. ed., (1988). *Parallel Computation and Computers for Artificial Intelligence*. Boston: Kluwer.

Kripke, S. (1972). 'Naming and necessity'. In Davidson, D. and Harman, G., eds. *Semantics and Natural Language*. Dordrecht: Reidel.

Lachmann, R, Lachmann, J. L., and Butterfield, E. C. (1979). *Cognitive Psychology and Information Processing*. Hillsdale, N. J.: Erlbaum.

Lachter, L. and Bever, T. G. (1988). 'The relation between linguistic structure and associative theories of language learning'. *Cognition*, 28, pp.195-247.

Lycan, W. G., ed. *Mind and Cognition. A Reader*. Oxford: Basil Blackwell.

Marr, D. (1952). *Vision*. San Francisco: Freeman.

McClelland, J. (1988). 'Connectionist models and psychological evidence'. *J. of Memory and Language*, 27, pp.107-123.

McClelland, J. L., and Rumelhart, D. E. (1986). *Parallel Distrib-*

uted Processing. Vol. 2. Cambridge, Mass: MIT Press.

McCulloch, W. S., and Pitts, V. (1943). 'A logical calculus of the ideas immanent in nervous activity'. *Bull. Math. Biophys.*, 5, pp.45-133.

Mehler, J. and Dupoux, E. (1987). 'De la psychologie de la science cognitive'. *Le Débat*, 47, pp.65-87.

Miller, G. A., and Chomsky, N. (1963): 'Finitary models of language users'. In Luce, R. D., Bush, R. R., Galanter, E., eds. *Handbook of Mathematical Psychology*. Vol. II. New York: Wiley, pp.419-49.

Neisser, U. (1963). 'The multiplicity of thought'. *British Journal of Psychology*, 54, pp.1-14.

Newell, A. (1990). *Unified Theories of Cognition*. Cambridge, Mass.: Harvard University Press.

Newell, A., and Simon, H. (1972). *Human Problem Solving*. Englewood Cliffs, N.J.: Prentice Hall.

Norman, D. A., ed. (1981). *Perspectives on Cognitive Science*. Hillsdale: Lawrence Erlbaum.

Nyíri, J. C. (1992). *Tradition and Individuality*. Dordrecht: Kluwer.

Osgood, C. E. (1963). 'On understanding and creating sentences'. *Amer. Psychologist*, 18, pp. 735-751.

Pfeifer, R, Schreter, U., Fogelma-Souli, F. and Steels, L, eds. (1989). *Connectionism in Perspective* Amsterdam: North-Holland.

Pinker, S., and Mehler, J., eds. (1988). *Connections and Symbols*. Cambridge, Mass: MIT Press.

Pléh, Cs. (1985). 'Brain language. On a book by Jerry A. Fodor'. *Folia Linguistica*, 19, pp. 539-5487

Pléh, Cs. (1987). 'On formal- and content-based models of story memory'. In: Halasz, L., ed. *Literary Discourse. Aspects of Cognitive and Social Psychological Approaches*. Berlin: de Gruyter, 1987, pp.100-112.

Pléh, Cs. (1992). 'The renaissance of association in contemporary cognitive psychology'. *Acta Univ. Budapestiensis de Lorande Eotvos Nom. Sectio Philosophica*, 24, pp.161-171.

Polanyi, M. (1962). *Personal Knowledge. London*: Routledge & Kegan Paul.

Polanyi, M. (1967). 'Life transcending physics and chemistry'. *Chemical and Engineering News*, 45, August, pp.54-66.

Polanyi, M. (1968). 'Logic and psychology'. *American Psychologist*, 23, pp.27-43.

Polanyi, M. (1969). *Knowing and Being*. Chicago: The University of Chicago Press.

Posner, M. I. (1986). *Chronometric Explorations of Mind*. Oxford: Oxford University Press.

Putnam, H. (1960). 'Minds and machines'. In Hook, S., ed. *Dimensions of Mind*. London: Collier Books, pp.138-164.

Putnam, H. (1988). *Representation and Reality*. Cambridge, Mass: MIT Press.

Pylyshyn, Z. W., (1972). 'Competence and psychological reality'. *American Psychologist*, 27, pp.546-552

Pylyshyn, Z. W., (1984). *Computation and Cognition*. Cambridge, Mass: MIT Press.

Rumelhardt, D.E. (1980). 'Schemata. The building blocks of cognition'. In Spiro, P. Bruce, B. and Brewer, W., eds. *Theoretical Issues in Reading Comprehension*, Hillsdale: Erlbaum.

Rumelhart, D. E., and McClelland, J. L. (1986). *Parallel Distributed Processing*, Vol.I. Cambridge, Mass: MIT Press.

Rumelhart, D. E., and McClelland, J. L. (1987). 'Learning the past tenses of English verbs: Implicit rules or parallel distributed processing?' In MacWhinney, B. ed. *Mechanisms of Language Acquisition*. Hillsdale: Erlbaum.

Ryle, G. (1949). *The Concept of Mind*. London: Hutchinson.

Schank, R., (1982). *Dynamic Memory. A Theory of Reminding in Computers and People*. Cambridge: Cambridge University Press.

Schank, R, and Abelson, R. (1977) *Scripts, Plans, Goals, and Understanding*. Hillsdale: Erlbaum.

Schiller, H.P., (1946). *Der Aufgabe der Psychologie*. Vienna: Segal and Lachman, 1977. Simon, H. (1979}. *Models of Thought*. New Haven: Yale University Press.

Smith, E. E., Shoben, E. J., and Ripps, L. J. (1974). 'Structure and process in semantic memory'. *Psychol. Review*, 81, pp.224-241.

Smith, J. C., ed. (1990). *Historical Foundations of Cognitive Science*. Dodrecht: Kluwer.

Smolensky, P. (1988). 'On the proper treatment of connectionism'. *Behavioral and Brain Sciences*, 11, pp.1-23.

Szentágothai, J. (1975). 'The module-concept in cerebral cortex architecture'. *Brain Research*, 95, pp.475-96.

Újlaki, G. (1992). 'A hallgatólagos s az explicit tudás. Polanyi Mihaly poszt-kritikai filozófiája'. ('Tacit and explicit knowledge. The post-critical philosophy of Michael Polanyi'). In *Polányi Mihaly Filozófiai Írísai* II. Budapest: Atlantisz, pp.276-328.

Wan, W. C. (1970). 'On two hypotheses concerning psycholinguistics', In Hayes, J.: *Cognition and the Development of Language*. New York: Wiley, pp.137-220.

Wertheimer, M. (1922). 'Untersuchungen zur Lehre von der Gestalt'. *Psychologische Forschung*, 1, p.47ff.

POLANYI'S DISTINCTION BETWEEN PURE SCIENCE AND TECHNOLOGY

P. Hammond

1. Introduction

Michael Polanyi's post-critical analysis of the structure of human knowledge has lost nothing of its original freshness and relevance since he formulated it almost half a century ago. The notions of focal and tacit knowledge provide a brilliant insight into the relationship between individual specialist subjects, which are studied separately, and their 'Sitzen im Leben' by which they combine to throw light on the human condition. Moreover the idea of 'commitment with universal intent' stressed by Polanyi joins the individual human explorer to his fellows and to the external environment and links meaning with value.

Polanyi is very much at home in his discussion of physics, chemistry and biology, where he had first-hand research experience. However, his touch is less sure when he comments on technology. He tends to regard technology as a craft rather than as a scientific activity and has no vision of the vast developments in control engineering and information technology, for which the foundations were being laid at the time when he was writing. Although a detailed study of such failures of foresight would have little interest except as a biographical footnote, an important purpose would be served if such a study could show that Polanyi's general epistemology had a much wider application than he himself realised and that it can be applied to modern technological society as well as to the pursuit of pure science in academic institutions, which was his primary interest. Not that he limits knowledge to the knowledge of science in an explicit manner. His own medical background prevented such narrowness and he gives many examples to a wide range of human activity. A particularly vivid one is the discussion of the method by which medical students learn to interpret x-ray photographs[1]. Nevertheless there is an implicit limitation which becomes apparent in his discussion of technology. The object of this paper is to examine Polanyi's distinction between pure science and technology and to suggest that it is in many respects invalid in the context of his theory of knowledge. If this is so, an unnecessary restriction can be removed from his theory and its usefulness.

2. Polanyi's views of the nature of technology

Although technology is not Polanyi's primary interest, he discusses it in some detail. The purpose of that discussion is chiefly to contrast technology with pure science, because it is the latter which he uses as the paradigm of his theory of personal knowledge. Thus he writes that whereas science aims at a deep understanding of nature, technology is concerned with contriving devices and processes which give economic benefits[2]. Technology works by means of inventions which are patentable in contrast to pure science which works through discoveries which are universal and therefore cannot be patented[3]. Although technology may use scientific discoveries, it relies heavily on indefinable knowledge[4]. There are three stages of learning which are well illustrated by experiments with animals. The first is trick-learning, the second sign-learning and the third latent learning. Trick-learning is a contriving, sign-learning an observing, and latent learning an understanding[5]. Technology is akin to trick-learning[6], whereas pure science is a form of latent learning. Since technological success depends on economic advantages, such knowledge lacks the value of scientific knowledge which is independent of external factors. Inventions can be rendered useless and even farcical by changes in the price of materials, whereas discoveries are valuable in themselves because they increase knowledge[7].

The economic constraints on technology require an organisational structure which is controlled by a central agency. Hence technological tasks are essentially subordinate and subject to central direction[8]. Science on the other hand cannot be planned. It works through spontaneous co-ordin- ation among scientists and requires independent action and originality. It follows that the proper organisation of technology is in centres of industrial production and the proper home for scientific work is provided by universities[9]. In education also there is a clear distinction between pure and applied science. Technology is generally taught in separate technical institutions[10]. There is a similar clear distinction between scientific and technical journals[11]. Scientists and technologists find it difficult to co-operate. The scientist is irritated by the practical pre-occupation of the technologist, while the latter thinks that scientists are blinkered[12]. All in all, science and technology are completely separate domains[13] .

Occasionally Polanyi softens these sharp distinctions. Electro-technics and aero-dynamics can be cultivated like pure science and the study of materials is a technically justified science[14]. There is an activity intermediate between sci-

ence and technology which can be described as engineering[15]. The theoretical principles of engineering are a branch of applied mathematics. They possess intellectual beauty and are best studied within an academic community. Nevertheless their intrinsic interest is linked to external usefulness, whereas pure mathematics and the natural sciences have an interest which lies wholly in themselves[16]. Besides engineering there are also medical subjects which overlap both science and technology as for example pharmacology. The observation of the effects of a drug is a fact of nature, while the prescription of the drug for producing this effect fulfils a practical purpose. Nevertheless the two aspects of pharmacology are distinct[17].

So far I have been careful to present Polanyi's views of technology without comment. There is, however, one very strange matter which cannot be mentioned without drawing attention to its contradictory nature. Polanyi frequently refers to operational principles which govern technology and are absent in pure science. These principles he regards as being linked to the economic purpose inherent in technology[18]. The curious thing is that in his view such principles cannot be studied scientifically, because contrivances cannot be understood by science[19]. At times his reason seems to be that many contrivances are human artefacts, but he also admits that living beings exhibit operational principles[20]. No reasons are given why operational principles should be excluded from scientific enquiry. If science stops short of studying such principles, it cannot study any systems whether these are techno- logical or biological. It is even doubtful whether language itself can be understood apart from its operational principles[21]. If this is so, the use of the presence and absence of operational principles as a demarcation between science and technology does not seem possible and Polanyi's contention that science is to be valued above technology in the pursuit of knowledge is unconvincing.

3. Polanyi and the freedom of science

Polanyi's published work and his career throughout his long active life show him to have been a man of exceptionally wide interests and sympathies. He had experienced racial persecution and was deeply aware of the malpractices of totalitarian societies. Moreover he could understand the fatal attraction of a utilitarian and positivist philosophy to western liberal intellectuals[22]. In particular he was aware of the danger of Marxism with its limitless moral demands on society coupled to a rejection of moral values in individuals. The chapter on 'Conviviality' in *Personal Knowledge* gives his searching analysis of Marxist ideology with its consequence in the destruction of liberty. He was very concerned with the use made of such terms as 'scientific materialism' and the political overtones in Marxist discussions of scientific matters. During a visit to Russia in 1935, Polanyi was able to observe the working of Marxist scientific policy in that country. He was told that science was in reality the same as technology. The notion of pure science was nothing but a bourgeois perversion[23]. Scientists in Russia were forced to make false declarations of practical usefulness to justify their work[23]. Polanyi saw clearly that the very existence of pure science was in danger from this ideology and that the pursuit of science might be abandoned altogether where this utilitarian doctrine prevailed. He worked tirelessly to defend the freedom of science, which he knew to be linked to all other human freedom. It is very likely that he regarded the strict separation of science and technology as a prerequisite for the preservation of independent scientific research. In this he may have been mistaken. It may well be that science and technology are natural allies because technology no less than science requires liberty to pursue the truth. It is relevant that the collapse of the Soviet Union was largely due to the subordination of technology to politics. In the context of this paper we shall stress the point that Polanyi's distinction between science and technology may be a hindrance rather than a help in the defence of intellectual integrity and that it imposes an unnecessary restriction on the application of his ideas of the nature of human knowledge.

4. Personal knowledge

Let us now enquire whether Polanyi's distinction between science and technology follows from his central investigation into the nature of human knowledge. In a short paper like this it is impossible to do justice to the intricacy and persuasiveness of Polanyi's diagnosis and we shall have to be content to select some of the important strands of his discussions. The first of these is that knowledge requires a learning process guided by commitment. Polanyi deliberately fuses together the objective and subjective aspects of learning and therefore of knowledge. There is a world to be discovered but it needs an explorer to make discoveries and a community of explorers to share and evaluate them. These explorers have to be strongly motivated and committed. First he illustrates the process from observations on animals where the motive is generally the obtaining of food. There he examines infant behaviour as described by Piaget, in which the motive of the child is the establishment of a fixed interpretative frame-work for understanding his environment. As already mentioned, Polanyi describes three levels of learning. First there is trick-learning which involves skill, then sign-learning which involves observation, and thirdly latent learning which leads to a deep understanding of a situation.

This classification of types of learning is valuable but in my view it can become misleading if pressed

too far. There is surely a great deal of overlap between all three types of learning. There is also a danger in the manner in which Polanyi attempts to distinguish between the motives in such learning processes. Trick-learning he associates with purpose, sign-learning he describes as guided by the achievement of a strained attention and latent learning he attributes to a heuristic process of intellectual re-organisation[25]. There seems to be no reason why purpose should be confined to trick-learning especially as Polanyi discourses most eloquently on the passionate commitment required by people involved in scientific discovery. Moreover, human motives and purposes are always mixed and are not totally defined by the task. The desire for promotion or recognition is likely to be present in scientific work alongside the more noble motive to make explicit some of the beauty and order of natural processes. Conversely, a desire for good workmanship is generally present even in very simple tasks which may come under Polanyi's somewhat derogatory description of trick-learning. All kinds of learning involve effort and effort requires a purpose. That consideration is relevant to Polanyi's distinction between science and technology. Science no less than technology requires motivating purposes. When Polanyi describes scientific knowledge as being independent, whereas technology is governed by economic factors, he comes dangerously close to the positivist view of knowledge as an object in the manner of Sir Karl Popper's account of objective knowledge stored in libraries[26]. In general, this is contrary to Polanyi's considered opinion, because he strongly argues that scientific knowledge needs a scientific community to sustain it[27]. If knowledge cannot be detached from human beings, it is not surprising that the pursuit of scientific knowledge is subject to a greater or lesser extent to non-scientific influences. Where such influences are dominant they may require special kinds of organisation, but it seems contrary to Polanyi's central thesis to assert that knowledge is degraded by being useful. Of course pure science needs to be defended against economic planners who are concerned only with monetary return. However, a defence based on an intrinsic absence of useful applicability is likely to be misunderstood. It may even lead the planners to seek to prove their point by cutting off economic support for scientific research. That would be misguided but it would show that economic factors influence science as well as all other human endeavours. It is also worth remembering the story of Michael Faraday who, when asked by Mr Gladstone what use electricity was, is said to have replied, 'You will soon be able to put a tax on it'[28].

5. Focal and tacit knowledge

One of Polanyi's most valuable insights into the nature of knowledge is that personal knowledge is necessarily of two kinds. As an illustration he uses the act of driving a nail by hitting it with a hammer[29]. To perform this action skilfully one has to attend to both nail and hammer, but in a different way. Attention has to be focused on the nail, but the way in which the hammer is grasped by the hand and moved by hand and arm is also part of the knowledge required. However, this know-ledge is subsidiary and tacit. Another telling example is the use made by a blind man of his stick, which he uses as a probe to explore his surroundings[30]. Focal and tacit knowledge co-exist and cannot be separated, nor is there any sharp demarcation between them. However, they are also mutually exclusive in the sense that an examination of the tacit component shifts the focus and involves new tacit components. It is impossible to have focal knowledge without its tacit background. This destroys the possibility of isolated objective knowledge and severely limits the value of critical methods of analysis such as the method of deconstruction.

The example of the hammer and nail is clearly a technological one, whereas the activity of exploring by means of a probe is more akin to science. The blind man with his stick can be taken as a type of scientist, if the stick is replaced by a particle accelerator or a computer, or a mathematical hypothesis. In terms of focal and tacit knowledge there is a complete identity between the technological and the scientific activity. Polanyi tries to rescue the distinction by saying that hammering a nail produces a material change and is an achievement rather than knowledge. He writes that such achievements are characterised by success or failure, whereas knowledge is true or false[31]. This does not sound convincing. The blind man's successful negotiation of a street corner combines achievement with knowledge and so do all technological and scientific investigations and actions. Indeed, knowledge always involves successful action and is itself an achievement. Conversely, technological failure, such as the collapse of a bridge, is always associated with lack of knowledge.

6. The nature of operational principles

Polanyi's distinction between technological achievement and scientific knowledge is carried further in his discussion of operational principles which in his view are unrecognisable by science. Such principles he attributes to machines as against parts of machines[32]. The measurement of time by a clock is an operational principle. It resides in the clock as a whole and cannot be understood by considering its parts. This is undoubtedly a valuable insight. To tell the time we need focal knowledge of the clock as a whole, the parts provide a tacit field of knowledge. If we study a part like the minute-hand focally, we can no longer tell the time.

However, once the clock has been observed focally, it is useful for clock-makers and other specialists, such as antique dealers, to shift the focus, so that time-keeping becomes subsidiary and tacit. Undoubtedly the clock as a whole is an achievement, but so are its parts. For example, the case of a grandfather clock is a considerable achievement and can itself be considered to possess an operational principle. So do all the other parts. In terms of knowledge, Polanyi's distinction between the whole and its parts has little explanatory power.

However, he does have an important point in drawing attention to the relationship between the operational principle of the clock and the complexity of its parts. A clock measures time only when its parts combine and its operational principle does not reside in the parts by themselves nor in a collection of parts. There has to be an interconnection of the parts which can be described as complexity. It should be noted that a part without such interconnection is strictly speaking not a part but a whole. As such it has its own operational principles and can again be examined as a system of parts. Surely this is the reason why the so-called fundamental particles of physics can always be sub-divided.

This thought leads to a considerable extension of Polanyi's ideas. He showed that the analytical procedure of examining the parts of an object possessing an operational principle cannot define or explain the object as a whole. In technological terms the better description of such an object is as an interconnected system. What Polanyi apparently did not realise is that all objects are themselves systems and that this applies even to the fundamental entities used in pure science. The boundaries of such systems are similar to the boundaries between focal and tacit knowledge. Strictly speaking there are no isolated objects or systems, although they can be examined focally. Polanyi's separation of technology from science seems to imply that science deals with entities which cannot be sub-divided whereas technology deals with assemblies. The processes of technology are indeed assembly processes but so are many natural processes and particularly those described in terms of evolution.

