Part I Artificial Intelligence
A general survey by Sir James Lighthill FRS Lucasian Professor of Applied Mathematics, Cambridge University. July 1972.

Contents

1 Introduction

The Science Research Council has been receiving an increasing number of applications for research support in the rather broad field with mathematical, engineering and biological aspects which often goes under the general description Artificial Intelligence (Al). The research support applied for is sufficient in volume, and in variety of discipline involved, to demand that a general view of the field be taken by the Council itself. In forming such a view the Council has available to it a great deal of specialist information through its structure of Boards and Committees; particularly from the Engineering Board and its Computing Science Committee and from the Science Board and its Biological Sciences Committee. These include specialised reports on the contribution of Al to practical aims on the one hand and to basic neurobiology on the other, as well as a large volume of detailed recommendations on grant applications.

To supplement the important mass of specialist and detailed information available to the Science Research Council, its Chairman decided to commission an independent report by someone outside the Al field but with substantial general experience of research work in multidisciplinary fields including fields with mathematical, engineering and biological aspects. I undertook to make such an independent report, on the understanding that it would simply describe how Al appears to a lay person after two months spent looking through the literature of the subject and discussing it orally and by letter with a variety of workers in the field and in closely related areas of research. Such a personal view of the subject might be helpful to other lay persons such as Council members

in the process of preparing to study specialist reports and recommendations and working towards detailed policy formation and decision taking. The report which follows must certainly not be viewed as more than such a highly personal view of the Al field. It owes much to the study of published work and of private written communications and spoken comments by numerous individuals, including the following:

J.Annett, H. G. Barrow, S. Brenner, D. E. Broadbent, R. A. Brooker, 0. P. Buneman, R. M. Burstall, A.D. B. Clarke, M. B. Clowes, A. H. Cook, D. C. Cooper, J. E. Doran, J. F. Duke, E. W. Elcock, J. Good, C. C. Green, R. L. Gregory, P. J. Hayes, A. L. Hodgkin, J. N. Holmes, J. A. M. Howe, D. H. Hubel, S. Isard, H. Kay, T. Kilburn, J. Kulikowski, D. N. L. Levy, H. C. Longuet-Higgins, D. M. Mackay, D. Marr, J. McCarthy, B. Meltzer, D. Michie, M. Minsky, D. Mollison, E. Moore, J. S. Moore, R. M. Needham, N. J. Nilsson, C. Oldfield, J. V. Oldfield, I. PohI, R. J. Popplestone, B. Raphael, J. A. Robinson, C. Strachey, N. S. Sutherland, M. M. Swann, H. P. F. Swinnerton.Dyer, D. Wilshaw, T. Winograd.

The author is grateful for the large amount of help and advice readily given in reply to his many requests. He must emphasize, however, that none but himself is responsible for the opinions expressed in this report. They represent merely the broad overall view of the subject which he reached after such limited studies as he was able to make in the course of two months. Readers might possibly have expected that the report would include a summary, but the author decided against this partly because considerable material is summarised already in almost every paragraph. Furthermore, he believes that this kind of report can be valuable only to those who read it all, and for this reason preferred to avoid attempting a condensation.

2 The ABC of the subject

There is a general consensus about which main areas of research are to be grouped within the broad field of Al. This section lists briefly these main areas and divides them further into three categories, A, B and C according to the long-term motivations for the three different types of work.

Here, categories A and C have clearly distinct motivations: each has a well defined general direction of its intended objectives, but the two directions are quite different. In both these categories a certain amount of rather respectable progress has been made during the subject's twenty-five years of life (which may be taken as beginning with Turing's 1947 article 'Intelligent Machinery'), although expectations have as we shall see in section 3 been frequently disappointed. During the same period a further category 'B' of researches has been pursued: a 'bridge' category where aims and objectives are much harder to discern but which leans heavily on ideas from both A and C and conversely seeks to influence them. Research in category B, if acceptable arguments for doing it can be agreed, works by its interdependence with studies in categories A and C to give unity and coherence to the whole field of AI studies. There is, however, a widespread feeling (Section 3) that progress in this bridge category B has been even more disappointing, both as regards the work actually done and as regards the establishment of good reasons for doing such work and thus for creating any unified discipline spanning categories A and C.

