An Axiomatic Approach to the Study of Mind

Petros A. M. Gelepithis

Director, Cognitive Science Laboratory, Kingston University, KT1 2EE, England.
E-mail: Petros@kingston.ac.uk

Abstract

Cognitive Science is the study of intelligent systems. Methodologies for their study abound. Bechtel and Graham [1998] devote over 120 pages in describing 10 methodologies for the study of mind. In sharp contrast Wilson and Keil [1999] have not a single entry on this topic. From the system science perspective, [e.g., Delgado 1993], the interrelated consideration of the deductive, inductive and experimental approaches is a must for the study of mind.

The aim of this paper is to illustrate an axiomatic approach to the study of Mind. Specifically, in the first section we introduce our approach and its wider context within the Unity of Science movement. In section two, we specify necessary and sufficient conditions for intelligence as part of a foundation for a theory of mind. Finally, in section three, we provide a brief summary of further work.

Keywords: Cognitive science, axiomatic approach, unity of science, intelligence, theory of mind, understanding, communication.

1. Introduction

My working hypothesis is that the existence of mind implies the existence of an intelligent system. It follows that any theory of mind should be based on a sound theory of intelligent systems. The aim of this paper is to specify, *axiomatically,* necessary and sufficient conditions for intelligence. The rest of this section places our approach into its wider context.

The space of intelligent systems is huge, complex, and extremely varied [see, e.g., Gelepithis 2001]. Not surprisingly, methods and techniques for their study abound. Bechtel and Graham [1998] devote over 120 pages in describing a large number of them –classified in 10 methodologies. Such a classification is a recent approach to integration, namely, local integration on the basis of: (i) tools and techniques, (as exemplified by the rise of cognitive neuroscience); (ii) regularity clusters [as described by Newell 1990]; or (iii) interfield connections [Darden and Maull 1987]. Interestingly, *The MIT Encyclopedia of the Cognitive Sciences* [Wilson and Keil 1999] does not contain a single entry on methodologies.

Although both necessary and fruitful, local integration is inadequate for developing a theory of intelligence or mind. It has to be complemented by the development of unified theories (Newell, 1990). This objective and Newell's (ibid) call to arms for its achievement echoes the Unity of Science attempts pioneered by Carnap and Neurath, two prominent logical positivists of the Vienna Circle, [see, e.g., Carnap 1934]. It is also compatible with Wiener's [1948] cybernetic approach to the study of mind and von Bertalanffy's [1951] General Systems Theory. Delgado's [1993] instigation for a systemic methodology -requiring the interrelated consideration of the deductive, inductive and experimental approaches- confirms aspects of the aforementioned movement but he has not yet developed it into a specific theoretical framework.

In accordance with the overall objective of the Unity of Science movement that scientific knowledge should be integrated -a view rooted in Aristotle and the French Encyclopedists- our approach is characterised by:

1. Integration across disciplines in an *axiomatic* way that would not give preference to one discipline over another.
2. A shared base of observational phenomena that would serve as the inductive base of a unified, axiomatic theory.

On first reading of these objectives, the reader may be tempted to think that our approach is identical or at least too close to the traditional hypothetico-deductive (HD) model of science and he may even take sides along with Salmon [1990*1998] in considering the HD schema as an inadequate characterisation of the logic of science. This though would be a mistaken thought. To start with the reader may juxtapose the characteristics of our approach with the following two major senses of the HD approach. One, the term 'HD system' is used to denote a body of laws deducible from another body of laws [Ruse 1982*1998, p. 38]. Two, the HD model, in its simplest and most common form, "denies that there is any logic of discovery and affirms we are justified in accepting or rejecting theories only after they have been tested." [Curd and Cover 1998, p. 1298]. Second, the first characteristic of our approach is in clear contrast to the classical approach within the philosophy of science –reduction of all disciplines to physics. Finally, the reader is invited to consider the remarks 1-2, 2/3/4/5/-1, and 8-3 we make in the next section. They should dissolve any remaining doubts.

An additional, desirable characteristic for our approach would be the unification of the various levels of organisation in nature as it was aimed, and partially achieved, by Cybernetics and General Systems Theory. Unfortunately, we cannot currently demonstrate that this very desirable feature is an intrinsic characteristic of our approach and in particular, of the fundamental processes of understanding and communication.

The next section provides an outline of our approach as applied to the issue of the nature of intelligence. In other words, next section illustrates only the approach we follow and does not constitute a description of our theory so far. For some interesting results of our approach the reader is referred to Gelepithis [1991, 1997], Gelepithis and Goodfellow [1992], Gelepithis and Parillon [2002].