Nor does the criterion of the patentability of operational principles provide a clear demarcation between science and technology. If it did so there would be no role for patent lawyers. The trouble is that operational principles are always embedded in general physical laws which cannot be patented. Patents afford essential protection in the pharmaceutical industry but are not very important in engineering. The reason is that chemical substances can be defined by formulae, but engineering processes are difficult to define. It is fairly easy to circumvent a patent description of an engineering device and competing industrial firms often hold similar patents. Litigation to establish priority is relatively rare. The use of the patents is chiefly in marketing a product. Polanyi's extensive use of patentability in his discussion of the difference between technology and science is not borne out by industrial practice.

7. Emergence and complexity

Polanyi tends to regard his discussion of learning processes into trick-learning, sign-learning and latent learning as stages of development, but he is aware that this presents difficulties[33]. In particular there is a difficulty in regarding technology as trick-learning because it already makes use of scientific knowledge, which is acquired by latent learning. There is the further difficulty that operational principles depend on complexity and yet in his view they are in some way inferior to scientific principles or laws. In the final chapter of *Personal Knowledge*, Polanyi deals with the emergence of machine-like operations in living things[34]. The term 'emergence' evokes a temporal process and is closely connected with the idea of evolutionary development. It is very difficult to reconcile evolutionary progress with a theory which regards machine-like operations as a preliminary and somewhat unsatisfactory form of trick-learning. However, this does not necessarily lead to the opposite conclusion that technical 'know-how' is superior to scientific generalities, nor that technical development necessarily comes later in time. The immense cathedrals of the middle ages were successfully constructed without a knowledge of structural analysis. On the other hand, modern tele-communication systems could not have been devised without the scientific knowledge of electrical phenomena. It seems right to conclude that technology and pure science are by no means 'totally disparate domains', but, on the contrary, that they are closely related in an intricate and complex manner.

8. Knowledge as an achievement

A recurrent theme in Polanyi's analysis is that knowledge always entails activity. Such activity involves personal commitment which is to be shared in community and which therefore has both objective and subjective aspects. Moreover, there is the possibility that the commitment may be mistaken. There is an inevitable risk in making a commitment. Knowledge is both either true or false and also successful or unsuccessful. Knowledge, as a successful activity, is an achievement.

How is success to be measured? Scientific knowledge will be tested experimentally and will also judged by its inherent beauty and its fruitfulness in suggesting further research. In mathematics the tests will be consistency, elegance and intrinsic interest.

What about technology? Polanyi is uneasy about technology because he distrusts economic tests. Scientific discoveries have permanent significance, but technological inventions become obsolescent. Steam engines are replaced by electric motors. Science is pursued for what it is in itself: technology is pursued for gain. Clearly Polanyi is reacting against an overbearing utilitarianism and many academics will have sympathy with his view. Nevertheless in terms of the theory of knowledge as achievement this rejection of economic motive is unnecessarily restrictive. There is more to technology than monetary advantage, although such advantage often provides a useful test of successful achievement. Monetary advantage is an aspect of a more general test in terms of 'added value'. That will involve 'fitness for purpose' and will include such features as use of natural resources, social effects, environmental effects and safety. Academics need to be reminded that the provision of research facilities is a technological achievement, just as politicians need to remember that scientific work deserves support even if it confers no immediate economic benefit.

9. Conclusion

Polanyi's theory of personal knowledge is an enormous achievement. It sets us free from restrictive critical views which seek to equate knowledge with sense data and which turn their backs on human imaginative powers. By stressing commitment and achievement Polanyi shows the uselessness of a corrosive analysis which replaces whole persons by collection of parts and in so doing destroys both meaning and purpose. By reminding us that our focused vision is always surrounded by a tacit field, Polanyi enables us to accept tradition and use it as a means for discovery. It is a pity that, in his struggle against the implications of Marxism, Polanyi sought to separate scientific knowledge totally from technology. Such a separation is not an integral part of his theory. When this is realised, the power and scope of his views become more clearly visible.

Department of Electrical Engineering
The University of Southampton

The author would like to thank Dr R.T. Allen for letting him use his annotated bibliography of Polanyi's published work on society, economics and philosophy.

Notes:

Where only page numbers are given these refer to *Personal Knowledge*.

1. p.101.
2. p.178.
3. p.177.
4. p.52.
5. pp.71-77.
6. p.76.
7. pp.177-178.
8. 'Pure and applied science and their appropriate forms of organisation', Occasional Pamphlet No.14, Society for Freedom in Science, 1953.
9. 'Science: Academic and Industrial', *J. of Institute of Metals*, Vol 89, 1961, pp. 401-6.
10. p.180.
11. p.180.
12. p.178.
13. p.179.
14. p.179.
15. See n.9.
16. pp.186 and 178.
17. p.179.
18. p.176.
19. p.175.
20. p.383.
21. pp.77-82.
22. pp.227-233.
23. p.238.
24. pp.180-181.
25. pp.71-77.
26. K. R. Popper, *Objective Knowledge* (Oxford, Clarendon Press, 1972).
27. p.182.
28. J. Kendall, *Michael Faraday* (London, Faber and Faber 1954).
29. p.55.
30. pp.55-56.
31. p.175.
32. pp.328-331.
33. pp.174-175.
34. pp.401-402.

CIVIL SOCIETY IN MICHAEL POLANYI'S THOUGHT

Endre J. Nagy

1. Introduction: civil society in Eastern Europe

Why should we deal with Polanyi's views about civil society if Polanyi never used this term? As is amply pointed out in the literature, Michael Polanyi was not a 'professional' philosopher and 'gave little detailed attention to other philosophers writings' (Allen, 1990, p.15). He worked out a vocabulary of his own, beginning in the thirties when struggling against totalitarianism, and he developed a way of looking at things which organises the position and the place value of every particle of his theoretical edifice. The 'Polanyian flavour' means that even the categories and concepts he had taken over from the traditional language of philosophy are through and through imbued with his idiosyncratic perspective. They always need interpretation. Yet, it is also acknowledged that any themes he entertained deserve attention, for his arguments and conceptions are often very instructive to anybody attacking a particular problem. We may presume that this holds too for the topic of civil society.

In speaking of civil society in this paper I intend to refer to citizens freely forming alliances, institutions and associations designed to further private ends and social interests. I shall focus upon the recent history of the notion of civil society, especially as it applies to Eastern Europe.

The idea of civil society was taken up by Polish 'dissident' intellectuals (Michnik, Kuron, etc.) who were followed swiftly by Hungarians (Bence, Kis) in the late seventies (see Frentzel-Zagorska, 1990; Arato, 1990). Civil society as a topic was a magnet for those with political aspirations and engaged in discussions of human rights, rights of association, alternative media and free public inquiry, but not seizing political power. Far from it. The representatives of opposition intellectuals gave the Communist leadership the greatest surprise by their taking seriously the 'socialist' Constitution that, imitating the institutional framework of Western parliamentary democracy on the 'surface', seemingly guaranteed human rights. However, these rights were confiscated in the details of the jungle of law, or else, to give a striking example, in the measures taken by a policeman. Thus the conception of civil society signalled a 'velvet' struggle against the Communist regime, inasmuch as it only sought to eliminate its 'anomalies' and 'dysfunctions'. As the opponents of the régime, no more than anybody else at that time, could not expect the dissolution of the Soviet empire, the suggestions were subsequently approved.

If one opens a book at random about the changes in Eastern European since 1989, one will find an interpretation of why civil society became a battle-cry heralding the end of Communism. B. Ackermann locates the problem of civil society within the context of Marxist teachings about bringing up the 'socialist type of man' who would sacrifice his private interest on the altar of the general interest of the whole of society. He then argues as follows:

> Eastern Europeans have developed a rich literature of 'civil society', emphasising the crucial need to construct institutional spaces that might serve as a buffer against the totalitarian pretension of the state. The aim here is to describe a more modest kind of political involvement in which citizens can control the state without merging their identities into the collective whole (Ackermann, 1992, p. 33).

As an example, leading representatives of the Hungarian opposition suggested to the Party leaders, in a pamphlet, that they should divide the political power so that the Communist Party would maintain its sovereignty over a part of foreign policy (mainly bearing upon affairs relating to the Warsaw Pact and Comecon) while letting the rest of the issues be handled by society (see, Kis, Köszeg, Solt, 1987, p. 10). This proposal was put forward in the second half of the eighties and, as could not have been imagined earlier, would probably have led to a change of the political system in the long run if the Party élite had accepted it. As such this conception really transcended the role of civil society as being the 'buffer' ascribed to it by Ackermann.

* * *

Yet, an older conception of civil society existed in Eastern Europe alongside the tempered political aim as outlined above. This was its economic side. It is known that Hegel and Marx spoke in their theories, not about civil but about bourgeois society, *bügerliche Gesellschaft*, and Marx identified it with the sphere of capitalist economy. Drawing on the possibility inherent in the French language, he, firstly opposed, in his philosophical period, the selfish bourgeois to the altruistic *citoyen* and wanted the latter to prevail over the former in a true Rousseau-like, direct democracy or else in a democracy as found in the Greek

polis. Later on, in his *Grundrisse* of the early sixties, he despised even the citoyen, and tried to point out that the endeavours of the citoyen to support equality, liberty, and fraternity were but manifestations of the concealed aims of the selfish bourgeois for the rule of society. Since the existence of the bourgeois—and capitalism itself—was based on private property, the latter must be eliminated by a workers' revolution.

It was then but a logical outcome of civil society movements that when, in the late eighties the change of system was put on the agenda in Eastern Europe, the other side of civil society, the economic one with the requirement of the rehabilitation of private property and the market economy, became subsumed under the concept of civil society. An intimate observer of Eastern Europe, Timothy Garton Ash, saw this conjuncture of the two sides of civil society:

> For what most of the opposition movements throughout East Central Europe and a large part of 'the people' supporting them were in effect saying was: Yes, things are intimately connected—and Marx is right, the two things are intimately connected—and we want both! Civil rights and property rights, economic freedom and political freedom, financial independence and intellectual independence, each supports the other. So, yes, we want to be citizens, but we also want to be middle-class, in the sense that the majority of citizens in the more fortunate half of Europe are middle-class. We want to be *Bürger* AND *bürgerlich*! Tom Paine but also Thomas Mann. (Ash, 1990, pp. 148-149).

* * *

This view was not without antecedents. Sociologist noticed already in the seventies, especially in Hungary, an 'economic civil society' in the making. There had been a 'shadow economy' beside the state controlled one all along during the history of Socialism. It was not a 'black market economy' in its entirety. The most important form of the legal and partly private economy was small-holding. The land was owned by the *kolkhoz* but privately used by families. Small-holding closely connected with the *kolkltoz* provided the family with seed-corn, animals for breeding, crops for selling, etc. This symbiosis of two kinds of economy has remained the most successful economic branch ever since in Hungary, whereas other East European countries had often to cope with food shortages. In the mid-eighties a third of the production in certain agrarian sectors was yielded by small-acreage farmers, while the entire agriculture of Hungary became one of the most export-oriented branches of the national economy. Furthermore, the small farmers were followed by private industrial associations, economic co-operatives, factory co-operatives, small co-operatives, etc. They were joined by a tiny fraction of craftsmen and tradesmen who have always existed in Hungary. Essentially, these economic activities were described as a 'second economy'.

It was I. Szelényi—a sociologist expatriated in the mid-seventies, now a professor at UCLA—who named the whole process just outlined, 'embourgeoisement'. He saw it as the formation of the new middle class that T. G. Ash was talking about in the above quotation. However, Szelényi had precursors who already recognised the rise of the new middle class (Zsille, Juhász, Lengyel, Magyar). Yet, they also described it as being 'distorted', 'incomplete', 'one-sided' under the given political circumstances. As I. Kemény articulated clearly in the title of his essay (which gives an outstanding summary of the relevant literature; see Kemény, 1991, pp. 131-146), the process involved was 'embourgeoisement without civil rights'. But Szelényi mostly emphasised the economic side. He took over the theory of P. Juhász about the so-called interrupted embourgeoisement (the claim that it was stopped by the Communists at the end of the forties) and tested it empirically. While he corroborated the emergence of new middle classes, he transposed the results onto another, rather 'ideological', level by stating that it originated a 'Third Road' between Capitalism and Socialism, a theory which was in fruition among Hungarian 'populist' writers before World War II. While most of the sociologists celebrated the process because it produced economic actors as entrepreneurs independent from the state, Szelényi considered them, and named them, 'socialist entrepreneurs' (Szelényi, I. 1987). Socialist embourgeoisement, he argued, was a successful 'silent revolution from below', since

> classes struggle to achieve compromises to alter the distribution of power between classes at the point of production and by establishing an alternative economic system.

He added that 'in Eastern Europe it is the public-political sphere where the least action occurs!' (Szelényi, 1987, p. 8). Thus Szelényi hoped that Eastern European societies would not join the Western type of market economy but, in the long run, transform themselves into a new social-economic formation which would be neither Capitalist or Socialist After Communism collapsed he adjusted his tenet so that it was this revolution of small people, peasants and workers that undermined the régime, since they did not make a frontal assault on the castle of the state but 'came around by the back door'.

Yet, there was less talk of the problematic side of embourgeoisement, and Szelényi himself hinted at the time that the bourgeois needs to become not only an entrepreneur but also a citoyen. In a broadcast interview Szelényi wittily stated that 'a bourgeois was one whose grandfather had already been also bourgeois'. However, the moral and political deficiencies of embourgeoisement matter. Indeed, there were disquieting harbingers of

problems. An American ethnographer of Hungarian extraction, M. Sozan, outlined the numerous kinds of thefts wide-spread among members of co-operatives and concluded his empirical research by making the final statement:

> Everyone knows the system well. Accordingly, it is not the fact that here everything reposes upon stealing that demoralises the members but the fear that they might be unable to steal enough (Sozan, 1985, cited by Kemény, 1991, p. 108).

But one could cite other kinds of 'negative reciprocity' by referring to the American social-anthropol ogist, M.D. Sahlins, who, following A. Gouldner, carried out the topology of the exchange relations of goods prevailing within primitive tribes (Gouldner, 1960, Sahlins, 1965). According to Sahlins negative reciprocity occurs, as against generalised and balanced ones, when partners in the exchange of a commodity or good act immorally. It goes without saying that the economic mentality of negative reciprocity has not been confined only to the members of co-operatives, but is found inside circles of society. Of course, there were exceptions, for the author of this essay also found sporadically in his empirical researches entrepreneurs who went beyond their direct material interests and established themselves as being both bourgeois and citoyen. Yet, this does not call into question the validity of the statement that the whole of society was informed by the mentality of negative reciprocity.

This problem was a crucial point which, in my opinion, theoreticians and sociologists dealing with civil society and embouge-oisement did not take into account. Nor can one find any serious tackling of this problem by glancing at the huge literature of the political sciences. Political scientists laid stress always upon the so-called 'interest articulation' of various social bodies, whether already existing, like 'official' unions, professional chambers, or those to be established. For existing bodies had always to yield to the Party-line. The keen advocates of civil society wanted these bodies to become independent from the Party by letting them articulate their interests. But the question, 'How then could these bodies of civil society be prevented from falling into local, professional parochialism?', had never been raised. To sum up, there were two points missed in these theoretical and empirical assumptions. The first one bears on the economic side of civil society, the process of embourgeoisement. The economic mentality of negative reciprocity must not be conflated with Max Weber's 'spirit' of modern capitalism. Weber described 'adventure capitalisms' (like trade capitalism, war capitalism, etc.) that served as obstacles to the formation of modern capitalism. To characterise the spirit of these pre-modern capitalisms, Weber recalled as a symbol the Dutch captain who 'was ready to travel for the gain down into the underworld, even if his sails would get burnt', a mentality, he adds, that is well-known up till now to anybody getting acquainted with coachmen and boatmen in Naples, Italy, not to mention similar types in Asia. The unlimited (e.g., morally not moderated) 'drive for gains' and 'desire for acquisitions' has in itself nothing to do with the spirit of modern capitalism that had to defeat them for its coming into being at the dawn of European modernity. It was the Protestant ethic that helped the emerging modern capitalism to win over the unlimited 'desire for money acquisition'. However, once capitalism was firmly rooted at the onset of the nineteenth century, it needed not this religious basis any more, since thereafter fair and decent economic behaviour was obtained by the force of the impersonal economic mechanism. But, henceforth, can the overall negative reciprocity, the innumerable manifestations of which are being experienced both by tourists and businessmen getting in touch with all sorts of economic conditions and actors throughout Eastern Europe, be eliminated at all if there is no religious support to economic activity under the conditions of a secularised world?

* * *

The second point missed by sociologists and political scientists closely relates to the first one: the miserable state of morality in general in Eastern Europe. Undoubtedly, there were authors who emphasised the part the moral component played in the revolt against the communist rule. However, even those who took notice of it, gave more attention to the social and economic background. For example, Jadwiga Staniszkis, a former adviser of Solidarity in Poland, did not overlook the moral component and put it this way:

> It was a moment when individuals, brought up in the totalitarian situation (characterised by a gap between common-sense morality and rules imposed in the name of 'objective reason') were able to overcome their own moral indifference to society and to themselves. The moral experience of Solidarity strikes was also a peculiar cognitive experience. People discovered that they were ready to run risks in the name of values and that others would act in similar way (Staniszkis, 1991, p. 237).

In a similar vein, B. Ackermann recalls Vaclav Havel's call to 'live in truth' (see Havel, 1988). In Ackermann's interpretation it 'evokes the best of the liberal spirit in opposition to the oppressive banalities of bureaucratic totalitarianism' (Ackermann, (1992, p.32). He also hinted at the term 'anti-politics' introduced by the Hungarian writer, György Konrád, who, by this term, promulgated an 'ethical revolution' politics as such. He scornfully looked down upon politicians of East and West equally and, addressing himself to individuals, called them to change their personal lives, from the most intimate sphere to the realm of environmental protection. Thus, intel-

lectuals in Eastern Europe proclaimed a moral revolution since, at that time, the Soviet Empire did not seem ready to be overthrown in the foreseeable future. And at the 'moment of truth', 1989, when Communist régimes crumbled at astonishing speed, such moral factors like longing after truth were causally as instrumental in the elimination of the system as they were in Hungary in 1956. Remember 'The Message of the Hungarian Revolution' and 'Beyond Nihilism' by Michael Polanyi, who quoted Hungarian and other East European writers and poets who had decided to form 'a firm alliance for the dissemination of the truth' (Polanyi, M., 1969, p. 20). Coincidences between the two events, 1989 and !956, are shiningly clear. It has always been important to reassert moral values like truth, justice and solidarity. And ordinary people understood the teaching of intellectuals and revolted against the structure of organised lying, against the 'devastated moral environment' (V. Havel).

However, the fusion between intellectuals and the masses could not endure for ever. That correlates to the very nature of any kind of 'moral revolution'. The change of system in Eastern Europe was the work of intellectuals in the last analysis (see Ash, 1990, pp. 135-6). There were playwrights, historians, actors, philosophers and economists who initiated, promoted, proclaimed the 'moral revolution', stirred up the masses and settled down to the peaceful taking over of power ('round tables' in Poland and Hungary). But at the onset it was all 'anti-political', since there seemed not to be any real force that one could expect to begin the fight with a real hope of success.

This anti-political, moral attitude has been characteristic of intellectuals throughout history, as Max Weber demonstrated in his *Religionssoziologie*. Weber argues that intellectuals seek to fill the chaotic universe with 'meaning', and when this endeavour ought, of necessity, to face the social and economic order prevailing in the real world, a desire to 'escape the world' overcomes intellectuals. Intellectuals can escape into 'unspoiled nature' (Rousseau), promulgate 'moral self-perfection' (Tolstoy); or join the 'people' untouched by distorting human traditions (the Russian Narodniki), and this escapism can take on an overtly religious shape, aim for individual redemption, or else, aspire to change the world, not with social and political revolution, but in a collective ethical-revolutionary way (see Weber, 1964, pp. 381-392).