Category A

Here, letter A stands for Advanced Automation: the clear objective of this category of work being to replace human beings by machines for specific purposes, which may be industrial or military on the one hand, and mathematical or scientific on the other. The work looks beyond automation of the type that is widely adopted at present in control engineering and data processing, and aims to make a far fuller use of the general-purpose digital computer's logical (as opposed to arith- metical) potentialities. Nevertheless it must be looked at as a natural extension of previous work on the automation of human activities, and be judged by essentially the same criteria.

Industrially important purposes include, for example, machine recognition of printed or typewritten characters (an area where good progress has been made) and of handwritten characters (incomparably more difficult), as well as a much wider range of pattern-recognition activities. The auditory equivalent to this visual area is speech recognition and synthesis. There are great economic incentives for work in machine recognition of speech, as well as in machine translation between languages, although progress in both has so far been very disappointing.

A further industrially important aim is to go beyond the automation of component design and manufacture, towards automation of design and assembly of whole products. It is argued that the complex spatial relationships involved in assembly processes put them far beyond the scope of conventional control engineering and require a much more advanced logical structure in the controlling software. Similar arguments may apply to problems of improving packing ratios in parcel con- tainerisation.

The level of automation that can be called advanced has to be placed higher in the military field with its remarkable achievements both in cryptography and in guided missiles. A modern missile's capability to move in response to its own perception of its target against a noise background is highly reminiscent of the way in which a predator uses its complex central nervous system to home on to its prey. Beyond this the military have an incentive, however, to build less specialised devices that might be programmed to perform in hostile environments a far wider range of actions in response to information from organs of perception. Space exploration and, perhaps, some parts of industry (including firefighting) may look for a similar hostile-environment capability.

In the meantime, the application of digital computers in mathematical work has gone beyond the mere organisation of numerical calculations on a large scale and includes for example some very effective programmes for massive manipulations in algebra and analysis. Category A looks well beyond these, however, to the automation of problems of logical deduction including theorem proving, and still further to the automation of inductive generalisation and 'analogy spotting'. In scientific applications, there is a similar look beyond conventional data processing to the problems involved in large-scale data banking and retrieval, The vast field of chemical compounds is one which has lent itself to ingenious and effective programs for data storage and retrieval and for the inference of chemical structure from mass-spec- trometry and other data.

Information retrieval is, indeed, one of two dominant themes under- lying all work in category A: this work is found to depend essentially on a 'knowledge base' which the program causes to be stored in the computer, and the 'file structure' of this knowledge base is of crucial importance in determining how data is accessed and used in the machine's operations. The other dominant theme is problem solving. This goes beyond mathematical theorem proving into the solution of numerous common- sense problems' such as may arise in industrial and other applications. They can often be represented as problems of 'transversing a graph', using 'graph' in the specialised mathematical sense: an assemblage of points or nodes representing states of the system studied, some but not all pairs of nodes being linked by a line representing a permitted transition between states. Programs may be sought for solving problems in the sense of finding 'optimal' (eg shortest) paths between remote nodes on such a graph.

Longer-term objectives in category A include that of combining a well structured knowledge base and an advanced problem solving capability to generate improved methods for industrial and economic planning and decision making, although admittedly there will always be serious difficulties in establishing that any particular program must necessarily have an acceptable output of plans and decisions! Another longer-term objective permeating all work in category A, furthermore, is to incorporate into programs an increasingly greater capability of 'learning', so as to reach improved levels of performance in response to experience with tasks already undertaken. Efficient modes of learning will, however, be seen in section 3 to remain somewhat elusive. To sum up, category A is concerned with Advanced Automation aimed at objectives such as written character recognition, pattern recognition, speech recognition and synthesis, machine translation, product design and assembly, container packing, exploration and action in hostile environments, theorem proving, inductive generalisation, analogy spotting, information storage and retrieval, analysis of chemical structures, problem solving, graph traversing, learning and decision making. In marked contrast to the diversity characteristic of all these numerous objectives, whether practical or scientific, is the much more unified fundamental-research aim of category C: computer-based studies related to the Central Nervous System (CNS) in man and animals.