2. An Axiomatic Approach to the Study of Mind

Presently, the study of mind is still dominated by the computational paradigm [Newell 1990] although work based on the dynamical hypothesis [van Gelder 1998], and drawing upon the traditional artificial neural networks approach, is hotly pursued by an increasing number of centres around the world. The term third way is the name coined by Gelepithis [2002] to refer to the various attempts aimed to develop an alternative, more satisfactory, theoretical framework for the study of mind. So far, to my knowledge, the three most interesting attempts exemplifying the third way are those of Barsalou [1999], Edelman [1992], Edelman and Tononi [2000], Gelepithis [1989, 1991, 1995, 1997, 2001], and Gelepithis and Parillon [2002].

As we remarked in the introduction, a theory of mind presupposes a sound theory of 'intelligence'. The study of the latter pivots around three questions:

(a) What is the nature of intelligence?
(b) What are its mechanisms?
(c) What would constitute an objective measure of intelligence?

In what follows we focus on the first of these questions.

Over the years 'intelligence' has been defined in many ways [see, for example, Gregory 1987; Sternberg and Detterman 1986]. Not surprisingly, theories of intelligence have proliferated too. In AI and the psychology of human intelligence at least seven main types of theories may be found: genetic-epistemological [principal proponent Piaget 1947*1972]; physiological [e.g., Hebb 1949]; factorial [e.g., Guilford 1967]; information processing (IP) based [e.g., Newell and Simon

1972; Hunt 1980]; logical [principal proponent Nilsson 1991]; connectionist [e.g., Rumelhart, Hinton and McClelland 1986]; and functional [e.g., Chandrasekaran 1990].

Recently, both the scope and the approach to the study of 'intelligence' have changed dramatically. One may distinguish three types of extension of the scope of 'intelligence'. First, there is the move towards studying real life situations and problems, i.e., practical or everyday intelligence [e.g., Sternberg and Wagner 1986; Sternberg et al 1995]. Second, there is the move, especially in AI, of actually building entities able to interact among themselves and eventually with the world on the basis of increased intelligence -given or acquired [see, for example, Brooks 2002]. Finally, there is the move of broadening the study of intelligence by bringing in skills which are of fundamental importance in the shaping of one's intelligence. Such skills come under the generic term of emotional intelligence [Goleman 1995]. With respect to approach, one may distinguish two changes. First, there is the move towards a unified treatment of 'intelligence' [e.g., Sternberg 1985]. Second, there is the increasing interest in theoretically sound theories of intelligence, what Newell [1990] calls theory-bound conceptions of intelligence. It should be noticed that none of the extensions of scope or change of approach has resulted into a new type of theory. For example, Sternberg's triarchic theory is a mixture, but not synthesis, of elements of factorial, functional, and IP-based theories, and Newell's [1990] conception of intelligence is of the information processing type.

The rest of this section is a considerably revised version of the third section of Gelepithis [2002] and provides a delineation of the class of intelligent systems as foundation for the development of an axiomatic, unified theory of mind. We start with the specification of necessary and sufficient conditions for membership to the space of intelligent systems.

Definition-1:

A system, S, is intelligent if and only if:

a) It possesses sensors.
b) It is able to act on its environment.
c) It possesses its *own* representational system R_s, i.e., R_s is independent of the language of another kind of system S^*.
d) It is able to connect sensory, representational, and motor information.
e) It is able to communicate with other systems within its *own* class.

Remark 1-1: For a justification of this definition the reader is referred to Gelepithis [1984, 1991, 2001].

Remark 1-2: The above definition is independent of human and animal characteristics and, therefore, applicable both across disciplines from Psychology to Artificial Intelligence and the full range of intelligent systems. It, therefore, fulfils the first characteristic of our approach.

Consequence 1-1: It follows, from Definition-1 that the space of intelligent systems is extremely varied with nearly impenetrable regions of intelligence. For instance, a system, S, is human –level intelligent if and only if it is intelligent and it is able to communicate in a human-equivalent language.

Our next move then is to define the notions of 'communication' and 'representational system'. Definitions 2-5 deal with the latter.

Definition-2:

Let R_e be the symbol for the representational system of entity E. R_e is a representational system of E if and only if R_e is a thought system of E able to create representations.

Definition-3:

For an entity E, a representation of a situation, say, S_1 *is* another situation, say, S_2, characterised by the properties:

S_2 simplifies S_1; and
S_2 preserves the essential characteristics of S_1.

Remark 3-1: The situations S_1 and S_2 can be either external or internal to E excluding only the possibility of being both internal. Specifically, if S_1 is external then S_2 can be either external or internal. If S_1 is internal then S_2 can only be external.

Definition-4: $\mathcal{T}$ is a thought system of E if and only if $\mathcal{T}$ is a system of thoughts of entity E.

Definition-5: □ is a thought of E if and only if □ is an ordered n-tuple of meanings of E.

This brings us to the definition of meaning.