Thus, the intellectuals of Eastern Europe who proclaimed the moral revolution joined a particular style of escaping the world characteristic of certain intellectuals under certain historical conditions. Yet, this assumption must be qualified. It is true that the intellectuals involved promulgated a 'collective-ethical revolution' in Weberian terms, but, while they repudiated the whole of the oppressive state and totalitarian power, they did not reject the public sphere or 'society' as such. And it was just this moment of their 'ideology' that led them to proclaim a free civil society. It follows that the slogan of 'anti-politics' was substantially but a postponed politics, an oblique politicisation opposed to totalitarian rule. As J. Staniszkis plainly puts it, civil society movements were 'to take over the function of the state' and it turned out to have been 'an effective instrument of political struggle' (Staniszkis, 1991, p. 182).

* * *

However, moral ideals have been lost since 1989. Intellectuals who acknowledged the moral requirements to be instrumental in the collapse of the Communist régime and who, thereby, took a position of moral authority almost equal to that which the Church had in the Middle Ages, began to dispense with moral considerations, and, instead, to speak more and more in terms of mere economics, party politics and articulation of material interests. For example, J. Staniszkis, cited often above, finally reduces the moral component to a 'neo-traditionalism' that added to anti-secularisation, anti-individualism and, finally, anti-capitalism (see Staniszkis, 1991, p. 236). She disregards the fact that the moral co-efficient, according to Max Weber, was a pre-condition for the rise of a working market economy. In a similar vein, B. Ackermann restricts the scope of the validity of Havel's call to 'live in truth' by stating that

> the aim of the liberal revolution is not collective truth but individual freedom—freedom for each person to assert his or her moral ideals, even if a neighbour considers them 'wrong' (Ackermann, 1992, p. 32)

as if moral ideals such as truth were a private affair without intending universal validity.

Giving up moral considerations has been one of the strains characteristic of intellectuals since 1989. Another one is that intellectuals who gained the status of a secularised clerisy during the revolution have stripped themselves from the charisma of a moral authority by getting entangled with everyday party politics. Moreover, another, perhaps larger, segment of them have become disgusted with political altercations (which, by the way, unavoidably belong to any parliamentary democracy). As a result, ordinary people who fused with intellectuals and their moral ideals in the magnificent heyday of the revolution, were, by now, left alone. They had to face predicaments to which they were not at all prepared (impoverishment, unemployment, etc.).

And now, it also turns out that these people could not get rid of the backwardness of the political culture characteristic of the Eastern European region before the Communists seized the power at the end of the '40s. István Bibó, minister in the government of Imre Nagy in 1956, the last man to leave the Parliament building on the 4th of

November when it was already occupied by the Red Army, and whose sociological and political writings influenced public thought the most in the late '70s and '80s in Hungary, uncovered the 'inherently backward political culture' in this region of Central and Eastern Europe. We cannot digress in detail on the distresses, as Bibó called them, which tormented these peoples, but some of them can be mentioned which clearly show the deep historical soil on which the peoples of Central and Eastern Europe are acting now that they have regained their freedom. According to Bibó, the backward political culture manifested itself in 'confused, vague and false political philosophies', 'anti-democratic nationalism' and 'deformation of political character'. Bibó sums up concisely the consequences as follows:

> The distorting psychological symptoms of the inability to keep a healthy balance between desires and reality are clearly evident in the self-contradictory behaviour of the peoples of this region: exaggerated self-documentation and inner insecurity, oversized national vanity and sudden self-humiliation, the constant mentioning of achievements and the obvious decline in the real value of achievements, moral demands and moral irresponsibility . . . Under such conditions, the sense of political values is pushed into background (Bibó, 1991, pp; 45-46).

Could we believe that all these deformations of political culture vanished without trace just in the years of Communist rule? It can hardly be so. Under changed historical conditions they would have survived in the souls, minds, attitudes, mentality of the population involved. In 'Beyond Nihilism' and 'The Message of the Hungarian Revolution' Polanyi listed some possible dangers threatening the future. Polanyi was afraid of religious bigotry and especially of national feeling that 'has proved in the past no safeguard against the descent of dynamism into moral inversion' (*KB*, p. 22). Furthermore, he also reminded us in *Personal Knowledge* of the possible perils of distortion of 'civic thought' by insisting that the institutions of loyalty, property and authority can be distorted into local parochialisms (say nationalism), greedy appetite and sheer violence, respectively, since they rely 'ultimately on coefficients that are essentially at variance with the universal intent of intellectual and moral standards' (*PK*, p. 215). And having taken a look over the mental panorama of Eastern European countries, one can sense the foreboding presence of factors that Polanyi guessed and contemporary observers (see Ash, 1990, pp. 143-149; Staniszkis, 1991, pp. 236-7) took notice of.

* * *

Thus, having given a brief account of the actual state of the morality of the masses in Eastern Europe and having taken into consideration the attitude of intellectuals to it, we can rightly state with Polanyi that 'all the logical antecedents of inversion are present today as they were before' (*PK*, p. 22). That is why we reconstruct Polanyi's view of civil society even though he never used this term. For, in contrast to the prevailing conviction according to which free society can repose in the free articulation of interests and observation of the rules of the free market, he always and frequently emphasised that a free society must be guided by 'transcendent values' and dedicated to 'moral and intellectual order', and, thereby, the market can also accomplish its practical task and 'provide a framework for making a living' (Polanyi, 1945/III, p. 1).

In the following sections of this essay we will analyse the development of Polanyi's social theory from the '30s up to *Personal Knowledge,* focusing on the elements which relate to the web of civil society. Because elsewhere (though in Hungarian, see Nagy, 1991) we outlined the historical paradigms of civil society, it will be sufficient to present the paradigm that we use in our investigation. This is the Hegelian construction of civil society. As an outstanding German interpreter points out, Hegel

> separates the political sphere of the state from the realm of 'society' which has become 'civil'. In this way the expression 'civil' gains a primarily 'social' content as opposed to its original meaning and it is no longer taken to be synonymous with 'political' as it was in the eighteenth century' (Riedel, 1984, p. 139).

Civil society is structured for it encompasses 'the system of needs' (identical with political economy, viz. 'pure' economics), the corporations (professional chambers, guilds), jurisdiction by virtue of which the *Bürger* make decisions in suits among them, the police who watch over the public order. Hegel considers the whole of civil society as a mediating sphere between the family and the state, for individuals develop step by step from their selfish particular interests to an insight into the general interest that is presented by the bureaucracy, the representatives and the monarch. However much Marx criticised Hegel's conception and historical conditions have changed since the death of Hegel, we are still living according to this paradigm. Even totalitarianism did not add to the conception of civil society, since it completely did away with it. At most, one can state that totalitarianism extended the border-lines of civil society for it prohibited the citizens from organising even the most 'innocent' union, e.g. a union of philatelists, if it did not want to integrate its activities into the official framework of institutions. Thereby in totalitarianism each small circle could eventually take on a 'public' character when it became a hot-bed—as Polanyi would say—of 'independent thought'.

In what follows we shall confine ourselves to the reconstruction of Polanyi's concepts relating to prob-

lems which arose and still are arising in civil society. We have made an effort to concentrate on the unpublished manuscripts for they can throw light on the words published and already known.

2. Polanyi's anti-total- itarian starting-point

As is well documented in literature and also repeatedly discussed by himself, M. Polanyi became concerned with the miserable plight of government-run science in the Soviet Union and Germany (*TD*, p. 3), and that caused him to fight against it with intellectual means. Reflecting on the experiences gained on the spot in the Soviet Union, he recognised a close connection between the corruption and backwardness of science and the peculiar type of society functioning there. So he drew the consequence that the denial of the very existence of independent scientific thought follows inevitably from a philosophy he called in the context 'Social Absolutism', and this closely correlates with the organising principles put into effect in that society. On the one hand, Social Absolutism endows itself with the right to represent society's interest and to judge good and evil, truth and falsehood, assuming intelligent control of all affairs including the. ideas of the citizens (Polanyi, 1941/ V p.15). On the other hand, Social Absolutism conceives central planning as the comprehensive organising principle to be practised throughout society, including science. Planned society coupled with planned science was such a mortal danger for independent science and free society that Polanyi risked his entire scientific career to fight it.

This anti-totalitarian motive provided Polanyi with a starting point from which he approached social issues (cf. Prosch, 1986, pp. 176-199; Gelwick, 1976, pp. 35-41). This anti-totalitarian starting point can be said to be his 'Archimedean immovable fixed position' from which he was going 'to jolt the world' by constructing a new theoretical space that he equipped with proper notions having suitable positions and specific weight. Consequently, we have to suppose civil society to fit into this theoretical space.

Polanyi's leading ideas were centred around science which provided him with a vista constituting, thereby, a field of elements uniquely characterising his conceptual framework regarding totalitarianism. Being concerned about independent science he inferred that free thought was an inescapable presupposition for its existence. Looking at things from this vista he was actually forced to conclude that other intellectual spheres cultivated by society like art, religion, law, etc., also required free thought that, in turn, presupposes free society, too. Polanyi equated the issue of the freedom of science with the destiny of Western civilisation (cf. Prosch, 1986, pp. 176). For he was seeking 'to restore and reconsolidate the essential elements of a civilisation which is in the process of breaking down' and 'to reconsider the very purpose of society' (Polanyi, 1945/1, p. 1). This consists of recognising 'transcendental obligations', e.g. 'moral and intellectual purposes' (Polanyi, 1945/III, p. 1). Any society fostering, defending and disseminating free thought in various shapes can be called a 'dedicated society'. And at this point civil society, as explained in the introduction in the terms of post-Marxist thinkers of the '70s and '80s, appears in Polanyi's thought. Civil society is an aggregate of smaller and larger circles, groups of individuals who independently govern themselves for their own sakes under the rule of law. We will realise that Polanyi elucidated these circles as 'effectively fostering the intrinsic power of thought' and that is why they are not tolerated in totalitarian régimes.

> They are feared more than any scientific associations, because the truth of literature and poetry, of history and political thought, of philosophy, morality and legal principles, is more vital than the truth of science. This is why the independent cultivation of such truth has proved an intolerable menace to modern tyranny (*TD*, p. 84).

We have to stop at this stage of our discussion. We have anticipated to some extent the outcome to which we are going to come at the end of our argumentation. It has proceeded without proof, on trust that the reader will find our arguments sound. The final results of our discussion of Polanyi's conception of civil society must subsequently corroborate all that we outlined so far if it is to be successful.

3. The logical time-structure of Polanyi's thought

We consider *Personal Knowledge* to be the culminating fruition of Michael Polanyi's scientific career. All trends of thought emerging from the second half of the thirties flow into *Personal Knowledge* and occupy their proper place in its conceptual space. We would argue that all notions uniquely characteristic of Polanyi will not undergo any considerable change in specific weight or in significance hereafter. What will happen to them in *The Tacit Dimension* (1964), in the essays of *Knowing and Being* (1969) or in *Meaning* (1975) are a deepening or an expansion into new domains. For instance, the term 'from-to knowledge' appears in *The Tacit Dimension* (*TD*, pp. 3-25) but anybody who would try to deny the presence of equivalent thought in *Personal Knowledge* (*PK*, pp. 55-58) would be wrong. We suppose the same is true of civil society in Polanyi's thought Its most mature structure can be found in *Personal Knowledge*. However, he omitted certain elements from the final composition of the theory of civil society in *Personal Knowledge* because, partly, he dealt with some of them in *The Logic of Liberty*, and partly, the topic in question had to be submitted to the requirements of the internal logic of

the work. Yet parts omitted have to be integrated in the interpretation which seeks to focus on the nature and meaning of a specific concept like civil society. For this reason we shall follow the path of thought from the second half of the thirties up to that of *Personal Knowledge*.

After wrestling with the monster of totalitarianism, Michael Polanyi seemingly felt prepared to outline a plan of a book that remained in manuscript. The intended book was given the title: 'The Struggle of Man in Society.' The summary of the seven chapters shows the main issues to which detailed argumentation and commentaries were put forward. The text is revealing for both the genesis of Polanyi's social theory in general and that of civil society in particular.

In the notes Polanyi unfolds his 'Archimedean fixed position' of attacking unshakeably the totalitarianism that utterly permeates every particle of argu- mentation. He also undertakes to set out his basic 'musical themes' henceforth, underlying each topic of his *oeuvre*. If we look back from the position reached in *Personal Knowledge* to this stage of his conceptual development, we see that the notions vaguely elaborated so far can be considered as being forward-pointing clues to the set of notions laid down later on. Polanyi now sees clearly why and what he wants to express but without yet finding the proper terms. He provides the conceptual scope of notions flexibly circumscribed and puts certain contents into them. However, both their border-lines and contents will incur considerable shifts and substantial modifications before *Personal Knowledge*. Thus this never published essay shows itself to have an abundant ensemble of groping formulations, several of which will be echoed by matured forms. We 'only' need to unstitch these strands still wrapped in the texture of argumentation.

When sketching his outline, Polanyi was intending to write a comprehensive book that would unify scientific, social and economic matters in itself. This is worth mentioning, for he abandoned this plan in 1944/45 and published three books (*Free Trade and Full Employment*, 1945, VI; *SFS; LL*) instead of one (see Polanyi, 1944). 'The Struggle of Man in Society' throws light on a unifying principle, on a 'Weltanschauung' that makes the essay an organic unity. Several formulations which appeared earlier can be integrated in later contexts though they do not figure explicitly in the latter.

For example, the notion of 'sectionalism' played an important part in 'The Struggle of Man in Society' but ceases to operate in *The Logic of Liberty*, turning into 'polycentricity', and reappears in *Personal Knowledge* under the term 'administration of undivided and civic culture'.

4. The development of Polanyi's thought in the '30s and '40s

Concentrating now on the main strain in 'The Struggle of Man', we realise that Polanyi postulates free society as an ideal and opposes it to totalitarianism. Briefly, he saw science to be threatened by a totalitarianism that caused corruption, backwardness and loss of independence. Man should revolt against it. But the struggle for independence for science as part of any free society needs intellectual arms. Polanyi searches for the means by which co-operation, coordination and integration in society can be achieved. His endeavour implies a critical element on this level of problem posing. He tries to point out that public good can be wielded by one centre but cannot be put directly into effect. In contrast, the supporters of totalitarianism claimed the superiority of central planning over liberal ways by trying to demonstrate the possibility of direct governing of public good. Already in a manuscript prior to 'The Struggle of Man in Society' Polanyi affirmed that the idea of revolution first implemented in the Soviet Union

> consists of the belief that gradual conciliation cannot lead to an effective improvement of social justice and that therefore human considerations have to be temporally set aside in order to establish better institutions Polanyi, 1938, p. 1).

Following the original Marxist tenet about the priority of the economic substructure to the social and ideological superstructure, the Communists first undertook to set up a new economic system (planned economy) to substitute for the market economy full of 'cash nexus', 'commodity fetishism', in brief, 'anarchism' (Polanyi, 1939, pp. 3-4). The Marxists, Polanyi maintained, concealed their moral claims by evoking an explicit contempt for moral values as being merely derivatives of class interests (moral inversion; see Polanyi, 51, p. 106; *PK*, pp. 227-237; Gelwick, 1977, pp. 6-14; Prosch, 1986, pp.86-8). They were driven, so to say, by the strength of logical inference to conclude that, not only that economics should be submitted to a central authority to plan for the welfare of the community, but so should the domains of society in which moral and intellectual claims are embodied. As these spheres of society left alone would run wild in anarchy, they ought to be abolished in favour of a planned economy (Polanyi, 1939-40, pp. 69/81), Marxists maintained.

Polanyi also strives to uncover the different kinds of integrative coefficients operating in the making and unified functioning both of totalitarian and liberal society. Therefore, he investigates social co-operation which can occur on a smaller and larger scale. In his initial investigation in 1939-40 he found that there was no general idea discovered so far 'by which co-operation of a multitude of men can be achieved in an impersonal way' (Polanyi, 1939-40, p. 34) as opposed to a personal one. However, one can find two *different* kinds of principles of freedom

which cannot be reduced to a common ground. The first one is the market operating for production and goods and services by the division of labour. The other is found in the scientific community by the help of which the co-ordination of scientific production is secured.

Henceforth Polanyi himself set out to investigate the principles of co-operation underlying both the economy and science. Undoubtedly, the machinery of Adam Smith's 'invisible hand' served him as a guiding ideal for not only economic but general freedom. He stated in 1938 that by the 'production of commodities for the market, the acquisition of money is turned by the proverbial invisible hand into the service of the community' (Polanyi, 1938, p. 3). The performance embodied in this statement cannot be overvalued because it postulates the possibility of wielding the public good without mediation by a central authority that would be uniquely opposed to carrying it out, as proponents of central planning assert. Consequently, Polanyi goes on to conclude that 'collectivism as a claim of supremacy of general interests over that of the individual concerned, can be manifested in various examples which are not related to planning' (Polanyi 1939-40, p. 40). To illustrate this he opposes *growth* to *construction*. The condition of growth, he argues, lies in the capacity of the community to improve through a very large number of gradual changes, each of which is due to the action of an individual and is profitable to that individual In contrast, construction is a change in which essentially incomplete stages appear under immediate domination. The latter seems to be obscurely defined, but Polanyi explains that construction operates as described because the planners must necessarily leave a large part of their work undefined since they lack the precise knowledge of all single facts.

Since at this stage of his intellectual advancement Polanyiconsiders growth and construction as distinguished mostly by the difference of view-points from which they are envisaged, we may see them retrospectively to be antecedents of two kinds of *order* that Polanyi will usher in the next year, 1941. These will play a decisive part in his social theory up to 1951. Already at the time in question he is seemingly intent on relating growth to the ideal of free society and construction to that of totalitarianism. But it is not yet clear. As he analyses the problem of co-ordination on an abstract level he has to acknowledge that a community continuously has problems to tackle by the way of construction (e.g., if the birth rate is too low) when an organiser of the general interest of the community might step in. Even more, he does not intend to exclude authorities from both forms of organisations. Far from it, he indicates that there are 'influential' persons, briefly, 'authorities' within the scope of growth, as we shall see soon.

* * *

Coming now closer to our subject, civil society, we can find important passages in the text in question. We have seen that Polanyi was striving to outline the main features of the general principle organising society. Let us switch our attention to how individuals grow. Polanyi points out that the new-born gradually grows into an *impersonal* complex social context made up by various sectors. In one sector (*section*) individuals act but within the scope of their profession and are judged publ icly without regard to the rest of their lives. In turn, they demand their private lives to be respected by others. Within 'office hours' they are submitted to leadership of different kinds. Yet the question now arises as to what co-efficient makes these independent sectors (or spheres) hold together, or, in other words, what is the organising general principle that Polanyi is going to establish? The answer is that there is competitive leadership within each autonomous sector and it, as such, is opposed to the 'immediate domination of totalitarianism' (Polanyi, 1939-40, p. 8). Competitive leadership is a function discharged by a wide range of influential persons seeking to rival each other. Within and between sectors they are continuously about 'to modify the structure and contents of their spheres in accordance with the standards of public opinion, and of those closely participating in the particular spheres' (Polanyi, 1939-40, p. 48). Competitive leadership as a motive force of free society infiltrates all parts of it, such as the proceedings in the law courts, rival canvassing of parties, charities, and schools of thought, with a public spirit that enters upon the struggle against totalitarian domination. That is why 'all prospective dictators try to abolish parties which are the matrices of competing leadership' (Polanyi, 1939-40, p.46). In such a free society it should be taken for granted that the State itself may be but one of the authorities in society, and that it is distinguished from the others by the fact that everyone belongs to it and that it is the main source of organised physical power (Polanyi, 1939-40, pp. 4&49). We have seen an impersonal *sectionalism* emerging in free society. This sectionalism is essential in our search for the place and significance of civil society in Polanyi's thought. Therefore it is indispensable to quote a long passage of 'The Struggle of Man' in which the substantial elements of sectionalism as related to civil society are revealed:

> Sectionalism implies the existence of a number of unrelated authorities . . Every section has its hierarchy . . , so that it really has a multiple internal structure. The various spheres competing for the attention . . . are trying to maintain or to exalt some sectional ideal of their own: even those concerned with the whole of society are, in reality aiming at a section only, though the section might go right

> across society . . . The wealth of liberty which society can offer increases directly with the number of separate authorities . . . The basis for these quasi-independent authorities must be that they present partial aspects of life each of which has a system of its own by which it sets its standards and carries out its functions. It is the assumption that the true way of living is not wholly known but must work itself out, by giving scope to the activities of more or less autonomous interests . . . By the same token which causes the free citizens to enter impersonally into the various equations left open to him by society, he is also deprived of aiming in the resulting activities at the good of society as a whole. The supreme good is not institutionally represented. There is no profession of guardianship of supreme welfare in a free society . . . In this sense the complex physiognomy of society is a democratic representation of the response of society to the channels to it by its social inheritance (Polanyi, 1939-40, pp. 57).