Category C

Thus, letter C stands for Computer-based CNS research. In a lay person's report the extended term central nervous system is used in preference to the term brain which to a lay person may have subjective associations with the more conscious, or more consciously 'brainy', parts of the brain activity, as against the emotional parts or those associ- ated with perception and movement. The co-ordination of perception and movement in animals generally is a particularly significant area of research which the lay person (unlike the biologist) might be tempted to forget if the word 'brain' conjures up for him the specifically human aspects of brain activity.

Category C is concerned, then, with theoretical investigations related to neurobiology and to psychology. The word 'theoretical' is used here to emphasize that we are concerned, not at all with the use of computers merely to process experimental data, but with their use to build models of CNS processes whose performance can in due course be compared with experimental data - a phrase carefully chosen to be uncontroversial as between critics of work in category C who argue that it takes insufficient account of existing data and some of the research workers who feel that the experimentalists will need the stimulus of revolutionary theoretical ideas to produce their best work.

It must be emphasized that the use of computers in building and evaluating theories of neurophysiological and psychological phenomena is a trend in no way out of the ordinary: the great majority of theories in physics and chemistry are built up and evaluated on computers, and similar habits are now increasingly permeating the biological sciences. Biologists generally accept that computer-based theories in their field, far from implying any disrespect to the special characteristics of living matter, may have quite as much value as in physics and chemistry for stimulating understanding and suggesting new kinds of experiment - provided only that the theoretical work takes proper account of available observational data.

Category C is especially concerned with theories to interpret neuro- biological data on specific areas of the CNS, using computer-based models of neural nets to test out particular hypotheses on (say) the functioning of the cerebellar cortex. Other theories, of (say) parts of the visual cortex, may seek to relate both to neurobiological and to psycho- physical data. Generally speaking, mathematically educated persons may be most effective in this field after prolonged study of CNS anatomy and physiology. Conversely, experimental psychologists and neuro- physiologists may in several cases become expert in the construction of computer models from which new theoretical concepts may develop.

Other important aims in category C include the development of computer models related to observations of a strictly psychological nature, such as data on visual pattern recognition and scene analysis, on visual and auditory memory, on general aspects of associative recall. A further series of aims refer to specifically human types of CNS activity: thus, psycho-linguistic studies are concerned with theories of the psychological processes concerned in the use of language, while other studies probe similarly the processes involved in classification and inductive generalisation. These are areas where the computer-based models of neural activity are inevitably remote from the hard facts of neurobiological observation, but where contact with the data of experimental psychology is of crucial importance.

Some workers in this field identify the essential long-term aim as 'understanding the human intellect', but they mean this only in the sense that the aim of cosmology is 'understanding the past, present and future of the universe'. There is no implication that such generality is apparent in any one group of research lines, merely that a general direction of desired improvement of knowledge is common to many such groups.

One more group of category C researches is concerned with how the human intellect acquires knowledge and skills, and this is related to educational psychology. For example, behavioural data on the order of acquisition of different abstract concepts in childhood may be studied in relation to models for the structuring of such concepts within the CNS.

To sum up, category C is concerned with basic research on Computer- based studies of CNS function, including the function of particular areas like the cerebellar cortex or parts of the visual cortex, and also special functions like visual pattern recognition and scene analysis, visual and auditory memory, associative recall, psycho-linguistics classification, inductive generalisation and learning. This is work' essentially within the life sciences and involving the pursuit, for its own sake, of knowledge which must appear to us as introspective living beings particularly desirable of attainment.