Definition-6:

The meaning, M, of something s, in the context C_s, for the entity E, at time t -symbol M (s, C_s, E, t)- is the *prevailed* formations of $R^m{}_e$, at t. ($R^m{}_e$ the representational material of E).

Remark 6-1: Our definition assumes the existence of a whole class of $R^m{}_e$ formations out of which a particular formation eventually prevails.

Remark 6-2: The term prevailed is used to indicate the potential complexity involved in the struggle for selection. Such complexity depends on the complexity of the R_e. In the case of humans, R_e = the human brain whose complexity is known to be hyper-astronomical.

Remark 6-3: It can be shown, but it is well beyond the scope of this paper, that concepts and instances of concepts can be represented topologically in terms of neighbourhood systems. Specifically:

(i) tokens (i.e., instances of concepts) are meanings represented by neighbourhoods; and

(ii) concepts are meanings represented by neighbourhood families.

Remark 6-4: Mathematically speaking, neighbourhoods are members of the power set $P(R^m{}_e)$, and neighbourhood families members of the $P(P(R^m{}_e))$.

Remark 6-5: It should be noticed that the cardinality of $R^m{}_e$, $P(R^m{}_e)$, and $P(P(R^m{}_e))$ changes over time through processes like development, learning, growth and understanding.

We turn now to the definition of communication:

Definition-7:

An entity E_1 communicates with E_2 on a topic T if, and only if:

(i) E_1 understands T -symbol: U (E_1, T);

(ii) E_2 understands T -symbol: U (E_2, T);

(iii) U (E_1, T) is describable to and understood by E_2; and

(iv) U (E_2, T) is describable to and understood by E_1.

This brings us to our last definition:

Definition-8:

An entity E has understood something, S, if and only if, E can describe S in terms of a system of *own* primitives (□ is a primitive if and only if E's understanding of □ is immediate).

Remark 8-1: In other words, understanding is a thought process characterised by reducibility to a system of primitives.

Consequence 8-1: Genuine understanding of S by entities E_1 and E_2 does not imply that E_1 and E_2 can communicate nor that U (E_1, S) and U (E_2, S) are necessarily identical.

Remark 8-2: For some major consequences following from our theory so far the reader is referred to Gelepithis [1991, 1997], Gelepithis and Parillon [2002].

Remark 8-3: Definitions 6-8 can jointly provide a shared base of observational phenomena in the form of, *eventually,* widely accepted sense and linguistic primitives and, therefore, satisfy the second characteristic of our approach.

With this definition our task of delineating the class of intelligent systems is now complete.

3. Summary of Further Work and Conclusion

As we remarked earlier, the axiomatic approach we illustrated in the previous section needs both to be applied to the theoretical framework we have developed so far and to be extended. The following agenda items provide an indication of the type of the extent required.

Apply our axiomatic approach to the notions and associated phenomena of perception, action, growth (inc. self organisation), purpose (inc. expectation), emotion, human language, consciousness, beauty, and ethical principles. This is a tall order but one needs to address as many of these issues as possible to develop a really unified theory of mind.

Investigate whether the processes of understanding and communication could be used as tools for unifying the various levels of organisation in nature.

Expand the implementation of the Understanding/Explanation matrix we developed in Gelepithis and Goodfellow [1992] to become the central subsystem for both types of the artificially intelligent systems introduced in Gelepithis [2001].

Specify the mechanisms for the genesis and storage of representations and investigate how does an entity select *essence-preserving simplifications* within a representational system [after Gelepithis 1995].

In summary, a systemic, unified theory of mind derived through the axiomatic approach illustrated in this paper will considerably enhance our understanding of minds (whether human or artificial).

References

Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences,* **Vol. 22**, pp 577-660 (including commentaries and response by the author).

Bechtel, W., and Graham, G., eds. (1998). *A Companion to Cognitive Science.* Blackwell Publishers Ltd.

Brooks, R. A. 2002. *Robot: The future of flesh and machines.* Allen Lane, the Penguin Press.

Carnap, R. 1934. *The Unity of Science.* Kegan Paul, London. (Translated from the German by M. Black).

Chandrasekaran, B. 1990. What kind of information processing is intelligence? A perspective on AI paradigms and a proposal. In Partridge, D., and Wilks, Y. (eds.). *Foundations of Artificial Intelligence: A Source Book.* Cambridge University Press.

Curd, M., and Cover, J. A., eds. 1998. *Philosophy of Science: The central Issues.* W. W. Norton and Company.

Darden, L., and Maull, N. 1987. Interfield Theories. *Philosophy of Science,* **Vol. 43,** pp44-64.

Delgado, R. R., ed. (1993). Systems theory, Basic concepts. In *International Systems Science Handbook: An Introduction to systems science for everybody.* Systemic Publications, 28010 Madrid, Spain.

Edelman, G. M. (1992). *Bright Air, Brilliant Fire: On the Matter of the Mind.* Basic Books, Harper Collins.