We emphasise that though the system of thought which Polanyi reached at that time will become more sophisticated, new notions will be introduced and worked out, and the established concepts will incur modifications, the important and main characteristics that are ascribed to civil society will not be essentially changed.

* * *

Between 1941 and 1951, in the second stage of Polanyi's intellectual development one of the most important events consisted of Polanyi's taking over the notion of 'order' from Köhler. This took place sometime in 1941 (Polanyi, 1941/IV, p. 4). Following Köhler he elaborated order in a more general sense (order in space and time; Polanyi, 194/1V) and then defined two kinds of order in society. The first one was identified in 'The Struggle of Man' as 'construction', the other as 'growth'. Now 'construction' is replaced by 'corporate order' and 'growth' by 'dynamic order'. Furthermore,what was designated as 'competitive leadership' assumes the new term 'mutual adjustment'. We will meet it again in *Personal Knowledge*, where, after it had played an important part in *The Logic of Liberty*, it ascends to a central position while the two kinds of order will cease to operate in *Personal Knowledge*.

Corporate order consists, in the more abstract sense, of sorting out things and assigning to each a place according to a prearranged plan. This kind of order in society is put under one authority and exists on a vast scale, e.g., in the army and various governmental departments. In contrast, dynamic order is brought about if one leaves things alone, since then, order is achieved by *internal forces* between particles, that is, by *mutual adjustment.* Competing private enterprises work in dynamic order. Yet Polanyi emphasises that there are other examples, of greater importance than those of material production:

> literature, art and sciences, custom, and law, in fact the entire progress of our cultural heritage as achieved by a co-operation based on the mutual adjustment (Polanyi, 1941/II pp. 2, 4).

These spheres of autonomous activity are given a new term: city. Cities can be broader or narrower in scope, but each incorporates mutual adjustment. Cities of businessmen, clergymen, etc. form a more or less independent body politic.

> There are certain general principles common to all these structures and these can be best made clear by systematically comparing and contrasting the various spheres. This book traces many such comparisons, all centring around science, and attempts to gain as a result, a complete analysis of the scientific body coupled with at least a general characteristic of the other bodies which surround it and form, in an aggregate, the rest of society'.

Polanyi conceives so in an introductory chapter of a planned book drafted on December 16, 1941 (Polanyi, 1941/III, p. 1) that throws light on the further crystallisation of the concept of the structure and dynamics of free society. Notice how firmly Polanyi sticks to the idea that the co-ordinating principle of science has to be expanded to the rest of society and to prevail there (though corporate orders can occur sporadically in free society, too; Polanyi, 1945/II, pp. 7-8).

In the '40s Polanyi went on with developing his conception and disentangling hidden implications from it. It is worthwhile presenting some of them. Firstly, moral arid intellectual standards which were already called into being in 'The Struggle of Man' Polanyi, 1937-40, p. 38) now become conspicuous as 'definite principles'. They afford protection. Even more, free society

> is not based on freedom of individuals but on freedom of principles . . . Absence of government control is balanced and may be outweighed by the acceptance of control by definite principles (Polanyi, 1945/IV, pp; 1-2).

Polanyi minutely points out that neither the economy nor jurisprudence can operate unless they incorporate into their functioning the ideals of fairness, decency and truth (Polanyi, 1945/V, pp. 4-5). To sum up: definite principles are the basis for the social edifice, and without putting them into effect no sound society can prove well grounded or habitable.

Secondly, Polanyi makes a clear distinction between individual and public functions, which may be considered as pointers to the distinction between individual and civic culture made in *Personal Knowledge*. Furthermore, he divides dynamic order into three kinds:

> Intellectual: of individual efforts contributing to an established system of ideas. Productive: of producers managing resources for commercial profit. Distributive: of consumers sharing current productions and awarding rewards to producers through the market (Polanyi, 1945/

IV, p. 3).

These activities are subsumed under individual functions. Public functions are also threefold: cultural, when the individual participates in preserving, transmitting or expanding the intellectual heritage; civic, when he appears as a witness or acts as a juryman; democratic, when he responds with criticism or support of political leadership, or offers himself as a leader and achieves leadership (Polanyi, 1956/V, p.1). Polanyi adds to this that

> The range and complexity of these functions, particularly in the cultural field, is a measure of the extent to which society depends for its very life on the liberal way (Polanyi, 1945/IV. p. 2).

Thirdly, a novel trait of the State deploys into the system. Polanyi perceives that the State must enter the scene of public interest adjudicating between competing interests for they are not able to judge between themselves. 'It would suppose a choice between rival interests on the basis of principles which override interests' (Polanyi, 1945/IV p.3) If some of the interests directly involve the public they should be properly safeguarded by the State. Consequently, the State may be said to be responsible, in the ultimate resort, for 'definite principles'.

All intellectual achievements and orientations elaborated in the thirties are collected and brought to a fitting climax by Polanyi in *The Logic of Liberty.* As Harry Prosch presented a large-scale interpretation of it in every respect, we need not provide a detailed analysis of its inherent conceptual arrangement, We merely recall that Polanyi in this book enlarged the number of examples, consolidated the contours of notions, reasserted tenets in more sophisticated ways, and gave more striking examples, but he did not essentially modify the results up to the fifties. It is sufficient to mention that he replaced 'dynamic order' by 'spontaneous order' and ushered in the term 'polycentricity' (*LL*, p. 184) for describing the interrelating aggregate of 'cities' joined to each other by mutual adjustment. We may therefore pass on to *Personal Knowledge.*

5. The fruition of Polanyi's thought in Personal Knowledge

With *Personal Knowledge* the reader enters a new world of Polanyi's theory. New terms, notions and concepts overwhelm the philosophical scene so as to enlarge the perspective. We have to suppose the reader to be familiar with this world. Let us focus on issues closely related to the subject of civil society so as not to stretch the frame of this paper too far. Thus we will be confined only to novel elements germane to our subject matter. Nevertheless, these new elements are not without antecedents, since almost all of them appeared earlier at least in germinal shape. Taking the chapter on 'conviviality', we face a notion we have not yet met. Polanyi ushers in the convivial element after he has dealt with intellectual passions of universal intent by bringing to light the civic co-efficients of those intellectual passions. He affirms that 'an intellectual passion can survive only with the support of a society which respects the values affirmed by these passions' (*PK*, p. 203). Thus conviviality, apart from its bearing on intellectual passions, is similar to the *Vergesellschaftung* described in German sociology (Simmel, Weber, etc.).

As for the content of conviviality, Polanyi uncovers it as the procedure through which fellowship comes into being. Two forms of it are sharing of experience and participation in joint activities, e.g., in rituals. These are instances of 'pure' conviviality as lower forms of coherence, preceding the stage of organised society. Pure conviviality embraces (by sharing of experience and rituals) 'a wide range of common values which are continuous with the impersonal appreciations laid down by morality, custom and law' (*PK*, p.212). It follows from this that

> the group has a claim to the conformity of its members, and that the interests of group life may legitimately rival and sometimes overrule those of the individual. This acknowledges a common good for the sake of which deviation may be suppressed and individuals be required to make sacrifices for defending the group against subversion and destruction from inside (*PK*, p. 212).

Thus for the administering the common good, convivial institutions take shape. This was not taken into account by Polanyi in the '40s, though a thorough investigation can possibly discover its seed.

As has already been said, pure conviviality does not suffice to bring about organised society. To make this it is necessary for four co-efficients to work in societal organisations, namely, sharing of conviction, sharing of fellowship, co-operation, authority or coercion. These are embodied in four kinds of institutions: those of culture (universities, theatres, churches, etc.); of conviviality (group loyalty, group rituals, common defence, social intercourse); of economics (property); of public power (which shelters and controls the cultural, convivial and economic institutions). The latter are civic institutions.

Remember the order as outlined in 'The Structure of Liberalism'. There we encounter three dynamic orders (intellectual, productive, distributive) and three public functions of the individual (cultural, civic, democratic). The positions of these elements within the theoretical framework are now rearranged in *Personal Knowledge.* Intellectual and cultural functions are fused into a new complex named 'individual culture', productive and distributive actions are fitted into economics, and the state that was appended to others from outside is a fourth institution. Convivial, economic and state institutions belong to civic culture. Though individual culture is guided by its own standards and

prompted by its own passions, it must be secured by established cultural institutions if its standards are to be socially cultivated. Yet, they are, at second hand, dependent on civic institutions, i.e., group loyalty, property and power. While civic culture is sustained by civic institutions, individual culture is not, as expected, sustained by cultural institutions alone, but also by civic institutions. It follows that, though culture, both individual and civic, proceeds under the influence of intellectual and moral standards, the civic pole relies ultimately on institutions of civic culture, that is, on group loyalty, property and power. Furthermore, as group loyalty is parochial, property appetitive and authority violent, civic culture can be at variance with moral and intellectual standards. As a result, 'the genuineness of moral standards will be rendered suspect when it is realised that they are upheld by force, based on property and imbued with local loyalty' (*PK*, p. 216). This distortion can take place at a critical age when civic institutions degenerate into local parochialism (ethnocentricity, nationalism), greedy economic interest (when actors of economics fall short of fairness, decency, and the running of the economy is not tempered by humanistic social policy), and the mere violence of state power. In this critical age 'this depreciation of thought will tend to spread and to bring about eventually the subjection of all thought to local patriotism, economic interest and the power of the State' (*PK*, p. 216).

This is one of the extremes of civic institutions. However, on the other hand, there can exist, it might be said, a positive extreme too. This is the case of 'a happy people' to whom their civic culture is their civic home, as Polanyi conceives it, for their intellectual and moral passions sustain the civic culture in an esoteric way. Morality (with its allies, custom and law) becomes an instrument of civic culture, and this intertwining of civic exigencies with the ideals of morality lets civic matters take shape by the same principles which effectively sustain the freedom of individual thought, i. e., of individual culture. The two kinds of culture need a twofold administration. Individual culture is administered by the mutual adjustment that Polanyi presents with the example of science. We are already familiar with it. Yet it is noticeable that the administration of civic culture has no peculiar machinery akin to mutual adjustment. On the basis of our tracing of the formation of Polanyi's thought, we venture to state that Polanyi did not build up any specific organising machinery for the administration of civic culture, not because of negligence, but because he took for granted that mutual adjustment was essentially inherent within civic culture too (think of the 'invisible hand' in economics). This is the reason why, when dealing with the administration of civic culture, he emphasises only the moral improvements which are embodied in humanistic laws and its institutions.

* * *

The question to be answered now arises: Where should we locate civil society within the context of conviviality? Undoubtedly, we may exclude the state. But convivial and economic institutions are organic parts of civil society as two fields of public life which, in a free society, are instrumental for sustaining individual culture (e.g. support for arts, churches, etc.). In totalitarianism or autocracy, autonomous convivial and economic initiatives belong to civil society rather than the private sphere. This is so that they can free themselves from the domains controlled by the state.

Yet, may individual culture be subsumed under the concept of civil society?

To gain an apposite answer let us remember what Polanyi taught about the republic of science. It should be taken for granted that a lab or a department of any science in any university does not imply, in itself, membership in civil society, but if they take part in the inner public life of a university (e.g. in decision about budgets) they form part and parcel of the university's 'civil society', for their activity in this case goes beyond that of strict science. Do not forget: 'civil' means activities turned into the public sphere beyond privacy, be the latter that of a private person or a body of any kind. Therefore, a citizen differentiates himself as a private person from himself as 'public person' by his participation in public affairs ('res publica'), and so do associations, unions, circles, etc. of any kind. However, there is no clear-cut or eternal demarcation between private and public. In totalitarian régimes the border-line of public shifts closer to the private sphere since totalitarian power seeks to control people's behaviour and thought, even in the family. The latter happened in the climax of Fascism and Bolshevism. Its extreme form ceased to hold in *Spätsozialismus*, but still worked in a more hidden form and between tighter limits until the collapse. Few know in Western countries that an officer of the political police (a 'communicator') was appointed in each university, whose task was to control the behaviour of students and the teaching staff. This person took the 'necessary measures' if he noticed some 'objectionable behaviour'. Polanyi knew fairly well the actual situation even of everyday life in Communism, and accordingly, he was able to state convincingly that circles and professional associations are feared and hated by modern totalitarian rulers (See *TD*, p. 84). This means that all elements of the theory which we can legitimately interpret as Polanyi's conception of civil society, are dynamic in Polanyi's thought. This is true, first, because a part of the activities of the associations in question is one mediating to public spheres, and second, because their activity, though being in itself 'neutral' can be threatened in particular

prompted by its own passions, it must be secured by established cultural institutions if its standards are to be socially cultivated. Yet, they are, at second hand, dependent on civic institutions, i.e., group loyalty, property and power. While civic culture is sustained by civic institutions, individual culture is not, as expected, sustained by cultural institutions alone, but also by civic institutions. It follows that, though culture, both individual and civic, proceeds under the influence of intellectual and moral standards, the civic pole relies ultimately on institutions of civic culture, that is, on group loyalty, property and power. Furthermore, as group loyalty is parochial, property appetitive and authority violent, civic culture can be at variance with moral and intellectual standards. As a result, 'the genuineness of moral standards will be rendered suspect when it is realised that they are upheld by force, based on property and imbued with local loyalty' (*PK*, p. 216). This distortion can take place at a critical age when civic institutions degenerate into local parochialism (ethnocentricity, nationalism), greedy economic interest (when actors of economics fall short of fairness, decency, and the running of the economy is not tempered by humanistic social policy), and the mere violence of state power. In this critical age 'this depreciation of thought will tend to spread and to bring about eventually the subjection of all thought to local patriotism, economic interest and the power of the State' (*PK*, p. 216).

This is one of the extremes of civic institutions. However, on the other hand, there can exist, it might be said, a positive extreme too. This is the case of 'a happy people' to whom their civic culture is their civic home, as Polanyi conceives it, for their intellectual and moral passions sustain the civic culture in an esoteric way. Morality (with its allies, custom and law) becomes an instrument of civic culture, and this intertwining of civic exigencies with the ideals of morality lets civic matters take shape by the same principles which effectively sustain the freedom of individual thought, i. e., of individual culture. The two kinds of culture need a twofold administration. Individual culture is administered by the mutual adjustment that Polanyi presents with the example of science. We are already familiar with it. Yet it is noticeable that the administration of civic culture has no peculiar machinery akin to mutual adjustment. On the basis of our tracing of the formation of Polanyi's thought, we venture to state that Polanyi did not build up any specific organising machinery for the administration of civic culture, not because of negligence, but because he took for granted that mutual adjustment was essentially inherent within civic culture too (think of the 'invisible hand' in economics). This is the reason why, when dealing with the administration of civic culture, he emphasises only the moral improvements which are embodied in humanistic laws and its institutions.

* * *

The question to be answered now arises: Where should we locate civil society within the context of conviviality? Undoubtedly, we may exclude the state. But convivial and economic institutions are organic parts of civil society as two fields of public life which, in a free society, are instrumental for sustaining individual culture (e.g. support for arts, churches, etc.). In totalitarianism or autocracy, autonomous convivial and economic initiatives belong to civil society rather than the private sphere. This is so that they can free themselves from the domains controlled by the state.

Yet, may individual culture be subsumed under the concept of civil society?

To gain an apposite answer let us remember what Polanyi taught about the republic of science. It should be taken for granted that a lab or a department of any science in any university does not imply, in itself, membership in civil society, but if they take part in the inner public life of a university (e.g. in decision about budgets) they form part and parcel of the university's 'civil society', for their activity in this case goes beyond that of strict science. Do not forget: 'civil' means activities turned into the public sphere beyond privacy, be the latter that of a private person or a body of any kind. Therefore, a citizen differentiates himself as a private person from himself as 'public person' by his participation in public affairs ('res publica'), and so do associations, unions, circles, etc. of any kind. However, there is no clear-cut or eternal demarcation between private and public. In totalitarian régimes the border-line of public shifts closer to the private sphere since totalitarian power seeks to control people's behaviour and thought, even in the family. The latter happened in the climax of Fascism and Bolshevism. Its extreme form ceased to hold in *Spätsozialismus*, but still worked in a more hidden form and between tighter limits until the collapse. Few know in Western countries that an officer of the political police (a 'communicator') was appointed in each university, whose task was to control the behaviour of students and the teaching staff. This person took the 'necessary measures' if he noticed some 'objectionable behaviour'. Polanyi knew fairly well the actual situation even of everyday life in Communism, and accordingly, he was able to state convincingly that circles and professional associations are feared and hated by modern totalitarian rulers (See *TD*, p. 84). This means that all elements of the theory which we can legitimately interpret as Polanyi's conception of civil society, are dynamic in Polanyi's thought. This is true, first, because a part of the activities of the associations in question is one mediating to public spheres, and second, because their activity, though being in itself 'neutral' can be threatened in particular

historical circumstances (e.g., in totalitari- anism), and it is just because they are threatened by public power that their activity can be qualified as public.

By pointing out the dynamism of Polanyi's conception of civil society (a conception that we took in our interpretation of recent discussion in Eastern Europe) we have come to the end of this essay. However, we have to deal briefly with some objections which cast doubt on Polanyi's constant endeavour to expound the republic of science as a model valid for civil society.

6. Can science be a model for society?

In an unpublished manuscript written by Polanyi in 1945 stands the following phrase: 'The autonomy of science as a paradigm of the liberal way' (Polanyi, 1945/II, p. 4). Polanyi also explains the reasons for it. He argues that both science and a free society can be guided by 'transcendent values' the observation of which by independent individuals establishes the moral and intellectual order. On the basis of it 'a considerable scope for civic functions' comes into being. By the same token there is 'a dynamic order of co-operating scientists, lawyers, artists, scholars, divines' (Polanyi, 1945/II, pp. 1-2). These thoughts of Polanyi were our directing stars in this essay, and in compliance with it we cannot agree with the views that attack the conception of 'The Republic of Science' for its suggestion that science resembles the body politic of, say, civil society.

Bertrand de Jouvenel questions whether 'the model of the dedicated company can, with proper adjustment, be used for civil society as Polanyi seems to suggest' (de Jouvenel, 1961, p. 136). De Jouvenel argues that since dedicated companies, like science, repose on principles diametrically opposed to those embodied in the rest of society, the latter has no common criteria of good and bad other than that of preservation of society itself. Accordingly, de Jouvenel maintains, the aim of preservation beyond morality irresistibly leads up to possible distortion of society, for 'where no truth is acknowledged, fanaticism is impressive'

Since we concerned ourselves with this view elsewhere (see Nagy, 1992) it is sufficient to confine ourselves to some comments. De Jouvenel seems to overlook that Polanyi was aware of possible distortions of civil society when he stated, e.g., in *Personal Knowledge*, that 'civic culture still remains dependent on force and material ends, and remains therefore suspect' (*PK*, p. 226), but, in spite of this (or: just for this reason) he declined to state that the operating principles of a free society should be akin to those of science. Why? If de Jouvenel would have thought over the possible answer to this question, he could have come to the solution as it is convincingly pointed out by H. Prosch when he interpreted Polanyi's message in this respect: 'a free democratic political republic cannot exist unless there are some basic ideals, principles held in common by all the members of political community' (Prosch, 1986, p. 279).