Evidently, there is a vast difference of approach between the practical, technological aims of category A (Advanced Automation of human activities) and the fundamental, biological aims of category C (Computer- based CNS studies). The aims are in each case perfectly clear, and perfectly distinct. The affinities in each case are much stronger with neighbouring fields (category A with general computer science and control engineering; category C with general neurobiology and psychology) than with each other. The appearance of a few common terms among the interests within the two categories (for example: pattern recognition, linguistics, inductive generalisation and learning) does admittedly indicate a degree of overlap, but may exaggerate its extent, as the problems of simulating these functions to achieve practical aims are not necessarily at all like the problems of studying how the CNS achieves them. If categories A and C were the whole body of research with which we had to deal we would recognise a minor extent of overlap of interest but regard the two areas of work as quite sufficiently distinct to warrant completely separate treatment in respect of research support, departmental organisation, etc.

Category B

Thus, the whole case for the existence of a continuous, coherent field of Artificial Intelligence research (Al) depends critically on whether between categories A and C there exists a significant category of research that may be described as a 'Bridge' category, B, as well as on the strength of the case for any researches in that category. The existence of research work in this category is hardly in dispute: such work, as stated earlier, has been voluminous for many years, but there are much greater difficulties in any attempt at clear identification of good reasons for putting resources into those researches. The activities and stated aims of work in category B are described in the remainder of section 2.

Here, letter B stands not only for 'Bridge activity', but also for the basic component of that activity: Building Robots. The whole concept of Building Robots is, indeed, seen as an essential Bridge Activity justified primarily by what it can feed into the work of categories A and C, and by the links that it creates between them.

Thus, a Robot in the sense used here, and by most workers in the field, is an automatic device that mimics a certain range of human functions without seeking in any useful sphere of human activity to replace human beings. Work in category B (Building Robots) is frequently justified because it simultaneously supports category A (Advanced Automation), in the sense that generalised information on automatic devices may emerge which can be used in practical problems of Automation, and supports category C (Computer-based CNS studies), in the sense that devices that mimic a human function may assist in studying, and in making a theory of, that function.

These are serious arguments, that will need to be considered seriously in sections 3 and 4. On the other hand, they are probably by no means the only reason why Building Robots is a popular activity. At the other extreme of the spectrum of reasons we have to remember the long- standing captivation of the human imagination by the very concept, as shown by its continual prominence in literature, from medieval fantasies of the Homunculus through Mary Shelley's 'Frankenstein' to modern science fiction. To what extent may scientists consider themselves in duty bound to minister to the public's general penchant for robots by building the best they can?

Incidentally, it has sometimes been argued that part of the stimulus to laborious male activity in 'creative' fields of work, including pure science, is the urge to compensate for lack of the female capability of giving birth to children. If this were true, then Building Robots might indeed be seen as the ideal compensation! There is one piece of evidence support- ing that highly uncertain hypothesis: most robots are designed from the outset to operate in a world as like as possible to the conventional child's world as seen by a man; they play games, they do puzzles, they build towers of bricks, they recognise pictures in drawing-books ('bear on rug with ball'); although the rich emotional character of the child's world is totally absent. Builders of Robots can justly reply that while robots are still in their infancy they can mimic only pre-adult functions and a limited range of those at most, and that these will lead on to higher things. Nevertheless, the view to which this author has tentatively but perhaps quite wrongly come is that a relationship which may be called pseudomaternal rather than Pygmalion-like comes into play between a Robot and its Builder.

General aspects of work in category B involve work on mimicking some special functions that are particularly highly developed in man: co-ordination of eye and hand; visual scene analysis; use of natural language; 'commonsense' problem solving. These areas for work in category B are evidently well chosen for giving good chances of feeding valuable results into the work of categories A and C.

Various reasons including limitations of computer power have restricted the 'universe of discourse' in which the functions just mentioned are exercised in existing robots to something like a chess- board, or a simple 'table-top world' on which coloured blocks are moved about and stacked on one another. Several workers have argued that games such as chess and draughts are ideal spheres for development of robot potentialities because there is great scope for ingenuity but little waste of programming effort on inessential features resulting from too extensive a universe of discourse.