Edelman, G. M., and Tononi, G. (2000). *Consciousness: How Matter Becomes Imagination.* Allen Lane, The Penguin Books.

Gelepithis, P. A. M. (1984). *On the Foundations of Artificial Intelligence and Human Cognition.* PhD Thesis, Dept. of Cybernetics, Brunel University, England.

Gelepithis, P. A. M. (1989). Knowledge, Truth, Time, and Topological spaces. *Proceedings of the 12th International Congress on Cybernetics*, pp 247-256, Namur, Belgium.

Gelepithis, P. A. M. (1991). The possibility of Machine Intelligence and the impossibility of Human-Machine Communication. *Cybernetica*, **Vol. XXXIV,** No 4, pp 255-268.

Gelepithis, P. A. M. (1995). Revising Newell's conception of representation. *Cognitive Systems,* **4,** No 2, pp 131-139, (Special issue on Representation).

Gelepithis, P. A. M. (1997). A Rudimentary Theory of Information: Consuequences for Information Science and Information Systems. *World Futures* **49,** pp 263-274. Reprinted in *The Quest for a Unified Theory of Information.* Hofkirchner, W. (ed.). Gordon and Breach Publishers.

Gelepithis, P. A. M. (2001). *Intelligent Systems Volume-1: Knowledge Representation, Social and Psychological Impact of Artificial Intelligence.* Stamoulis Editions, Averof st., Athens , Hellas (Greece). (In Greek).

Gelepithis, P. A. M. (2002). Computing, Dynamics, and The Third Way. Proceedings of *The Second International Conference on Neural, Parallel, and Scientific Computations,* 7-10 August 2002, Atlanta, USA. (Invited Address).

Gelepithis, P. A. M. and Goodfellow, R. (1992). An alternative architecture for intelligent tutoring systems: Theoretical and Implementational Aspects. *Interactive Learning International*, Vol **8,** No 3, pp 171-175.

Gelepithis, P. A. M., and Parillon, N. (2002). Knowledge Management: Analysis and some Consequences. In *Knowledge Management and Business Process Reengineering,* pp 68-81, Hlupic, V. (ed.). Idea Book Publishers.

Goleman, D. 1995. *Emotional Intelligence: Why it can matter more than IQ.* Blumsbury, London.

Gregory, R. L., ed. 1987. *The Oxford Companion to The Mind.* Oxford University Press.

Guilford, J. P. 1967. *The nature of intelligence*. McGraw-Hill.

Hebb, D. O. (1949). *The Organisation of Behaviour.* John Wiley and Sons.

Hunt, E. B. (1980). Intelligence as an information-processing concept. *British Journal of Psychology*, Vol. **71,** pp. 449-474.

Newell, A. (1990). *Unified theories of cognition.* Harvard University Press.

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

Nilsson, N. J. (1991). Logic and Artificial Intelligence. *Artificial Intelligence,* Vol. **47,** pp 31-56.

Piaget, J. (1947*1972). *The Psychology of Intelligence*. Routledge and Kegan Paul. (Original in French).

Rumelhart, D. E., Hinton, G. E. and McClelland, J. L. (1986). A General Framework for Parallel Distributed Processing. In D. E. Rumelhart, J. L. McClelland and the PDP research group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Vol. 1, pp 318- 362.

Ruse, M. 1998. Creation-Science Is Not Science. In: *Philosophy of Science: The central Issues,* ed. M. Curd and J. A. Cover. W. W. Norton and Company.

Salmon, W. C. 1998. Rational Selection. In: *Philosophy of Science: The central Issues,* ed. M. Curd and J. A. Cover. W. W. Norton and Company.

Sternberg, R. J., 1985. *Beyond IQ: A triarchic theory of intelligence.* Cambridge University Press.

Sternberg, R. J., and Detterman, D. K., eds. 1986. *What is Intelligence? Contemporary Viewpoints on its Nature and Definition.* Ablex, Norwood, NJ.

Sternberg, R. J., and Wagner, R. K., eds. 1986. *Practical Intelligence: Nature and origins of competence in the everyday world.* Cambridge University Press.

Sternberg, R. J., and Wagner, R. K., Williams, W. M., and Horvath, J. A. 1995. Testing common sense. *American Psychologist,* **Vol. 50,** no 11, 912-927.

Van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences,* Vol. **21**, pp 615-665 (including commentaries and response by the author).

Von Bertalanffy, L. 1951. *General Systems Theory: A New Approach to Unity of Science.* Johns Hopkins University Press.

Wiener, N. 1948. *Cybernetics: Or, Control and Communication in the Animal and the Machine.* Wiley.

Wilson, R. A., and Keil, F. C. eds. (1999). *The MIT Encyclopedia of the Cognitive Sciences.* The MIT Press.