Recently Richard Allen made an assumption similar to that of de Jouvenel. In his book on Polanyi, after having given a correct interpretation of Polanyi's views on the republic of science, he makes the objection:

> One may wonder, can actual society arise and he maintained by belief in and dedication to only formal and general principles? The republic of science exists within a wider society, or across several societies, and participation in it is voluntary and does not occupy the whole of one's life. And so it may not be quite so apt a model for that wider society' (Allen, 1990, p. 75).

In consequence, Allen goes on to argue that the wider society must cohere around 'a more specific set of beliefs', and it is, thereby, more like a religious body with its core of dogmas, rites and practices.

We cannot agree with this view. We have to point to the fact that modern societies are built up on norms which get more and more formal as opposed to those of religious bodies. The members of churches really have to abide by fairly strict norms, sometimes by dogmas. However, modern men belong at the same time to several groups which require sometimes different, sometimes contrasting behaviour (this is 'conflict of roles' in sociology), but this does not alter their conformity to more general norms (law, ethical norms, moral standards, all their 'mores' as Polanyi names them in *Personal Knowledge*). It follows from this that the wider society provides people a larger 'place for free play' than religious bodies do, but it does not mean that people cannot obey both kind of norms. If modern societies were regulated not by general but by specific norms, as Allen claims, they would fall back into the 'static' state of society preceding the modern one.

The other distinction made by Allen, namely, that participation in science is voluntary and does not occupy the whole of one's life, does not alter essentially the main goal of both science and society. It is their members' obligation to maintain these standards, no matter what their personal motives are for being members.

In a paper presented on the Polanyi centennial conference held in Budapest, R. Allen repeated his views about the difference between voluntary associations and the wider society:

> A state or body politic is not like a voluntary society within it, for it is constituted on the basis of historic or prescriptive right and obligation and not by contract. Our obligations to it and rights within it are inherited and not contracted (Allen, 1992, p. 97).

It seems that R. Allen overstates the difference between the two kinds of social aggregates, and his concep-

tion of tradition runs counter to Polanyi's teaching about the very nature of a tradition. In Polanyi's thought tradition is not as rigid as that of R. Allen. For example, in *Science, Faith and Society* Polanyi argues that, though scientists should assent to the premises of science, the great discoverers slightly modified them.

> There is in fact no aspect of science, including even mathematics, in which the fundamental presuppositions, the methods of investigation, and the criteria used for verification have not undergone a series of marked changes since the inception of modern science 300 years ago (*SFS*, p. 89).

The same holds for Common Law (see *PK*, p. 54) and, according to Polanyi to other fields of 'social lore'.

Already Tom Paine switched our attention to the fact, as against the far-fetched traditionalism of Edmund Burke, that all tradition began once upon a time. And Polanyi did not say that Tom Paine should have replaced Edmund Burke, but rejected 'Paine's demand for the absolute self-determination of each generation', but does so for the sake of its own ideal of unlimited human and social improvement for his position 'accepts Burke's thesis that freedom must be rooted in tradition, but transposes it into a system cultivating radical progress' (*KB*, p. 71; see also *TD*, p. 63). Richard Allen is, of course, right in saying that even voluntary societies themselves depend upon traditions, but he disproves, just by this assumption, his other tenet which posits an unbridgeable gap between state, body politic and civil society, viz. voluntary societies. Both of them can be maintained by a creative reliance upon tradition. The latter is elucidated by Drusilla Scott in very illuminating examples in her book (see Scott, 1985, pp. 89-94). These examples clearly show that 'the same sort of lively tension between tradition and innovation can operate in these two fields as in the scientific field' (Scott, 1985, p. 88). For tradition should simultaneously be kept and renewed as Jesus puts it: 'every teacher of the Law who becomes a disciple in the Kingdom of heaven is like the owner of a home who takes new and old things out of his store-room' (Matt. 14: 52).

Dept of Sociology and Social Policy
Janus Pannonius University
Pécs, Hungary

Reprinted with permission from ***Polanyiana****, Vol 2 No. 4 (1992) & Vol. 3 No. 1 (1993) joint issue.*

Bibliography

Michael Polanyi's works (other than those listed on the inside rear cover):

1939 'It has been forcibly stated'. Polanyi Papers in the Joseph Regenstein Library, University of Chicago. Box 26, Folder 1

1939-40 'The Struggle of Man in Society'. Box 26, Folder 2.

1941/I. 'Plan for lecture on 7th October at SCM'. Box 26, Folder 8.

1941/II. 'Ideas and Cities'. Box 26, Folder 8.

1941/IV 'The Process of Legislation'. Box 26, Folder 9.

1941/V 'The Rights and Duties of Science'. Box 26, Folder 10.

1941/VI 'Order in Space and Time'. Box 26, Folder 6.

1942 'The City of Science'. Box 27, Folder 3.

1944 Letter to G. Stolper. 9th July, 1944. Box 4, Folder 11.

1945/I 'Programme of a Liberal Philosophy'. Box 30, Folder 1.

1945/II 'The Liberal and the Totalitarian Ways'. Box 30, Folder 1.

1945/III 'Liberalism—Rise and Decline'. Box 30, Folder 1.

1945/IV 'The Structure of Liberalism'. Box 30, Folder 1.

1945/V 'Our Times'. A proposed new quarterly journal to be published by The Manchester Literary and Philosophical Society. Box 4, Folder 12.

Other works referred to:

Ackermann, Bruce (1992), *The Future of Liberal Revolution*. New Haven and London: Yale University Press.

Allen, Richard (1990), *Polanyi*. London: The Claridge Press.

Allen, Richard (1992), 'The Limits of Contract', *Polanyiana* 1992/1-2, pp.90-98.

Arató, András (1990), 'Forradalom, civil társadalom és demokrácia' ('Revolution, civil society and democracy in Eastern Europe') *Mozgó Világ* 1990/8, pp.11-22.

Ash, Timothy Garton (1990), *We the People: The Revolution of '89*. Cambridge: Granta Books.

Bibó, István (1991), *Democracy, Revolution, Self-Determination: Selected Writings* (ed. Károly Nagy; trans. András Boros-Kazai) *Atlantic Studies on Society in Change*, No. 69. New York: Columbia U. P.

Frentzel-Zagorska, Janina (1990), 'Civil society in Poland and Hungary', *Soviet Studies* Vol. 42. No. 4, pp.759-77.

Gouldner, Alvin (1960), 'The norm of reciprocity: a preliminary statement', *American Sociological Review*, XXV, pp.162-178.

Havel, Vaclav (1988), *Living in Truth*. London: Faber and Faber.

Jouvenel, Bertrand de (1961), 'The republic of science', in *The Logic of Personal Knowledge: Essays presented to Michael Polanyi on his Seventieth Birthday*. Glencoe: The Free Press.

Kemény, István (1991), *Közeröl s távolról* (*Closely and Remote*). Budapest: Gondolat.

Kis, János, Ferenc Köszeg, Ottilia Solt, (1992), 'Társadalmi szerzödés' ('Social contract') *Beszélö II*, pp.749-791. Budapest: AB-Beszélö Kiadó.

Nagy, Endre (1991), 'A civil társadalom paradigmái' ('Paradigms of civil society'), (unpublished).

Continued on p.43

WORKING PAPERS

TOWARDS A NEW METAPHYSICS (1)

G. G. Constandache

'Metaphysics of Computers' defines a technical field that we are going briefly to refer to by recalling Mole's definition of the General System Theory as 'a new philosophy of the man-environment relationship'[1]. L. von Bertalanffy stresses the urgency of supporting that conceptual attitude that opposes reductionism and which he called 'perspectivism'[2].

All joint research endeavours entitled Artificial Intelligence, through a willingly conducted effort of de-anthropologising present us novel representations of certain rational aspects of the ambient reality. Computer soft-ware, identified by some theoreticians with either the mind or the spirit, is essentially nothing else but a presentation of symbolic algorithms adapted to human aims.

The main reasons for the present concern with the metaphysics of computers are quite evident. This is not only a matter of the inquisitiveness that cognitive sciences have awaken, or the temptation that certain metaphors may arouse, such as that of the computer and the brain, but also of the progress of computer technologies, as well as of a vague feeling that we are witnessing the birth of a new paradigm. Besides, I consider that the inner evolution of metaphysics can be described in such a way as to display its gradual impact on all these external factors. Thanks to cognitive sciences we can attain conscious cognition through ceaselessly musing on our own functioning conditions, as P. Suppes proved[3]. Technique itself participates indirectly, and in a complex manner, in the reconstruction of what has always been the object of metaphysics, just as M. Bunge pointed out[4].

The new metaphysics does not limit itself to objective and reductionist cognition, so very much peculiar to traditional systems, being thus capable of approaching the singularity and complexity of phenomena (as J.M. Besnier has recently shown). This is exactly what these papers are meant to illustrate, viz. various current problems and questions that can be put under the heading of *Metaphysics of Computers* (a volume published with the help of the Soros Foundation for an Open Society in Romania, Computer Publishing Centre, 1994).

1. TOWARDS A METAPHYSICS OF COMPUTERS

1. Hints and sources

If we are to translate Jean Piaget's dictum, 'No human being is fulfilled before having reached philosophy', we can say that no field of research has ripened unless a taste of metaphysical accounting for is felt. And by 'metaphysics' we generally mean a discipline focusing on philosophical presuppositions or 'hyper-hypotheses' derived from domains of advanced research. P. Suppes, for instance, proposed a 'metaphysics of probabilities'[3], based on probabilistic concepts and hypotheses, and meant to tackle epistemological and metaphysical issues, in an attempt to justify both ordinary models of thinking and methodologies of scientific investigation.

The field I shall refer to is being claimed by various *cognitive sciences*. It has raised heated controversies (some people hold that research in this direction is a matter of fashion, or an effective pretext to wring out subsidies, while others estimate it as a mere attempt to refresh worn-out traditions, if not downright imposture); but these very arguments offer excellent opportunities to identify new metaphysical presuppositions, which may gather weight to the extent that they are systematically approached and their concepts elaborated. In fact, any significant field of research that has reached a mature stage (recognisable by its claim of universality, and of laying a new basis for co-operation among sciences) is challenging and stimulates speculation. For example, M. Bunge designed a 'metaphysics of technology'[4], starting from general technology, and developed a series of hypotheses which he called 'techno-ontological'. In his view, the investigation of philosophical presuppositions goes as far as the epistemic background of technology, and techno-axiology.

However, the actual validation of a new area of cognition will always imply a critical assessment of research results. And the attempt to clarify various problems of cognition in a meta- physical perspective can reap only benefits from challenges of either scientific or philosophical nature. With H. Dreyfus,

intellect is not *something* within the body, which we might imitate *irrespective* of the body itself. Because the body is not only a source of information about the environment, but also the source of our needs and impulses further to investigate the surrounding world (1972, 'What computers can't do: a critique of artificial reason'). Similarly, J. Searle maintains that it would be a mistake to assume that *simulation* meant the making of a duplicate, since computers cannot *duplicate* the causative potential of the brain (1982, 'The myth of the computer').

2. Main themes

Continual questioning of the big metaphysical issues reveals the resourcefulness of cognitive sciences. After all, metaphysics is no more than a preliminary systematising of knowledge, prepar- the ground for philosophical options, that is global view of the universe. What really counts at this point is to scoop out possibilities of renewing our vision of the world, starting from specific experiences in the field mentioned above. Of great relevance in this respect are such artefacts as computers, whose performances, when fed 'intelligent' programs, can be effectively compared with the procedures of the human brain which they apparently reproduce.

The current opinion about this domain is that it is governed by criteria coming from natural sciences or computer engineering. We must admit that cognitive sciences do rely on data provided by studies in artificial intelligence, psychology of cognition, linguistics, logic, neuro-science, and so on. This does not necessarily lead to the conclusion that epistemic differences among the various disciplines of cognition are disappearing. On the contrary, what I wish to demonstrate is that all these disciplines are expanding modern awareness for a new philosophy to build on.

Here now are several significant contributions that have re-opened discussions on major issues:

1. Intellectual activity, as being or not being specifically human (A. Turing, 'Computing Machinery and Intelligence', 1950).

2. God's creation as compared to man's creation and the output of machines (N. Wiener, 'God and Golem', 1964).

3. Technological patterns of the use of computers in various arts (A. Moles, 'Art et ordinateur', 1971).

4. The spirit (or soul) as a programming system (D. Dennet, 'Where am I?', 1978).

5. The critique of the anthropocentric outlook on reason (D. Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid*, New York, Vintage Books 1980).

6. The naturalisation of artificial intelligence (C. Giumale, 'Generatia a cincea de calculatoare, un pas spre programarea naturală?'/'Is the fifth generation of computers a step forward towards natural programming?', 1985).

7. The naturalisation of cognition (R.J. Bogdan (ed.), *Mind and Common Sense*, 1991).

3. Predecessors and subsequent reconsiderations

The variety and lack of homogeneity of the approaches which I am going to review (some of them with certain reservations) under the heading of 'Metaphysics of computers' is probably due to the fact that most of them buttress their claims of novelty with arguments taken from the history of philosophy.

Even those promoters of cognitive sciences who are trying to put an end to metaphysical speculations are constantly worrying about the question of the philosophical legitimacy of these disciplines. Hobbes was initially claimed as a forefather, since he had defined reason as a calculus of symbols and precise rules. He had urged his contemporaries to give up vain efforts to explain significance through the resemblance between the ideas and the things they represented. At the same time he authorised formal playing with ideal elements, to which any significance might be attached. J. Haugeland ('L'Esprit dans la machine', 1989) shows that Hobbes is superseded by Hume in the vision of those scientists who invoke the autonomy of reflexive heuristic computers, in support of a wholly dehumanised conception of intelligence. Hume conceived the development of intelligence as a result of a mechanical association of ideas (his model being inspired by the dynamics of planets subject to gravitation). His hint at a non-intentional finality made him a fore-runner of the theory of self-modelling.

H. Altan, a biologist and cybernetician, reconsidered the relation between soul and body in functionalist terms, outlining a sort of 'sweetened materialism with a structuralist flavour' (1986, 'À tort et à raison; intercritique de la science et du mythe'). Yet, even the simplest information networks could set forth the limits of strict mechanism, while the latest developments in automation and artificial intelligence have quite discredited reductionism, at least in its classical, hard-line version. In J. M. Besnier's view, soft reductionism is reminiscent of Kant's 'maxim of reflexive judgment' (non-objective), present in any mechanist approach to nature. Monism will always be the basis of reductionist theses, and although it is illegitimately accredited as an actual reality, it remains an effective method (1990, 'Les sciences cognitives, ont-elles raison de l'âme?').

D. Laplane, a neuro-psychologist, pleaded for the rehabilitation of Bergson's parallelism, restating—in the name of human liberty and dignity—the absolute incompatibility of mechanist sciences with spiritual matters. This anti-reductionism highlights the special, irreducible character of our higher faculties, and shows that our statements about them can be neither validated nor dismissed by the *objective* science of today or tomorrow (1987, 'La mouche dans le bocal.

Essai sur la liberté de l'homme neuronal').

The resort to precedents in order to persuade the public of the novelty of present research in this possible (virtual) field, still lacking a central category to induce a structure and give it a style, is the most striking symptom of current endeavours. No individual approach has reached maturity so far, nor any new essential feature of the above-mentioned artefacts has been revealed yet and awaits a new concept to express it and a new philosophical category to be worked out in answer to yet another fundamental question about human existence.

These frequent references to the past seem to indicate that an integrative category is assiduously sought. Several versions of the theory of identity (asserting that 'mental states are a sub-group of brain states') require a psycho-neural monism or an emergent materialism, but they are vigorously challenged by the supporters of interactionist dualism (J. Eccles and K. Popper, *The Self and Its Brain*, 1977).

The assertion that any mental event can be identified with a physical one allows for a direct correspondence between brain functioning and the structure of its neuronic network. But in that case, J.P. Changeux concludes, 'man has nothing in common with spirit any more, so we can deal with as a *neuronal man*' (1983, L'Homme neuronal'). Such ideas recall La Mettrie, who used to say that

> the soul is no more than a principle of movement, or sensitive material part of the brain . . . . Man is a machine, and in the whole universe there is a single substance diversified through modifications' (1747, *L'Homme machine*).

Philosophers today still cherish Vaucanson's project of building an 'artificial man' (18th C.). A proof of this is the equally naive, and recently concocted theory of the limiting case (which, we must admit, draws upon valuable scientific data): it states that 'The more we learn about organisms, the closer we come to the conclusion that they are not merely analogous to machines, in fact they *are* machines' (McCulloch, *Embodiments of Mind*, 1965).

Artificial intelligence has usually been associated with the 'deconstruction of subjectivity', which is a new critique of the Cartesian dualism, as it gives a rigorous description of the *subjectless process*, a notion translating the structuralists' 'death of man'. Deconstruction expanded to subjectivist metaphysics resumes the psycho-analytical argument that volition is the result of a mechanism, the same as perception, reason, and even conscience, being the 'visible top of a pyramid of functions describable in mechanical terms' (Freud).

An overall assessment of research on artificial intelligence shows it has developed in three directions: technical, psychological, and philosophical. Widening the range of interpretations in this field used to be regarded as prejudicial to man's metaphysical status, which explains why suggestions were put forward that its scope be limited, as for example to 'Solving logical or ludic problems, translating languages, representing and processing information with the purpose of investigating and recognising forms' (G. Boss, 'Les machines à penser: l'homme et l'ordinateur', 1987). The danger here is quite obvious: neuro-biology and artificial intelligence are being used to justify a counter-sense—'anyone can see that there is no spirit' (J. Arsac, 'Un informaticien, Entretien avec J. Vauthier', 1989).

The tendency to generalise the use of computers designed to reproduce human performance is challenging the thesis of man's supremacy on Earth, as the only being endowed by God with a creative intelligence. A new era of self-regulating and potentially self-reproducing machines will swivel on the paradigm of autopoiesis; then the definition of the nervous system as a closed operational system is likely to cause trouble.

> It may seem presumptuous to propose another perspective, almost opposite to the generally accepted view, and say that the nervous system operates with neither inputs nor outputs. Its cognitive function reflects nothing but its own structure, and it does not collect but imposes its own information upon its environment. (A. Varela, *Autonomie et connaissance. Essai sur le vivant*, 1989).

Such a statement obviously falls wide apart from the philosophical choice of the contemporary connectionism, which accounts for the cognitive system as a whole in terms of the mechanics of the neuronic network (P. Livret, 'Cybernétique, auto-organisation et néo-connexionnisme', 1985).

What is still required is a knowledge of the unity of this field, to be described systematically, by means of *a new category* as it previously occurred with the concept of random happening, mathematically expressed through probability, which led to a new type of determinism, rendered through a new categorical structure of general philosophical concepts (drawn up by P. Suppes).

So, if we admit that 'no human being is fulfilled before having reached philosophy', and that a metaphysical insight is a prerequisite of maturity in any field of research, we cannot help observing that the various *cognitive sciences*, gravitating towards the simulation of human intelligence, are actually expanding, exploding. Controversies and challenges in this particular field testify that a 'metaphysics of computers' is being worked out. Its very name points to the impact of research in artificial intelligence and connected areas on the traditional philosophical disciplines (ontology, epistemology, etc.), as well as to the relevance of a metaphysical outlook for the progress of computer science and technology. (For a bibliography on this topic see, G.G. Constandache, *Noologia abisala si metafizica ordinatoarelor*, Bucharest, COS, 1991).