To sum up, category B is a Bridge Activity concerned with Building Robots for purposes which include the feeding of information into the work of categories A and C; each Robot is designed to mimic some group of human functions, including functions such as eye-hand co-ordination, scene analysis, use of natural language, problem solving, etc, within some limited universe of discourse such as we may exemplify by a game (chess, draughts, etc), a puzzle, a table top on which blocks are moved about, or a drawing-book. One's views of the fundamental coherence of the whole field of Al spanning categories A, B and C must depend on one's opinion on whether the arguments for this Bridge Activity in category B are sound enough for it to be regarded as a necessary concomitant to, and link between, the rather different and rather easily defensible activities in categories A and C.

3 Past disappointments Most workers in Al research and in related fields confess to a pro- nounced feeling of disappointment in what has been achieved in the past twenty-five years. Workers entered the field around 1 950, and even around

1960, with high hopes that are very far from having been realised in 1 972. In no part of the field have the discoveries made so far produced the major impact that was then promised.

The disappointment felt may be analysed into two kinds: work in the categories A and C of section 2 has some respectable achievements to its credit (and achievement in such categories of work with rather clear aims is clearly discernible), but to a disappointingly smaller extent than had been hoped and expected, while progress in category B has been even slower and more discouraging, tending (as explained in section 2) to sap confidence in whether the field of research called Al has any true coherence. In the meantime, claims and predictions regarding the potential results of Al research had been publicised which went even farther than the expectations of the majority of workers in the field, whose embarrassments have been added to by the lamentable failure of such inflated predictions.

These general statements are expanded in a little more detail in the rest of section 3, which has been influenced by the views of large numbers of people listed in section 1 but which like the whole of this report represents in the last analysis only the personal view of the author. Before going into such detail he is inclined, as a mathematician, to single out one rather general cause for the disappointments that have been experienced: failure to recognise the implications of the 'com- binatorial explosion'. This is a general obstacle to the construction of a self-organising system on a large knowledge base which results from the explosive growth of any combinatorial expression, representing numbers of possible ways of grouping elements of the knowledge base according to particular rules, as the base's size increases.

Category A

Achievements within the sphere of the Advanced Automation (category A) have to be judged in competition with what industry has been able to achieve during the same period by perfectly conventional methods of control engineering and data processing. We may remind ourselves of the toughness of this competition by two examples. The human skills required to land a large aircraft reliably and safely are complex and intricate; yet the Automatic Landing System of Smith's Aviation Ltd., which uses classical control technology, has a better than human performance and has now been certified by the Air Registration Board, which for the purpose had demanded to be convinced of a less than 1 in (10 to the power 7) failure rate. Another British firm, Image Analysing Computers Ltd., has had a considerable commercial success using conventional programming methods to analyse images (eg microscope slides) as scanned by a television raster and to give numerical data (eg on metallographic grain shapes and sizes, or on cell characteristics in blood samples) without human intervention; automatic cervical-smear analysis now seems achievable by these means.

Workers in category A, while recognizing the effectiveness of such conventional control-engineering and data-processing methods applied to particular specialised tasks, have tended to emphasize the likelihood of Advanced Automation techniques of far more general applicability emerging from their work. The concept of automatic devices or methods with general capabilities is certainly a most attractive one. It is therefore particularly disappointing that the experience of the last twenty-five years has increasingly forced workers in category A to conclude that Advanced Automation techniques are successful not when they are developed with a high degree of generality of application, but only when a large quantity of detailed knowledge about the problem domain is utilised in the program design.

While this conclusion, which is rapidly gaining acceptance, has been undermining one of the clearest overall justifications for work in category A, performance of Advanced Automation systems developed at great expense in problem domains of particular economic importance has generated a still stronger sense of disappointment. Work in the pattern-recognition field has not yet proved competitive with con- ventional methods: even the recognition of printed and typewritten characters posed a quite surprising degree of difficulty, while the recognition of handwritten characters appears completely out of reach. Speech recognition has been successful only within the confines of a very limited vocabulary, and large expenditure on schemes to produce machine recognition of ordinary speech has been wholly wasted. Learning techniques, by which a machine's performance at recognising words might improve on receiving identified words from more and more individual speakers, appear feasible only for an exceedingly small vocabulary (such is the power of the combinatorial explosion) like the decimal digits!

