Authors' affiliation

Center for Biological Information Processing

Department of Brain and Cognitive Sciences

MIT E25-201

Cambridge, MA 02139

representation in biological systems becomes more important. In neurobiology, the new field of computational neuroscience embraces the development of computational models of neural function on all levels, from membrane channels and receptors to (real) neural networks. Conceptually, computational neuroscience brings together the ideas of representation and computation that originated in AI, the mathematics of complex systems, and the tradition of neurobiological modeling that goes back to McCulloch, Hodgkin, Huxley and Hebb.

Further reading

Fischler M and O Firschein (1987): *Readings in Artificial Intelligence*. San-Mateo, CA: Morgan Kaufman.

Marr D (1982): *Vision*. San-Francisco, CA: Freeman.

Posner MI, ed., (1989): *Foundations of Cognitive Science*. Cambridge, MA: MIT Press

Artificial Intelligence. Daedalus, Winter 1988. Issued as Proc. of the American Academy of Arts and Sciences, vol. 117, no. 1 (1988).

Trends

Representation, learning and memory. The evolution of the notion of representation and the development of new paradigms for learning and self-organization affect some of the basic issues in AI, outlined above. One example of this is the development of a computational theory of probabilistic learning, which seeks to identify classes of representations that can be learned from examples by an efficient procedure. Another example is the understanding of representation and learning in distributed dynamical systems that begins to emerge from analyzing ANNs using standard mathematical tools provided by linear algebra, statistics, theory of dynamical systems and function approximation. Such an understanding could lead to an increase in the perceived role of memory in intelligence. In particular, mechanisms based on memory of specific experiences rather than on intensive computation may prove useful in applications ranging from automated reasoning and expert systems to visual recognition.

Computational neuroscience. The difficulty of devising efficient representations for a variety of problems in reasoning, memory, perception and motor control motivates a closer cooperation within different branches of cognitive science. At the behavioral level, the contribution of psychological studies of

attempts to embed AI systems in real-world environments (as opposed to artificially limited problem domains). Making AI work in the real world appears to require a knowledge base that is both extensive and ecologically valid. In visual recognition, for example, this problem is reflected in the need for large libraries of object and scene representations.

A natural way to meet these challenges is by making the AI systems (1) *learn* the format of the representations they employ, and (2) amass the necessary real-world knowledge through experience. Automatic learning, long studied as a discipline within AI, may thus be a prerequisite not only for adaptive behavior, but for the very acquisition of knowledge necessary to support complex functions such as object recognition or reasoning. Computational studies of symbolic learning algorithms are now complemented by an intensive investigation of a class of parallel computing devices called artificial neural networks or ANNs. In ANNs (which at present bear little architectural or functional resemblance to real neural networks) representation and computation correspond, respectively, to patterns of activity and to the spread of activation in networks of simple processing elements. Learning in ANNs is achieved by modifying connection strengths between elements in a manner that brings the system's performance closer to a preset goal.

cumbersome to use, even though in principle equivalent to the map. A second consideration relevant to the choice of representation is its computational tractability. Thus, a planning module that maintains an explicit representation of all consequences of its actions may require exponentially growing resources when applied to real-life situations.

Challenges to the classical view

In recent years, the standard view of AI, as outlined above, has been repeatedly challenged. First, the need for elaborate representations has been questioned. For example, it has been assumed since Marr's work that the ultimate product of a visual system should be a three-dimensional reconstruction of the environment. The research program based on this assumption led to substantial progress in the understanding of low-level visual processes such as stereopsis and motion perception, where powerful mathematical methods were invoked to identify and apply computational constraints relevant to the initial stages of visual reconstruction. In contrast, it becomes increasingly apparent that higher-level visual processes such as navigation and recognition may be better mediated by representations that do not attempt full three-dimensional reconstruction.

A second challenge arises in connection with the difficulties faced by the

computational-level theory of information processing, are addressed.

Representation as the basis of intelligence. The majority of researchers in AI and related fields currently believe that internal representation of relevant features of the outside world is a prerequisite for intelligent behavior in both artificial and biological information processors. Learning to represent the world, maintaining the representations in the face of the changing situation and using them to achieve behavioral goals are considered to be the basic ingredients of intelligence.

The dogma of independence of information-processing and implementation levels leads to the emergence of theories of representation centered around the notion of abstract computational utility. One computational factor in assessing the suitability of a representation scheme to a particular task is the kind of information made explicit by the scheme. For example, a chess-playing system is expected to represent the state of the board in a manner that would directly show impending threats and attack opportunities (that is, a one-dimensional vector of symbols standing for the different pieces will not do). Similarly, a robotic navigator should maintain an analog, map-like representation of its environment, since any alternative (e.g., a set of propositions encoding the locations of the various obstacles) would be more

soning about knowledge, heuristics and search methods, knowledge-based systems).

General notions

Levels of understanding of intelligent systems. Intelligence, considered as an information-processing phenomenon, may be studied at three distinct levels. The top level is concerned with the computational theory, in which the problem is characterized as a mapping from one kind of information to another, and the abstract properties of this mapping are analyzed. At the middle level, there is the choice of algorithm for the mapping, determined, among other factors, by the nature of representation at its input and output. At the lowest level, there are the implementational details, such as the physical realization of the representation and the algorithm.

It is the existence of the abstract computational level in the description of information-processing systems that provides ground for the assumption that intelligence can be studied in separation from its natural substrate, the nervous system. In an example due to David Marr, the study of intelligence is likened to the problem of understanding the phenomenon of flight in birds. Both wing anatomy and the microstructure of a feather are better understood once the relevant problems in aerodynamics, analogous to the

Artificial Intelligence – An Update

Shimon Edelman and Tomaso Poggio

Since its emergence in the 1950s, the field of Artificial Intelligence (AI) has been shaped by the interaction of two goals. The first of these is to substantiate the basic tenet of AI, namely, that intelligence can be given an information-processing formulation that is independent from the physical substrate that implements it. The second goal, linking AI to neuroscience and in a sense complementary to the first one, is to understand the characteristics a physical system should possess to be able to support intelligence. In contrast with the traditional approach in the philosophy of mind, AI combines theoretical study of the relevant questions with attempts to replicate faculties such as memory, reasoning and learning that are an integral part of intelligence. AI is also concerned with theoretical and practical aspects of perception and action (e.g., vision and visual recognition, hearing and speech understanding, planning and motor control). The combined theoretical and applied strategy essential to AI provided a new angle on some old problems in epistemology and led to the development of entire new disciplines in mathematical logic and computing science (e.g., truth maintenance, rea-