2. COMPUTER METAPHYSICS:

THE RENEWAL OF A TRADITIONAL DISCIPLINE?

1. The revival of metaphysics

The question mark in the title suggests that, today, we can ask ourselves whether a new metaphysics is possible starting from the existence of computers.

Taking metaphysics as a particular means of philosophical reflection, we accept that it overruns, by assuming the axiological and hemeneutical perspectives, rigorous scientific conclusions. In so far as the metaphysician suggests and discusses a project which engages him with original presuppositions, that is to say, superhypotheses, he is attracted by relevant issues beyond the limits of empiricism or scientific experiment. We could most properly characterise metaphysics as a preliminary effort at systematising knowledge, and going on subsequently to establish a certain philosophical option. 'Getting rid of essentialist automatisms and illusions, post-Wittgensteinian language and discourse can separate themselves from the false vision of a *universal order* which would mask reality and betray facts by presenting them as regular incorporations of particulars into universals' (Adrian-Paul Iliescu, *Filosofia limbajului si limbajul filosofei* (*The Philosophy of Language and the Language of Philosophy*), Editura Stiintifica si Enciclopedica, Bucharest, 1989, p.311).

Metaphysics appeared in Antiquity, thanks to Aristotle (even though the name, as the inventor suggests, was given on), having as its object 'being as being', and specifically the supreme being. Modern metaphysics includes the contributions of several great thinkers, of which we mention here Descartes (dualism), Leibniz (monadology) and Kant (criticism). Yet it is Hegel who adds a particular (non-dialectical) note to 'metaphysics', causing some of his disciples (mainly the 'Left Hegelians') to consider it obsolete.

The metaphysical constructions in contemporary philosophy, especially the work of Bergson, Whitehead and Heidegger, were discredited in many ways by those who upheld the idea of unified science. Otto Neurath called the preference for a totalising vision over existence an 'attraction to combine concepts and statements which are not empirically tested' (Neurath, Carnap, Morris (eds), *International Encyclopaedia of Unified Science*, University of Chicago Press, 1938, Vol. I, Pt I, p.5). But in recent decades we have witnessed an improvement in the reputation of metaphysical approaches whereas 'logical positivism has nearly faded from the scene' (P. Suppes, *Probabilistic Metaphysics*, p.10).

There have been long and conflicting discussions of this problem. For Paul Ricoeur (*La philosophie*, UNESCO 1978, p.1129), if philosophy cannot include science in a larger and more fundamental view of reality (the 'synthetic' project), none the less it cannot limit itself to the role of criticism of language (the 'analytic' project), and it would not be able to explore human subjectivity ('the phenomenological-existentialist' project).

The revival of metaphysics is defended by the French 'new philosophers'. A veiled ethnocentrism has to give place to a relativism of 'respect for polyphony, for the diversity of cultural ensembles in the ever renewed effort to prevent the reduction of the Other to the size of the Same' (Jean-Marie Benoist, *Marx est mort*, Paris, 1970, p.132). It is suggested that a European confederation is desirable and that it should be based on decentralisation and cultural polyphony, an open system which should have respect for each country, those regions so vivid and rich in memories. Bernard-Henri Levi, the champion of the French New Philosophers, characterises the thinker thus:

> before being on the right or the left, well or ill intentioned, before being anti-social or asocial, literary creation—writing in itself —is, before anything, a metaphysical activity[5].

An indirect but sensitive appeal is found in the environmental sciences, through the anti-reductionist attitude of ecology. Research into the properties which characterise only the parts of a complex system can offer us useful explanations, but never a full understanding of the whole.

> Reductionism tends to separate the scientific disciplines from each other and all of them from the real world. Lack of communication between such special sciences generates a great source of difficulties in understanding the environmental problems . . . . Reductionism tends to separate the scientific disciplines from the problems affecting the human condition (Barry Commoner, *The Closing Circle*[6]).

No scientific principle can guide us when we have to choose; value judgments appeal to persuasions and beliefs. In any genuine democracy, moral problems, of social and political judgment, belong not only to the competence of the experts, but make up the joint interest of the whole nation through its elected representatives. Besides, for exercising the right to choose, relevant scientific data are needed, and they become significant in the perspective of a philosophical or religious conception.

There are many signs of a new interest in metaphysics. And certain research programmes or the dominant topics of some publications

could be placed under the heading of 'computer metaphysics'. It is important to stress that these attempts originate, with almost no exceptions, in the scientific community.

2. Metaphysics and technology

Pointing out that numerous philosophical aspects are to be found in the elaboration of technological policy and projects, M. Bunge states: 'Technology inherited from fundamental science a part of its ontology and, in exchange, it gave birth to a remarkable metaphysics of its own'[7]. Before anything else, 'general technology is the most outstanding contribution of technology to ontology and metaphysics' (p.383). Besides, contemporary technology—far from only asking questions of the metaphysician—supplies him with new theories about systems in general , the theory of automata, machines, networks, control, etc. The affiliation of these theories to *ontology* was fully justified.

> They refer to entire genera (rather than species) of systems as interdisciplinary theories, applicable from one field to another . . . They are essentially independent of the nature of matter, and therefore of any particular physical or chemical laws. (They rather concentrate upon the structure and behaviour than upon the specific composition and mechanisms . . .) They are not likely to be tested without further hardships, as they do not generate predictions, but they can be so made that they yield projections (pp. 386-7).

Criticising neo-traditional metaphysics, P. Suppes wrote: 'Charles S. Peirce, is, probably, the most obvious predecessor I have to admit, because of the explicit stress he put on the accidental phenomena in nature' (*Probabilistic Metaphysics*, p.11). It is known that Peirce was a prominent member of a Metaphysicians' Club, c.1870, and that this title oscillated between irony (detachment from Hegelianism) and arrogance (less concerned with the Universe, but especially with abstraction as a cognitive approach, the compromised discipline had to be rigorously reconsidered). While treating metaphysical and epistemological problems, Suppes argues for a probabilistic empiricism which does not manifest reductionist propensities (p.2). And his main thesis refers to scientific plurality: 'The languages of the different branches of science are diverging rather than converging as they become more technical' (p.121).

As a metaphysician, Suppes admits that the massive undertaking of science no longer needs a special protection against some erroneous philosophical conceptions. This ontological pluralism is held without reservations: 'we cannot have a reduction of the subject-matter to the ultimate physical entities because we do not know what those entities are' (p.123). At the same time, we ought to stress the plurality of methods or 'the vast difference in methodology of different parts of science' (p.124).

3. The metaphysics of A.I.

Perhaps the best known advocate, among specialists in computers and artificial intelligence (A.I.), of a 'metaphysical' approach, is Douglas R. Hofstadter (Indiana University). He maintains that anything that can be made by a rational being such as man, can also be made by a computer.

He attempts in his works to discredit any 'anthropocentristic' view of reason, because activities we consider as rational should not be located exclusively in the human mind or brain.

His arguments are based on an emphasis on the strong relationship between the success of the mathematical logicians (setting laws for reasoning as a structural process) and the creation of the A.I. systems (mechanisation of the reasoning process). 'There is a specific characteristic of our intelligence—the capacity of looking from outside at the task going to be fulfilled (*Gödel, Escher, Bach: An Eternal Golden Braid* p.37). Because computers are nothing but materialisations as mechanisms of formal systems, any meaningful interpretation of these systems makes an isomorphism evident between the formal system and a section of reality. But computers dispose of a language hierarchy, the machine's language being situated at a lower level. The user never works directly over the machine's states, the instructions being transmitted to it in the language on top of the hierarchy. The more we move to the top of the hierarchy, the more the languages 'consider' the things, putting them together in ever larger categories, avoiding the contact with the details and complexities of the lower levels.

At the same time there are characteristics of a complex phenomenon which are not located in a certain place of the system, nor are they given beforehand in the constitutive rules of the formal system. Epiphenomena, or emergent (global) aspects, seem to be analogical to the phenomena of consciousness. And the logical systems can offer different means to get to a flexibility which apparently specific to human thinking. 'Neurons are not similar to the symbols of a formal system but they can be considered as active symbols' (p.337).

J.R. Searle does not accept this thesis that the human mind could not be detached from its material, physiological support. His criticism is aimed especially at Hofstadter who also compiled (with D.C. Dennet) an anthology (*The Mind's I: Fantasies and Reflections on Self and Soul*, New York, Basic Books, 1981) dedicated to the evidently 'metaphysical' thesis of the human mind as an equivalent of a programming system for computers.

The 'forte' version of A.I., maintained by several researchers in the cognitive sciences, was summarised by Searle in three theses:

> 1. Mental states mould computer states and mental processes are

computing processes. As a proof, the said programs are systems of representation which 'update' themselves and sometimes even 'design' themselves.

2. Since mind is something abstract, there is no direct connection between it and the brain. Thus the neuro-physiology of the brain is irrelevant; any computer that is equipped with programs has a 'brain'.

3. The presence of mental states and capacities in computers can be demonstrated by the classical test devised by Turing ('Computing machinery and intelligence', review, *Mind, 1950*).

Searle shows that the computer disposes only of a syntax and not any semantics, handling formal symbols without awarding them any meaning. It is only the human brain that 'produces' mind and reason, and the way it generates such activities does not consist only in implementing a computing program. The problem of mind and rational behaviour continues to be 'the scandal of 20th century intellectual life'. Since the solution should supply a duplicate of the brain's specific causal capacities, cognitive science therefore proves to be a mere name for a whole family of research projects and nothing like a proper theory (J.R. Searle, 'The myth of the computer', *The New York Review of Books*, vol. XXIX No.7, April 29th 1982).

Still, these few works, considered as clues or possible sources for a construction yet to be achieved, give us the opportunity to draw a first conclusion: cognitive sciences, including A.I. researches, do not repudiate metaphysics. It is true that A.I. is solidary from the beginning, not only with the positivist project of eliminating metaphysics, but also with the spiritualist view of the existence of the supreme functions of man (Jean-Michel Besnier, 'Les sciences cognitives, ont-elles raison de l'âme?', *Ésprit*, no. 161, May 1990, p.121). Anyway, the new (computer) metaphysics may not be considered without being based on other researches belonging to cognitive sciences in the widest sense.

4. Metaphysics and cognitive sciences

Who knows when we are going to have a greater chance of reaching an adequate view of the 'computer world', without reducing the existence of these artefacts to the lowest level of reality and their laws to 'absolutely' fundamental ones? Certainly, this metaphysical project cannot avoid the epistemological and axiological problems of the field. Further, on, we shall call upon the direct suggestions of scientists and engage more seriously in the discussion of the problems of the opportunity for, and the value of, the whole project.

Let us take as a starting point the conclusion of a minute investigation:

> The party is far from being gained by cognitive sciences and the philosopher—be he spiritualist or not, but, anyway, a little indulgent —is happy to notice that his old problems defy the scholars' ingenuity too much for them to be put under the heading of metaphysical antiquities (*ibid*, p.132).

The investigation had started in an attitude favourable towards the new sciences whose domain seems to be unlimited.

> Strictly speaking, nothing should be left out of the cognitive sciences; so the specialists should not hesitate to re-examine the problems which philosophy has declared insoluble (p.118).

Cognitive sciences are characterised by the interdisciplinary effort of evidencing the mechanisms of knowledge with the declared intention of modelling these mechanisms on the computer. Without analysing the number or specific relevance of co-operative disciplines in this effort, we bear in mind that artificial intelligence (A.I.) is the nucleus of researches involving both cognitive elements and technology, meant to ensure the capacity for computers to achieve operations which, being performed by man, should be called intelligent. After all, programs are searched to have the quality of granting the computer the performance of such intelligent operations in processing knowledge

Doubtless, research on the human mind which uses the concept of artificial intelligence is reducible, in the last analysis, to looking for certain programming systems for computers. Difficult to admit within this framework is the tendency of some authors to maintain the autonomy of mental phenomena, processes and states, in the sense of reproducibility independent of man's brain and body.

In a previous stage, the optimism of researchers in the field of artificial intelligence was related to the tasks of modelling and simulating cognitive processes on the computer and the processing of 'information', respectively. In the present stage, after the discussions related to designing the fifth generation of computers, the optimism concerns the processing of 'knowledge', that is to say, a more specific approach to the problems which appeals to 'their meaning, their contents, even their empirical aspects'. Yet, the older question of H. Dreyfus, about the two complementary aspects, is still valid: (a) to what extent is man following certain *formal rules* in his activity of processing data, so that this activity can be identified with the processing of data in the computer?; (b) to what extent can *human activity* be described by appealing to a formalism capable of implementation on a computer?

It is also Dreyfus who formulated four hypotheses associated with the attempts meant to obtain artificial 'reason' or 'intellect' ('What computers can't do', 1972). The suggestions of his analysis still retain their importance. The *biological* supposition gives credit to the analogy between the functioning of the human brain and processes in the computer. But the brain achieves a complex combination of numerical and analogical operations. Furthermore, the biology of the nervous system rejects this supposition. The

brain cannot be *identified* with a computer because 'the nervous system works as a closed system, without inputs and outputs'. Therefore, its cognitive functions do not reflect its organisation. It does not receive but only imposes its own information to the environment (F. Varela, *Autonomie et connaissance*, Seuil, 1989, p.145). The closed system is in a way a version of the *monad* imagined by Leibniz.

The *psychological* supposition admits the analogy between the psychic events and the processing of information in a computer: in both cases a string of rules seem to be applied in a determined order. It is true that there is a stability in the biological system which is often missing in the simple physical systems, so one cannot maintain 'any strong theses about reducing mental events to brain events' (P. Suppes, *Probabilistic Metaphysics*, p.132), even less so in relation to the processes in a computer.

The *epistemological* supposition gives credit to the governance of rational behaviour by a system of formalisable rules. But lately, the behaviourist thesis has become to be considered unacceptable. 'We will not reduce mental events to a characterisation in mere behaviouristic terms' (*ibid.* p.133).

Likewise, the *ontological* supposition asks us to accept that the world can be described in terms of discrete, atomic facts, transcribed in independent logical sentences, whereas in reality our knowledge is related to a continuum which is virtually infinite. Man starts from the whole in order to get to details, and details are interpreted according to the whole which gives them significance. The machine works quite the other way round: it always starts from details in order to get to the whole.

Besides, the role of the human body in the structuring of knowledge must not be neglected. *Body* is more than a source of information about the environment: it is a source of needs from which develop impulses that result in the advancement of knowledge. *Intellect* must not be considered as an addition to the body, as an entity that could be obtained independently from the existence of the human body. A robot which would imitate the human body should have organs just like human ones. In a way, man belongs to the world, which is his country or home. No machine is in that situation. J. Haugeland insists that computers do not achieve the activities of thinking because, after all, it is quite unimportant to them if they are thinking or not ('Understanding natural languages, *Journal of Philosophy*, 11, 1979). The specific exercise of thinking represents an elevated human satisfaction.

Beyond the optimistic initiatives of theoreticians who believe in the possibility of making computer programs capable of 'initiating' any type of intellectual activity, and beyond the severe criticism which qualifies them as alchemy or mythology, we find the conception of M. Bunge (*op. cit.* p.371). In his view, there are problems in the philosophy of technology, especially that of general or conceptual technology, which motivate a comprehensive programme of research, including not only semantics or epistemology, but also ontology as well as the theory of value. His manner of treating problems of technology in an epistemological and ontological perspective, and then in an ethical one, appeals to a new style of theorising in the fields traditionally assigned to metaphysics, which promises to bring into scientific research some domains only speculatively conceptualised so far.

5. Some conclusions

In order to award a more systematic character to the results of our discussion on 'computer metaphysics', let us try to draw some answers to a few problems formulated by M. Bunge. Here are some questions which belong to ontology and arise only because of the existence of computers.

The fundamental question seems to be the following: 'Do the man-machine systems belong to an *ontic level* proper?' This question contains, actually, the answer to another one which is first on the author's list: 'Do the artefacts possess particularities which are *not* common to natural objects, except that they were *designed* and *executed* by human beings or by other artefacts controlled by human beings?' We agree with Dreyfus' assertion that 'computers are instruments which intensify the human mind' ('What computers can't do?' p.274) But the detailed solution to the problem would require the analysis of the way we could characterise man's part and the machine's part, respectively, in this collaboration.

For the second question of great interest, 'Have the artefacts and the man-machine combinations their *own laws*, different from those studied by fundamental science?', the answer seems to be conditioned by another: 'Is there more than a mere analogy between the satisfactory operation of an artefact and the health of an organism?' Certainly, the concept of reliability does not overlap the concept of health in general for an organism and even less for a human being. Here is the paradoxical opinion of the Romanian metaphysician-poet Lucian Blaga: 'The disease which urges a creation is a higher-order health'.

The third significant question: 'Can machines have their own *will*?' seems to be a synthesis of two others. They are: 'Is it possible to design a machine capable of putting or avoiding original problems and to do right or wrong on its own initiative?' and 'Can anybody say that artefacts are *embodiments* or *materialisations* of ideals?' It is true that an affirmative answer to the last question entails a favourable answer also to the previous question. Taking an advantageous example for A.I. research, Paul Teller showed that, though machines will become in time an auxiliary, and that we could not do

without them in mathematics, mathematics after all 'remains a specifically human activity ('Computer proof', *The Journal of Philosophy*, 12/1980).

From project to reality there is a distance which cannot be underestimated, not even in the case of the most tempting hopes for the future. In other words, the seductive power of a project does not warrant its performance as such. The ideal always pays a serious tribute to the process of *materialisation* (generally appreciated as a decay) no matter how spectacular the effect could be for the technician. Lucian Blaga generously appreciated that 'technology dissolves the fairy-tale in our lives only to the extent that it is itself a fairy-tale come true'.

(To be continued)

Department of Philosophy, Logic, Psychology and Sociology
Polytechnic University of Bucharest.

Notes:

1. A.A. Moles, *Art et ordinateur*, Casterman, 1971.
2. *Théorie générale des systèmes*, Paris, Dunod, 1973.
3. *Probabilistic Metaphysics*, Oxford and New York, Basil Blackwell, 1984.
4. 'The philosophy of technology' in (ed) F. Suppe and P.D. Asquith, *Philosophy and Social Action*,2/1977.
5. *Éloge des intellectuels*, Grasset, Paris, 1987, p.77.
6. New York, A.A. Knopf, 1972.
7. 'The philosophy of technology',1977.

Continued from p.35

Nagy, Endre (1992), 'Introductory Remarks to the Study of the Theory of Law in Michael Polanyi's thought' *Polanyiana* 1992/1-2, pp.106-117.

Prosch, Harry (1986), *Michael Polanyi: A Critical Exposition*. Albany: State U. of New York Press.

Sahlins, Marshall D. (1965), 'On the sociology of primitive exchange', M. Banton, ed., *Monographs of the Association of Social Anthropologists*, No. 1. London: Tavistock and New York: Praeger, pp.139-236.

Scott, Drusilla (1985), *Everyman Revived: The Common Sense of Michael Polanyi*, Lewes: The Book Guild.

Szelényi, Iván (1987), *Socialist Entrepreneurs: Embourgeoisement in Rural Hungary*. Oxford: Polity Press.

Szelényi, Iván (1990), 'Új osztály, állam, politika' ('New class, state, politics'). *Replika* 1990/2.

Staniszkis, Jadwiga (1991), *The Dynamics of Breakthrough in Eastern Europe: The Polish Experience*. Berkeley: University of California Press.

Weber, Max (1964), *Wirtschaft und Gesellschaft: Grundriss der Verstehenden Soziologie*. Berlin: Kiepenhauer.

CRITICAL NOTICE:

A DEFENCE OF REALISM

Julian Ward

Marjorie Grene: *A Philosophical Testament*

Chicago and La Salle, Illinois: Open Court, 1995. Pbk £17.50. ISBN 0-8126-9287-X 1.