The most notorious disappointments, however, have appeared in the area of machine translation, where enormous sums have been spent with very little useful result, as a careful review by the US National Academy of Sciences concluded in 1966; a conclusion not shaken by any subsequent developments. Attempts based on classical grammar and syntax and on the transformational grammar of contemporary general linguistics have been equally unsuccessful in producing acceptable programs. Suggestions from recent research (see below), that analysis and use of natural language by computer succeed only when a very detailed knowledge of the universe of discourse is stored within the machine, augur badly for the future availability of machine-translation programs versatile enough to be commercially valuable.

Mathematical theorem-proving is another area of work in category A that has had its disappointments. Of course, conventional programming is used by many pure mathematicians with great success to generate examples suggesting, or counter-examples disproving, theorems; while conventional proofs that leave a finite residuum of cases unaccounted for may often be completed by a computational survey of those cases. The mathematician is then using the computer as a fast, reliable and 'biddable' number-cruncher, the role in which computers generally have been most successful.

In the nineteen-fifties and early nineteen-sixties, however, a great deal of optimism was generated from the concept of realising on a computer the algorithms for theorem proving suggested by the decid- ability propositions of mathematical logic, starting with the ‘completeness theorem of the first-order predicate calculus. Those most involved now emphasize that this is particularly an area where hopes have been disappointed through the power of the combinatorial explosion in rapidly cancelling out any advantages from increase in computer power. The modern trend is to ‘heuristic’ methods, which also are the only methods that have been found effective in the general areas of problem solving and graph traversing.

It is important to understand the meaning attached to this adjective ‘heuristic’ which increasingly permeates the Artificial Intelligence literature: it means that the program stores and utilises a large amount of knowledge derived from human experience in solving the type of problem concerned. Thus, it depends critically on data derived through the use of human intelligence, so that the widespread view that only heuristic methods are effective is a serious setback to more extreme versions of the Al philosophy. For example, a heuristic graph-traversing program requires stored values of a human estimate of the ‘nearness’ of each node to the desired goal.

Some interesting concepts from mathematical logic that have been influential on work in category A are those relating to two or more statements including variable elements: namely, their ‘resolution’ or greatest common instantiation, and its ‘dual’ their least common generalisation which represents a sort of ‘inductive inference’ from them. Algorithms exist for obtaining these but limits on their practical use again result from the combinatorial explosion.

There are, in addition, difficulties in using the techniques of mathe- matical logic in heuristic programs based on stored knowledge, particularly because the type of stored knowledge favoured by logicians, namely a set of axioms, is inconvenient for access by practical programs I

An excellent example of successful work in category A that has resulted from storage of as much detailed information as possible about the problem domain is the ‘heuristic dendral’ program for inference of chemical structure from mass-spectroscope data. Its output is a list of possible molecular ‘graphs’ (ie structures) in order of decreasing plausibility that are consistent with the mass spectrum and the empirical formula and in some cases data of certain additional types. It has been the extremely careful study of extensive detailed information affecting the relationship of chemical structures to mass spectra that has brought about the relatively good success in this field.

In just the same way, quite good performance has been achieved in complicated areas of data storage and retrieval where the problems were confined to data of very precisely defined and analysed types. By contrast, generalised information

-retrieval systems have been somewhat disappointing, especially when applied to research information involving relatively advanced ideas.

To balance seriously limited successes in achieving the longer-term objectives of work in category A, one must recognize a great deal of 'spin-off' from such work, and from associated work in category B, into the 'software industry' and into programming technique generally. Certain high-level programming languages developed for this work have proved invaluable in a wide range of programming activity. The list- processing languages have many advantages over conventional pro- gramming languages; for example they eliminate the labour of preliminary estimation and organisation of store space. Languages specially suitable for problem solving and for linguistic analysis have also been derived. Their advantages include 'automatic back-tracking' by which if a particular subroutine fails all activity is 'unwound' back to a specific point and then an alternative subroutine is tried. There is a very wide- spread appreciation of the many merits of this group of programming languages. It must be admitted, on the other hand, that excellent work on developing high-level programming languages has been done also in regular computing laboratories and in research groups devoted to general computational theory.