1. Marjorie Grene

This book is what the title says it is. It is not a philosophy book with carefully ordered arguments but a summarisation of a lifetime's work in philosophy, especially in modern philosophy from Descartes to Kant, epistemology and the philosophy of biology. Marjorie Grene (née Glicksman) was born in 1910, the daughter of a lecturer in English at the University of Wisconsin. After university training in biology in the USA she studied at the feet of Heidegger in Freiburg in 1931 and then under Karl Jaspers at Heidelberg in 1932. She then obtained her M.A. and Ph.D. at Radcliffe College with a dissertation on existentialism, after which she was told, 'Good-bye, you're a bright little girl, but nobody gives jobs to women in philosophy'. However, she spent 1935-36 in Denmark studying Kierkegaard's ideas. Then after 129 applications she obtained a post at Monticello College in Illinois only after the Principal was falsely assured by a sponsor that she was not a Jewess. From there, in 1937, she moved to a part-time post at the University of Chicago (the first woman appointed to the Department of Philosophy), where she attended Rudolf Carnap's research seminar, which cured her of a juvenile dalliance with logical positivism. She remained a lecturer at Chicago until fired in 1944. She then raised a family while engaged in farming in Illinois from 1944 to 1952 and then in Ireland. In 1958 she returned to academic life as the Senior Research Fellow at the Institute of Education at Leeds, where she examined the presuppositions implied in the teaching of biology. She taught modern philosophy from Descartes to Kant at Leeds in 1959-60 and then mainly Greek philosophy at Queen's University, Belfast (where she was debarred from teaching in the philosophy of science) from 1960 to 1965. She returned to the USA in 1965 and taught modern French and German philosophy and philosophy of biology at the University of California at Davis, where she retired as Professor in 1978. She has directed summer schools and been a visiting lecturer in philosophy at a dozen universities.

Marjorie Grene was employed as Michael Polanyi's personal research assistant in 1957-58. At the beginning of *Personal Knowledge* he acknowledges his debt to her.

> This work owes much to Marjorie Grene. The moment we first talked about it in Chicago in 1950 she seemed to have guessed my whole purpose, and ever since she has never ceased to help its pursuit. Setting aside her own work as a philosopher, she has devoted herself for years to the present enquiry. Our discussions have catalysed its progress at every stage and there is hardly a page that has not benefited from her criticism. She has a share in anything I may have achieved here' (p.ix).

She wrote 'The Logic of Biology' for *The Logic of Personal Knowledge: Essays Presented to Michael Polanyi on his Seventieth Birthday, 11th March 1962.* Grene compiled the extensive index of *Personal Knowledge* and edited a collection of essays by Polanyi under the title of *Knowing and Being* (University of Chicago Press, 1969). Grene came to realise that there was a convergence of Polanyi's thought with that of Merleau-Ponty, whose work she read for the first time in 1961. Grene assisted Polanyi in his work through much of the 1960s.

Grene has written twelve books, on subjects such as Aristotle, Descartes, Heidegger and existentialism, and the philosophy of biology, and about eighty technical articles, some of which are collected in *The Understanding of Nature: Essays in Philosophy of Biology* (Reidel, 1974) and *Philosophy in and out of Europe* (University of California Press, 1976). She has also edited ten other books. To mark her seventy-fifth birthday her colleagues Alan Donagan, Anthony Perovich and Michael Wedin edited *Human Nature and Natural Knowledge* (Reidel, 1986). It includes a personal reminiscence 'In and On Friendship' by Grene, in which she speaks appreciatively of the friendship of Richard Rorty, Alasdair MacIntyre and Rom Harré. Alarmingly, we read pejorative comments about Polanyi's thought. After saying that she has been seeking in her philosophising a 'true counter-Cartesian reform in thinking' that will lead to 'the kind of other-than-linguistic, other-than-scientistic turn' in consideration of what it is to be human, she comments,

> The point is: This has nothing special to do with the biological sciences but with the simple fact that we are *alive*. Polanyi's *Personal Knowledge*, of course, was an effort in the same direction, but I'm afraid that, as Körner once remarked, his rhetoric was hopeless, and in the end Polanyi, too, was too perversely subjectivistic' (p. 358).

However, she does allow that 'Po-

lanyi did try to ground knowledge in life, in human existence, and certainly with more subtlety and insight than selfish-gene reductivists have shown' (ibid). It is therefore with some relief we find that Grene is committed to Polanyian themes in *A Philosophical Testament*, although in it she does accuse Polanyi of a 'gross misunderstanding of evolutionary theory' (p.171). After being asked to contribute a paper (subsequently entitled 'The Subjective and the Personal') to the Centennial Conference on Polanyi at Kent State University, Ohio, 11th-13th April 1991, she re-read *Personal Knowledge* after a break of twenty years from it. In *A Philosophical Testament* she acknowledges Part Three of Polanyi's *magnum opus* as 'the best authority for the rather uneasy position I have been trying to describe' (p.171) but she is no longer able to give her unqualified approval to Part Four.

2. The primacy of perception

Chapter 1, entitled 'Knowledge, Belief, and Perception', is based on a lecture that Grene gave in 1978 at Tulane University, which subsequently conferred on her an Honorary Doctor of Humane Letters in 1980. In this lecture she discusses four different views on the relation between knowledge and belief. Two views affirm that knowledge and belief are different in kind and two do not. The first pair expresses the view most dominant in the western tradition from Plato onwards that knowledge assumes indefeasible certainty and belief evanescent opinion. The first of these two views is that certain knowledge is possible, as in Cartesianism and Husserlian phenomenology, and the second view denies that certain knowledge is possible, as in varieties of traditional scepticism and the modern relativising of science. But Grene protests that we can claim true everyday knowledge which we believe to be true. Thus she turns to two views that correlate belief and knowledge. The common claim has been that knowledge is justified true belief and, for a while, Grene affirmed this view herself. However, although considering the Gettier counter examples as trivial, she notes that 'justification' is rooted in combinations of perception, authority and traditions, as Merleau-Ponty and Polanyi so emphasised. But the real problem is how we know what is 'true'.

> What keeps eluding us is a way to check our beliefs against reality and find out, once and for all, whether they *are* true. It's been tried, both logically and quasi-psychologically, to devise a means for such a comparison, but it just hasn't worked. For we can't get *outside* all our beliefs at once in order to check them. That is the paradox Polanyi was trying to respond to in *Personal Knowledge* (p.17).

It is through our experiences that we shape and re-shape our beliefs, in accordance with the appropriate procedures and respected traditions of the relevant disciplines, and thus make contact with reality through submission to the values of honesty and integrity. Knowledge, then, is fallibly justified belief that is asserted, as Polanyi said, with universal intent, that is, with the confidence that anyone with the same evidence and the same standard of objectivity would make the same claim.

When it comes to knowledge by perception Grene notes that in ancients like Plato and Aristotle and moderns like Descartes, Spinoza and Husserl perceptual knowledge was likened to a form of intellectual seeing. But the majority of English philosophers have tended to forms of phenomenalism in which objects are the mind's constructs from sensations or sense-data, with Humean scepticism hovering in the shadows. But isolated 'sense-data' are the illusions of philosophers and can be no basis for a claim to knowledge. Accepting help from Aristotle one ought to realise that perception is 'awareness-through-interaction' (p.23) along with integration of the resources of our faculties. Better help is given by Merleau-Ponty's emphasis on the primacy of the perception of objects and Polanyi's doctrine of tacit knowing, for the latter's transformation of epistemology, by recognition of commitment with universal intent and our embodiment in a hierarchical universe, to 'ultrabiology' shows how his thought converges with that of his French contemporary. Perception of objects is thus the most primitive form of knowledge and the most pervasive, being foundational to our belief systems. It is surprising that Grene does not state what seems clearly implied, as indicated on p.130, that our persons assimilate sensations adverbially, e.g, 'I see that red box' means 'I see that object redly and cubely', as argued by R. J. Hirst in *The Problems of Perception* (Allen & Unwin, 1959) It is the primacy of perception as lived individuals in a changing physical and cultural world that guards against subjectivism and scepticism and assures us of living in a mind-independent reality populated by organisms that are irreducible.

From 1980 onwards Grene came to see the value of the ecological theory of perception of the eminent psychologist James J. Gibson, who died in 1979. Here again Grene finds a striking convergence with the thought of Merleau-Ponty. As opposed to the false modern notion that we apprehend directly atomistic sensations, asserted by Descartes and the British empiricists and still perpetuated in modern psychology texts by Richard Gregory and others, we apprehend objects, grasped through our perceptual apparatus ('grasped' is a tactual metaphor, as opposed to the visionary one of 'seeing' sensations as used by the phenomenalists), which is the product of our evolutionary history. Objects are identified by means of the invariants in the flowing array of sensations resulting from the

direct interaction of the organism with the surrounding objects. Perception is not a passive reception but an activity of discrimination and differentiation, a bodily achievement of the organism. For the phenomenalist Gregory perceptions are hypotheses, imaginative constructions or even fictions: passive sensations plus judgement. For Gibson perceptions are direct, active graspings of the world in which the senses actively reach out into the environment, e.g. in the apprehension of a three-dimensional world, to grasp what is significant for the organism: the 'affordances' that the environment offers it. 'It is not invariants as such that are perceived; they are the regularities that specify the objects and events whose affordances are directly perceived, whether as predator, as prey, as plaything, as potential mate' (p. 143). As Wittgenstein said, all seeing is seeing-as.

The artefacts of culture and science are not merely artificial additions to our environment but become part of that environment and also additional means by which we contact reality. The interaction between language and the world deepens our perceptual capacity and scientific research makes explicit invariants that are not usually observed in the everyday perception of affordances by making them, in Polanyi's terminology, focal rather than leaving them subsidiary. Likewise, the representational artist reconstructs invariants so that they become new affordances. The threefold structure of perception—information conveyed through invariants, significance as affordances, and significant objects of knowledge—is reduplicated in science and art.

3. From Kant to Merleau-Ponty

Grene has a great admiration for Kant but no patience with Hegel who produced a 'superteutonic professional perversion' (p. 30). 'Mind can be swollen up to Hegelian Spirit if one is willing to fool oneself sufficiently' (p.49). For Grene philosophising today should begin from Kant. Bertrand Russell's sarcastic comment about the obscurity of the *Critique of Pure Reason* was an unforgivable blasphemy (p.30). Grene's 'chronic fascination' (p.32) with the Transcendental Analytic section of that work leads to a brief exposition of its argument, a summarisation of a fuller exposition that had a central part in *The Knower and the Known* (Basic Books, 1966), written in the years 1961 to 1963 when she was at Queen's University. Kant was seeking to define the conditions necessary for us to be knowers and to rebuff pretensions to a metaphysical ontology. Grene believes she can take the argument of the Analytic and transform its context. Kant's transcendental unity of apperception becomes the living, embodied human being. Objects of knowledge are not Kantian 'appearances' but the real things-in-themselves that are significant for us. And the human being is not discarnate reason in temporal activity but a part of the nature that man's sciences investigate. Kant's argument, thus transmuted, allow us to assert three things: 'the active role of the knower in making experience objective, the inexhaustibility of the known, and the indissoluble connection between knower and known' (p.44).

Kant maintained that we are forced to choose between metaphysical realism combined with empirical idealism (cf. Descartes, Locke and its logical result, the scepticism of Hume and no basis for saying that science tells us about mind-independent reality) and his own transcendental idealism that he regarded as consistent with empirical realism. It is the genius of Grene's Polanyian transformation that we can hold to both empirical realism *and* metaphysical realism, which Kant considered to be an impossibility. Gilbert Ryle or J. L. Austin should have achieved this but were diverted from this goal by the then current obsession that philosophy only examines language usage. Wittgenstein was the best hope with his determination to escape the strictures of empiricism by emphasising the forms of life that are skills expressed in families of language games, and his views have some correspondence with Polanyi's thought. But Wittgenstein regarded the production of philosophical doctrines as anathema.

For Grene creative contributions have come from Helmuth Plessner's philosophical anthropology, Maurice Merleau-Ponty's phenomenological analysis of living and perceiving and Martin Heidegger's being-in-the-world, but, for Grene, the doctrines of the latter are, in retrospect, seen to be of very limited value. True, Heidegger overcomes Cartesian dualism but fails to give adequate emphasis to our 'livingness' and too much to death and German nationalism (p.17). For Grene Heidegger's project fails because of the confusions of his punning style, the disembodying of *Dasein*, and the glorifying of the authentic life in terms of Nazi-style heroism (pp.75-79). Grene spent immense labour unravelling the obscurities of *Being and Time*, but doesn't see now 'why sensible people .... should spend their time and effort in deciphering that crabbed document, let alone the rest of the *corpus*' (p.76). Sartre has nothing of value to add as his brilliant work proceeds from inadequate premises to 'an impassable dead end' (p.79). It was Merleau-Ponty who gave us 'the most effective account so far of what it is to be in the world' (p.80).

According to Merleau-Ponty, reflecting on vision, our prime sense, helps to overcome the subject-object dichotomy that 'has long distorted our philosophical reflections.' Explication of our bodily existence as spatial, motoric, grasping, sexual, social, gesturing, linguistic beings enable us to grasp that depth of reality that shows that

we are real people in a real world, which the abstractions of empiricism inevitably fail to do. Merleau-Ponty's explication of language constitutes 'the most effective refutation of Cartesian dualism in our literature' (p.82). My 'self' is no unextended thinking substance, but is a spatial history, partially constrained and partially transcending its constraints. Whereas Husserl bracketed out existence and ultimately lapsed into a convoluted Cartesianism, Merleau-Ponty keeps his philosophising in the real world.

4. Biology and persons

At the instigation of Polanyi, when helping him with the writing of *Personal Knowledge*, Grene read the works of heretical scientists who dissented from Darwinian and Neo-Darwinian evolution, such as Berg, Osborn, Schindewolf, Lillie, Dalcq, Vandel and Willis. In *Approaches to Philosophical Biology* (1968) she expounded the views of Adolf Portmann, Helmuth Plessner, F.J.J. Buytendijk, Erwin W. Straus and Kurt Goldstein. In *The Knower and the Known*, first published in 1966, Grene argued that organisms as a whole can only be understood as irreducible entities requiring a teleological explanation (p.239), with the implication that the course of evolution that has led to man is, as Polanyi maintained, teleological, 'a problem set to us'. But in the Preface to the 1974 edition Grene confesses 'to having overestimated the relevance of teleological thought for evolutionary theory—which I would now place at zero.' But she has permanently maintained, as did Polanyi, that one cannot undertake biology and expound evolutionary doctrine without presuming that organisms are basic irreducible realities in terms of which genetic mutations and adaptation to changing environments are described (*A Philosophical Testament*, pp. 90, 106).

This leads on to a discussion of the paradox that 'chance' and 'necessity' seem to become interchangeable terms in evolutionary explanation, if cosmic teleology is denied. Belief in creation by an omnipotent God must imply no possibility of contingency and freedom. Grene believes that no-one has yet produced a satisfactory definition of 'cause' (p.103). She concludes, 'Darwinism explains the origin of means-end phenomena by old-fashioned when-then causality. The teleological language belongs to the phenomena to be explained, not to its explanation' (p.104). Talk of progressive evolution is delusory.

Grene defines the human being as 'a biological individual capable of becoming a responsible person through participation in (or as one unique expression of) a culture' (p.107). Evolutionary epistemology rightly sees the *a priori* categories of the mind as the product of the evolutionary struggle for existence, but then wrongly dismisses the truly epistemological questions of the nature of knowledge and its justification (p.109) and questions of morality (p.111). Ethics is neither conditioned by our evolutionary past nor is to be seen as warring against it. It pertains to our cultural and historical situation within the constraints of nature. The fact of evolution, however, rules out the possibility of the Christian God but a religious awareness of the wonder of nature is still possible (p.112). Grene finds the Philo of Hume's *Dialogues* 'the safest guide in the philosophy of religion.'

For Grene the most significant epistemological consequence of an evolutionary metaphysic is 'an unwavering and unrepentant realism' (p.110). We are real people in a mind-independent real world: what Grene once called 'comprehensive realism' but no longer does to avoid confusion. Thought experiments about brains in vats, etc., are anathema to her. This leads Grene to undertake a critique of Arthur Fine's anti-realist operationalism. According to Fine, scientific theories can only be said to state the truth about mind-independent external reality, as scientific realism maintains they do, if we have a non-scientific justification of scientific theories more powerful than scientific procedures, but that kind of justification we do not possess. Scientific realism, Fine maintains, is supposed to hold to a vestige of the logical positivistic notion that there can be a strict demarcation between the subjective processes of discovery and the act of objective rational justification, but this is not true to the actual work of scientists. Moreover, a defence of realism, Fine thinks, will suppose that anything science says can be said explicitly, objectively and impersonally, with low-level statements grounded in infallible, explicit, high-level foundational statements. Fine knows that none of these things are true of science, so *a fortiori* scientific realism cannot be true. But, as Grene points out, Polanyi persuasively showed that, in science, discovery and justification are homologous, the statements of science are grounded in tacit knowing and cannot be wholly explicit, and science is rooted in real knowing persons, not impersonal factual statements. Due to these factors scientific realism can be justified.

Sociology of science tends to see science as a pragmatic, cultural or linguistic product with no justifiable claim to truth. But Polanyi shows that the practice of science is undertaken in a social tradition that bears on and is constrained by external reality and the language of science is not a veil that hides reality but is the pluriform means of access into that reality. Mainstream philosophy of science has lurched from one delusory dream (operationalism) to another (sociology of science) in this century and Polanyi can awaken us from such slumbers.

As against the sociobiologist E.O. Wilson who thought that human language was merely a developed version of animal signalling, Grene argues that it constructs a cultural world in which we live and in

which words derive their meaning. Human beings are subject to conflicting tensions within their natural and social worlds and their resolution comes by commitments to future behaviour that are expressed in kinship systems, language and rites. Moreover, cultural artefacts such as homes are extensions of ourselves, and also creators of ourselves, and language and rites are the symbolic representations of these quasi-independent creative powers on which society is founded. And these symbols have the power to create the realities they depict. But the activities of kinship and rites are promissory symbols with respect to future behaviour and our language is rooted in them. Hence metaphor and poetry are more fundamental for our language-bound lives than literal speech (p.166). But we need not fall into cultural relativism, for we can distance ourselves from our culture and reality can amend our fallible beliefs. As Polanyi emphasised, we live by the paradox of self-set standards to whose authority we submit in order to voice our claims with universal intent.

Grene concludes with a discussion of what it is to be a person. For her philosophers of mind have too often engaged in abstract problems disengaged from real people in the real world. A person could be defined as one who is able to make a responsible choice between alternatives, and could have chosen differently. Grene thinks that arbitrary (negative) freedom is unreal. Choice involves preferences and motives and responsible choice involves the (positive) freedom of choosing what we see to be right in a particular situation. There is a conformity between moral choice and commitments to knowledge claims with universal intent. Some have emphasised self-consciousness as the prime mark of personhood, but Grene's hostility to any form of Cartesianism means that, for her, 'the individual is not anything inward, but something like an ordering principle, a centre of responsibility to principles, or ends, or causes, something beyond myself to which I *owe* allegiance' (p.178). The sincere person gives himself in his knowing unconditionally and responsibly to the demand for truth and the truly authentic life lies in giving oneself to that which is seen to be totally worthy of one's self-giving. But Grene finally confesses that she has 'no well-worked-out answer' as to what the truly authentic life is (p.188).