Category C

The history of work in category C (Computer-based CNS Studies) has been somewhat similar to that in category A: in spite of a respectable volume of achievement resulting from such studies, most workers who ered the field around ten years ago confess that they then felt a degree of 'naive' optimism which they now recognise as having been misplaced. It is, once more, the most generalised types of studies whose end-products have proved most disappointing.

There is a consensus of view that benefits from this work in category C have flowed primarily to the science of psychology: in fact, a new range of attitudes to psychological problems has been generated. Computer models, although yielding no sudden breakthrough, have helped distinguish between theories of psychological phenomena which are possible candidates for consideration and theories that simply cannot be made to work.

As might be expected, some of the best work is by actual experimental psychologists with a good knowledge of a complex mass of data who have acquired the skills needed to build computer models for interpreting it: such work developed, for example, the concept of the visual buffer store. One school of thought emphasizes the value of intimate relation of computer models to detailed CNS data so exclusively as to propose denial of computer capacity to more theoretical groups until the demand for computer capacity from such experimental psychologists and from neurophysiologists is fully met.

Another school of thought sees a real place for the more speculative theorists, however, and points out the potential value of Current studies of the types of neural networks that might be effective for functions such as associative recall, classification by attributes and inductive generalisation. It is easy to believe that, as in physics and chemistry, the more speculative theorists do have a real role to play in generating ideas. On the other hand, some of the most significant work in these neural-net theories has been done in close association with local neurobiological data. Furthermore, some of the most remarkable neurobiological discoveries, including many on the structure of the visual cortex, have not required any computer-based modelling at all! A properly balanced view of work in category C may perhaps be that the besetting applied-mathematics sin of taking insufficient trouble to master the experimental facts needs to be carefully guarded against but that, if it is, the work produced can significantly help in the long process of moving towards better understanding of CNS function.

Psycho-linguistics is an area of psychology where this may particu- larly be the case. The algorithmic approach to the subject apparent already in transformational grammar and its syntactical theories of how sentences are generated is now being extended to involve algorithms taking into account more semantic information; that is, more knowledge about the universe of discourse. This type of algorithm looks much more promising as a model of how the CNS processes language.

The area nearest to an applied science which we listed in section 2 as coming within category C was educational psychology. There has recently been speculation on whether the time may be ripe for research aimed at direct application of Al research to educational method through the development of advanced forms of Computer Aided Instruction (CAI).

There is a well established, 'classical' approach to CAI that gives quite good results in educational areas that may be described politely as 'drill and practice', or less politely as 'cramming'! The 'teaching machine' is programmed to print out factual information interspersed with multiple-choice questions and to go into various alternative loops, in which it prints out encouraging or corrective comments (with additional questions in the latter case), according as the right answer or one of the wrong answers is chosen.

There are those who hope to go beyond this type of CAI to a type that might be suitable for a wider range of material than mere cramming of facts and might respond more sensitively to the abilities and difficulties of the pupil. It might depend not only on pre-stored material but also on a programmed capability to generate new material, using natural language, from a stored 'knowledge base'. These applied-research dreams are in the present author's view singularly untimely and unpromising. Taking into account the very large computer capacity and programming skill needed at present to achieve computer use of natural language on even a very small know- ledge base, and adding all the difficulties of structuring and accessing a larger knowledge base and monitoring pupil performance, one can only conclude that the nineteen-seventies are not the right decade in which to begin researches aimed at applying such techniques to the teaching of any body of knowledge big enough to be of practical interest. To avoid misunderstanding however, one should make clear that basic research on developmental psychology by the methods of category C would not on this argument be excluded.

Category B

The balance between numerous disappointments and certain solid achievements from work in categories A and C is, perhaps, typical of scientific research as a whole. It indicates only that these areas of research are not in one of those conditions of exceptional fruitfulness when everything seems to be going right. By contrast, the sense of discouragement about the intended Bridge Activity of category B, centred upon Building Robots, seems altogether more widespread and profound, and this raises doubts about whether the whole concept of AI as an integrated field of research is a valid one.