5. Philosophers and philosophy

In the course of her meditations Grene does not hesitate to make acerbic comments on philosophical styles that she sees as academic in-talk unrelated to the real world or devoid of lasting value. Some of Husserl's thought is 'insidiously subjective' (p.86). Heidegger's thought is 'vitiated by its fanatical nationalism' (p.17), the 'radical evil' of 'sick Germanism' (p.68), and a 'defective' exposition of being-in-the-world (p.69), which is 'fundamentally unacceptable' (p.75). His later writings are 'doing nothing but playing with words in a totally unconscionable manner' (p.76). Dewey and Mead are 'dim and dated' (p.54). Moore, Broad, Stout and Ward were 'dreary people . . . monosyllabic Englishmen' (p.53). Carnap produced 'sheer nonsense' (p.33) and the logical positivists 'were struggling with a pseudo-problem of their own ... based on such badly mistaken premises that [their discourse] could only be trivial' (p.116). Ayer was 'a pure posturer, skilful, in a glib way, at adopting and adapting the fashion of any given moment' (p.57). Oxford ordinary language philosophy was a superficial and surreal form of 'academic parlor games' (p.55). Unreal thought experiments are 'professional frivolity' (p.37) and obsession with language unrelated to reality takes us into 'a sort of Peter Pan world' (p.56). Grene will have no truck with dogmatic social Darwinism (p.90), the 'nonsense' of scepticism about the external world and the so-called problem of other minds (p.83), 'Freudian nonsense' (p.82), 'thoroughly unphilosophical' naive evolutionary epistemology (pp. 42, 109, 145), neuro-philosophy (p.155), inadequate sociobiology (p.156) and the 'nonsense-mongers' of social constructivists, deconstructionists and pragmatists (p.113).

Critics may take it that such pejorative terms show that irascibility has overcome scholarship in an aged woman. But for Grene reductionist analysis, divorced from human life as it is lived, leads to academic abstractions of no more value than word games. Perhaps the profession has long needed a lady philosopher who will tell male philosophers when they are producing mere verbiage. True, Kant, Wittgenstein and Ayer sought to do this, but then they didn't seem able to recognise *human* beings when they saw them. Perhaps women have greater insight into what it is to be human and so can make better philosophers. Grene's book is full of suggestive ideas that could prove fertile in future philosophical research.

For Grene the greatest iniquity in philosophy is Cartesian dualism and she agrees with Heidegger that 'consciousness' should be banished from philosophical discourse. This contrasts with Polanyi for whom many modern thinkers were wicked reductionists in denying the reality of consciousness. Certainly, the titles of some recent books include the dreaded word: *Conscious Experience* edited by Thomas Metzinger, *Ten Problems of Consciousness* by Michael Tye, and *Stairways to the Mind: The Controversial New Science of Consciousness* by Alwyn Scott, not to mention the international conference *Towards a Science of Consciousness* at Tucson, Arizona, in April 1996 with its fifty renowned speakers. Does the anti-reductionist Grene betray a vital insight of Polanyi in not carrying through his programme for

recognising the full range of an ontological hierarchy in human beings? Polanyi's vision led him to affirm a theistic God. I can't help feeling that it is Grene's agnosticism that makes her halt at affirming, with Kenny, that mind is no more than a set of abilities to behave in certain ways (p.158). But what of the values of justice, honesty and beauty by which we live and by which science is accomplished, as Grene avers? Are these values cultural constructs and, if so, what are they? Polanyi's vision gives a better answer if we also say, as Christians do, that man is made in the image of God. But then they say that, not because of philosophical argument, but because of the revelation of God in Jesus of Nazareth, for he is, as Paul says, the image of the invisible God (Colossians 1:15). Human freedom and biological evolution are not then to be seen as incompatible with belief in a Creator God, but then they are not to be seen as the whole story either.

BOOK REVIEWS

J. Misiek (ed)
The Problem of Rationality in Science and its Philosophy
Dordrecht, Kluwer (Boston Studies in the Philosophy of Science, Vol. 160), 1995; xii + 272pp.
ISBN 0-7923-2925-2

This volume contains 20 papers from two conferences, held in Poland, 'Popper, Polanyi and the Notion of Rationality' (1988) and 'The Aim and Rationality of Science' (1989). I shall comment only on the longer and less technical papers.

F. D'Agostino ('A deontological approach to the rationality of science') argues that rationality in general, and therefore that of science, does not arise from the pursuit of some specific aim, but from resolving disagreements in accordance with the rules of liberal conversation which enable one to express one's commitment as a rational agent. Yet, it strikes me, those rules arise from the commitment to the aim of seeking the truth and stating one's findings.

D. M. Armstrong ('What makes induction rational?') holds that it would be foolish to reject induction but that it is hard to see exactly why inductive arguments are sound. He argues that observed regularities are best explained by hypotheses of strong laws, holding between properties construed as universals, hypotheses which in turn entail conclusions about the unobserved. But perhaps this is not so much a justification of induction as an elaboration of what inductive arguments presuppose.

A. Cattani ('Popper, Polanyi and the notion of rationality') defends Polanyi against accusations of irrationalism by arguing that what he proposes is, in the terms of Aristotle's rhetorics, an 'enthymematic rationality', i.e., a reasoning that is not fully expressed and which is only probable. But that is to imply that any unexpressed elements can be stated, whereas Polanyi's argument is that many of them never can be. (It is significant that Cattani does not mention the *from-to* structure of tacit *integration* but only 'tacit knowledge'.) Cattani implies that in historical hindsight the missing steps in scientific discovery and explanation can be explicated, but Polanyi's thesis is that there is always a necessary limit to what can be said by anyone.

P.M. Churchland ('On the nature of explanation: a PDP [Parallel Distributed Processing] approach') proposes to explain explanation without reference to its linguistic expression and by analogy with perceptual recognition. He rejects Hempel's idea that explanation consists in bringing a particular fact under a covering law, because it cannot account for the immediacy and inarticulateness of both human and animal apprehension, in favour of computer simulations of the 'parallel distributed processing' whereby the brain learns to perform discriminations (e.g. between sonar echoes from mines and from rocks) by creating prototypes for each. Connected to these perceptual 'inputs' are behaviourial 'outputs'. But Churchland rejects the Behavourist S-R model because in a brain the prototypes are subdivided and many of the inputs come from within the brain itself (either from other neural networks or in reverse flow from within the same network) so that there is no simple and easily predictable correlation between sensory input and subsequent behaviour. It is these 'prototype activation vectors' which, when activated, constitute recognition and concurrent understanding of one's situation. Churchland claims that this model explains the different degrees of understanding and types of explanation by reference to the differing richness or character of the prototypes involved. But when he gives examples of the latter, we read no more of events in the brain but of the situations, problems and phenomena requiring explanation and their relation to other ones. That shows that Churchland has run together an explanation of *explanation*, in terms of our use of prototypes, which does seem superior to the usual models of explicit logic (inductive or deductive), with an explanation, in terms of neural networks, of the physiological processes which act as the vehicles of our thoughts. Likewise in attempting to account for the merits of one explanation over others, he claims to do this by reference to neural networks and *not* to notions of reference, truth, consistency, entailment and the like. Yet he moves back from the latter to the former in what Polanyi called a 'pseudo-substitution'. Thus he invokes

'pragmatic' and 'statistical' considerations to explain why one prototype is appropriate and another inappropriate. But those notions themselves implicity trade upon the very notions which they are supposed to replace, as when he says that one prototype *misrepresents* the situation. In other words, it is the *content* of the prototypes that matters. Hence he arrives, in effect, at a curious notion of *impersonal* tacit knowing, performed by *brains*, not people or animals, and contained within neural networks.

J. Misiek ('Personal Rationality') distinguishes between: $Popper_1$, a member of the Vienna Circle who likes 'logic' and logical analysis, fears psychologism, desires objective knowledge, apparently without a knower and residing in books, and thinks with Wittgenstein that 'there is no riddle'; and $Popper_2$, who is something of a Polanyian, considers preoccupation with logic to be unproductive, bodly explores problems involving a knowing subject, takes science to be rational and moral procedure of cognitive aims, acts and decisions, aiming at the description and explanation of reality, and holds theories to be testable and falsifiable only as we decide to allow evidence to count against them. Both Poppers need each other: $Popper_1$ needs $Popper_2$ to avow the values which $Popper_1$ tacitly upholds; and $Popper_2$ needs $Popper_1$ as a point of departure for more adventurous enterprises and to protect himself from criticism by neo-positivists. Popper, concludes Misiek, has failed sufficiently to free himself from neo-positivism, while philosophy of science should now follow $Popper_2$ and Polanyi: and focus upon research, the seeking of truth, and the personal and non-mechanical principles which govern it.

In a second paper, 'Assessment of theories', Misiek proposes to broaden Popper's and Lakatos' accounts of science on Polanyian lines, by distinguishing a higher, more demanding and thus élite rationality responsible for creating and extending scientific language and heuristic principles, as shown, for example, in Einstein's formulation and proposal of Special Relativity, and its acceptance by some but not all leading scientists, when there was little experimental evidence in its favour.

A. Musgrave ('Realism and idealisation') convincingly refutes Nancy Cartwright's claim that there are no exceptionless laws. Likewise, Z. Piatek ('Is evolutionism a scientific theory?') rebuts Popper's claim that the theory of evolution is not a scientific theory because it has no informative content (because it does not rule anything out, is not testable, nor allows for predictions to be made) by showing that it does have the features which Popper requires though often at levels higher than those which Popper considers.

J. Płazowski ('Rationality and beyond'), noting that today that science forms the ideal for rationality, rather than vice-versa, and how both are interpreted in an Objectivist sense, sketches how this has come about, the problems that it has met in the revolution within physics in this century, and the attempts to deal with those challenges: viz. verificationism, operationalism, and Popper's notion of an objective third world as the location of objective theories developing according to objective laws. But metaphysical conceptions, especially that of order, cannot be neatly separated from science. Nor is purely 'rational *ego*' sufficient for human creativity and discovery, within science as well as without it. Hence what is also needed is Polanyi's tacit knowledge. But at this point Płazowski seriously misinterprets Polanyi by equating tacit knowledge with a subconscious *id*, and by arguing that, in addition to Popper's rational ego and Polanyi's id, there is also needed a culturally determined *superego*, attuned to metaphysical conceptions such as order and symmetry. (In fact, Polanyi devotes Chap. 3 of *Personal Knowledge* to the idea of order and Chap. 7 to conviviality.) Tacit knowledge is not just a subconscious source, but a structure of integration of diverse clues and fundamental presuppositions, drawing upon the knower, his situation and the traditions wherein he stands, as well as the object known. The corrections that Płazowski wants to introduce to Popper are already provided by Polanyi. Nevertheless, he is right to conclude that rationalism is itself an ideology and that what is required is a less precise notion of rationality, a dialectical conception of philosophy that can cope with it, and a subject without an epistemology (an ideological epistemology) instead of Popper's subjectless epistemology.

W.T. Scott ('On Polanyi's notion of rationality') provides a masterly summary of Polanyi's epistemology and ontology of tacit integration, the role of personnal commitment and the corresponding kinds of rationality.

M. Zaberowski ('On the objectivity of the Popperian interpretation of quantum mechanics') argues that Popper fails to exclude all 'subjective' elements from his interpretation; that no sharp distinction between objective and subjective elements can be made in respect of our knowledge of physical processes; and than we need an anthropism which accounts for the objective recording of subject-object relations both when the subject is a real participant in the process and when the procees is similalrly considered as the subject and object of becoming, which is both an ability to recognise oneself in one's situation and an attribute of a process in which man does not participate.

J. Zycinski ('How to de-Ruse socio-biological theory') carefully considers the claims of socio-biology and its Darwinian and largely reductivist approach to scientific theories, and ('Tacit knowing and the rationality of science') provides a defence of science and Polanyi against those who misuse Polanyi's ideas in favour of subjectivism and relativism.

Inevitably a mixed volume, this does offer interesting comparisions of Popper and Polanyi, and treatments of other topics. But some papers should have been checked for spelling and grammar.

R.T. Allen

R.T. Allen
The Structure of Value
Ashgate Publishing (Avebury Series in Philosophy), 1993
ISBN 1 85628 458 1 £35

Richard Allen brings together a number of arguments which have often been put forward but not in a systematic way. The book is a comprehensive analysis of the structure of value, which challenges many conventional approaches. He challenges, for instance, the traditional fact/value distinction, the traditional concept of objectivity, and reductionism. A number of the ideas are derived from Michael Polanyi, in particularly the idea of perceiving patterns, *gestalten*, which Polanyi used to question the normal idea of scientific activity as searching for causal relations. He very much takes up the Polanyian idea of human activity as striving for achievement, which is used to develop his concept of responsible individualism and being a person.
In fact he uses these ideas to challenge the Kantian concept of autonomy. The Kantian individual appears as a mere technician obeying or not obeying rules rather than a real person who demonstrates his wholeness as an individual by actively taking responsibility for his actions.

In this first chapter he states that he is going to show that value and evaluation rest upon the category of activity, and that beings which show activity are of necessity achievements related to the realisation of their nature. To describe these achievements is to evaluate them in terms of the pre-moral and the moral.

The evaluation of human activity then is always the evaluation of the achievement of the activity in terms of responsibility.

The three tasks he sets himself are to discern the structure of value, to investigate the relation of value to reality, and to uncover the ways we apprehend reality, whilst recognising that different types of value may be related to reality in different ways. Value, he concludes, eludes a formal definition, because we can only grasps tacitly the indefinable unity between specific and wider types of value but we can say that values are specifications of goodness and badness. However, because we use the word 'value' in a whole lot of different senses, e.g. foundation value, utility value, or commercial value, when we say 'the value of x' it is not at all clear what we mean. The phrase can cover several logically distinct meanings and can emphasise some rather than others. It is also the case that objects can have value in different ways. The phrase can be replaced by 'the merits and demerits of x' in the first case to clarify its meaning, but in order to understand the second case the general ways in which things can be thought of as good, the dimensions of value need to be expressed.

Allen rejects the simple bad/good distinction, and finds three dimensions of value (good, worthless or unproductive, bad or counter-productive). He points out that being worthless or unproductive and being bad or counter- productive are distinct, e,g. one can be a worthless car driver, crashing gears, rolling back on hills, etc. but one can also be dangerous, hitting other cars and people.

There are also four categories of value: instrumentality or utility, foundation value, ingredience value, and performer and performance value. These four categories all presuppose the notion of activity, e.g. to perform is to perform an activity, ingredients arenecessarily ingredients in performances, when considered concretely, and ingredients in activities when considered abstractively. To be an ingredient is to be a necessary element in the whole. In the same way utility is related to purposive activities; and value as a necessary support or foundation is founded upon activities, e.g. food is necessary for living things, and living is doing things like eating, breathing, and reproducing.

The activities also set the standards of value in these categories: for instance, every performance is a good or bad performance of an activity. Activities set the standard of utility, for they state the degrees to which persons, things, events, processes, etc. aid or are irrelevant to, or obstruct the performance of activities or cause them to go wrong.

Activity then is the fundamental category in axiology, for it is activity which sets standards and requirements, and thereby generate the values, worthlessness and disvalues.

A special feature of an activitl is that it has a point or meaning. They have a 'goal', 'end', 'aim', or 'purpose', although the point of meaning can be as equally in itself and its performance as in something else. It can be realised in the activity itself, as well by the end point.

Allen argues, 'Activities have a point or meaning which their performers or performances either succeed or fail in realising . . . if there is a world without performers and performances, then that is a world without the possibility of value' (p.64). He points out that as there is life and mind in the world, and because values and disvalues are part of what there really is, it follows that the dichotomy of fact and value cannot be sustained. Or put in another way, 'If "things as they really are" are what they are totally independently of consciousness, then there really are no such things as machines and no such events as machines working nghtly and failing' (p.74).

He concludes his study of value by examining the structure of moral value, and argues that there are only two dimensions of value

within this sphere. The two dimensions are: conducting oneself in a morally responsible or morally perverse manner. The third possibility of acting irresponsibility or not bothering about the moral consequences of one's actions is really a case of moral perversity, and should be treated as such. There in fact can be no such thing as moral neutrality and mere worthlessness, for not doing good, and not containing evil is flouting what we ought to do.

However, to understand morality we need to distinguish a person from his mind and body, from his mental and physical powers, for he uses them and is not identical with them. He *'is* a value beyond any of his attainments . . . Persons, therefore, are to be *valued* and not only *evaluated* even in moral terms' (p.147). From this unique 'value essence' that is each person arise our duties toward persons.

In some ways the title of the book is misleading. It is a comprehensive study of value but goes far beyond this, for it puts forward an important message for responsible individuals. The message could have been easily lost in the dry argument of traditional philosophical analysis but in the last chapters takes off in a convincing and enthusiastic way.

R.J. Brownhill

David Selbourne
The Principle of Duty
Sinclair-Stevenson, 1994
ISBN 185619 4744 (hbk) £15
ISBN 185619 7204 (pbk) £10

Conservative critics of Liberalism have often argued that it takes for granted, and sometimes repudiates, traditions, especially of self-restraint, without which liberty declines into anarchy and thereby destroys itself. Dr Selbourne approaches Liberalism from the other side (and reminds this reader of Norman Dennis: *Rising Crime and the Dismembered Family*; and, with G. Erdos, *Families without Fatherhood*) but sustains a similar argument against the notion of 'dutiless rights' which dominate contemporary political thought and practice.

> If personal liberty is to retain its high ethical status, and the habitual use of force to maintain the civic order is to be disavowed, then the civic order must rest upon acceptance—preferably voluntary, but coerced by sanction if not—of the principle of duty, a principle which must be brought from the shadows into the political and moral light (p.33).

Old Socialists have created a 'universal plebeian', who receives public provision without reference to desert or merit and without any corresponding duties, because they equated duty with class obedience and had little regard for the total social order. Liberals, from Locke onwards, were aware of the principle of duty, but replaced pre-democratic obligations with rational self-interest and choice and put their faith in the citizen rather than expecting faith from him. On both sides there is lack of awareness of the need, on the part of all, for a sense of responsibility to the civic order.

The end result is the 'ostensible citizen' who is in fact a 'moral stranger' in a state of 'civil disaggregation', claiming his rights but owning no duties. Civic order rests upon the intuitive or instinctive recognition that one's own well-being is bound up with that of the whole, and not on any individualist calculation of interests nor collectivist priority of society over the citizen.

Unlike many cautious Liberals, who would agree with him so far, Dr Selbourne takes his argument a step farther. 'The obligation of the civic order to defend itself from assault has precedence over all obligations to citizens whatsoever' (p.86). That means that citizens have duties to protect that civic order, which is historically and logically prior to the state, whose function is to serve its interests. The aggregate of citizens in a civic order is an association which encompasses all other associations, including the state (p.90).

From these premises the author draws several instructive inferences, such as the need for social diversity within a moral unity, from which 'the moral stranger' ejects himself though not in the corrupted liberal order of the present day; the existence of obligations, not only on those in power to do justice, but the co-responsibility of all for the moral well-being of the civic order, and thus for positive commands of the law as well as prohibitions of unsocial conduct; for protection also against non-violent acts of oppression; that justice requires limits upon wants and wills as well as actions; obligations of all, and especially of parents and children,and to the natural environment and the patrimony of the civic order; that, in general, duties are prior to rights, which are bestowed by the civic order and not by contract; and the repudiation of the idea of a mere conglomeration of particular interests.

He specifies further duties of the civic order, aimed at arresting any further decline into disaggregation (e.g. to make divorce more difficult and by reviving the distinction between citizen and non-citizen or the citizen- turned-stranger), and of the citizens to each other and the civic order, and ways in which they may be fulfilled by citizens and not just officials.

The whole thrust of the book, and its specific suggestions, will doubtless be vigorously denounced by both collectivists and individualists. Likewise, because of its opening attack upon the amoral detachment and technical specialism of universities, and the absence of any notion of a principle of duty in current political philosophy (but *The Morality of Freedom* by J. Raz, is one unmentioned exception) it is unlikely to appear on reading-lists. Yet this reviewer hopes that Dr Selbourne (now in Italy) will further developed the principle of civic duty and provide further suggestions for overcoming civic disaggregation.

R.T. Allen