Quite possibly the sense of discouragement is greater in category B because still greater expectations have been sensed and voiced in this category than in the others. Some workers in the field freely admit that originally they had 'very naive' ideas about the potentialities of intelligent robots, but claim to recognise now what sort of research is realistic. In these circumstances it might be thought appropriate to judge the field by what has actually been achieved than by comparison with early expectations. On the other hand, some such comparison is probably justified by the fact that in some quarters wild predictions regarding the future of robot development are still being made.

When able and respected scientists write in letters to the present author that AI, the major goal of computing science, represents 'another step in the general process of evolution'; that possibilities in the nineteen- eighties include an all-purpose intelligence on a human-scale knowledge base; that awe-inspiring possibilities suggest themselves based on machine intelligence exceeding human intelligence by the year 2000; when such predictions are made in 1972 one may be wise to compare the predictions of the past against performance as well as considering prospects for the realisation of today's predictions in the future.

It certainly seems that early enthusiasm for programming and building a robot that would mimic human ability in a combination of eye-hand co-ordination and 'commonsense' problem solving has ended up gravely disappointed. The large amount of computer time needed to distinguish between everyday objects of markedly different shapes against a far from noisy background has been most discouraging; the engineering complications required to achieve eye-hand co-ordination (not of human standard but similar to what an octopus can learn) have been repellingly formidable. Reports from the world's different centres for this work are all disenchanting.

Some able research workers, who from their beginni,ng in the field regarded Building Robots as a precarious or even 'crazy enterprise but nevertheless were attracted to participate in such a long-shot or even 'shot in the dark' activity, have felt themselves driven now to recognise that the difficulty of achieving good hand-eye co-ordination in quite simple problem situations has proved unexpectedly great, and seems to hold out negligible hope of approaching human levels of achievement. In these circumstances, many good computational theorists are emphasizing that productive research on 'robot reasoning' (or, essen- tially, commonsense problem solving) does not necessarily need the physical presence of an eye-hand machine. This line of argument then branches in two directions, one leading to work properly in category A (directed to automating the solution of such problems as may arise in practical fields of application), and the other (which is our concern here) leading to programs for problem solving in an abstract 'play' situation: for example, in an abstract table-top world with data fed in not as television images but as statements about the positions of blocks on the table-top; or in a similarly defined chessboard or puzzle situation.

The 'Category B' research work on problem solving in these abstract play situations has produced many ingenious and interesting programs. A fair description of the success of these programs seems to be that they are effective when and only when the programming has taken into account a really substantial quantity of human knowledge about the particular problem domain. Just as in category A, the pure mathematical- logic methods suffer defeat at the hands of the combinatorial explosion, and have to be replaced by 'heuristic' methods. Some very interesting researches have been carried out to develop 'general' problem-solving programs, and such work can be of research interest to psychologists, but the performance of these programs on actual problems has always been disappointing. Students of all this work have generally concluded that it is unrealistic to expect highly generalised systems that can handle a large knowledge base effectively in a 'learning' or 'self-organising' mode to be developed in the 20th century.

Those wishing to decide as between this view and the quite opposite views of the 'awe-inspiring' future mentioned earlier can quite helpfully study the state of the art on chess-playing programs. This is partly because chess is a complicated enough game so that in a contest between a computer and a human player the computer's advantages of being able to calculate reliably at a speed several orders of magnitude faster need by no means be decisive (the number of possible positions being incomparably greater) and so there is real interest in whether or not they are outweighed by the human player's pattern-recognition ability, flexibility of approach, learning capacity and emotional drive to win. Another good reason for investigating chess-playing programs is that the long-term interest of the big international computer manufac- turers in bringing about some spectacular achievement of 'machine intelligence' against such a well developed human intelligence as an able chess player, in order to assist in selling more generally their products' potentiality for superseding human intellectual activity, has been an incentive to the devotion of quite considerable resources to producing an effective program.

It is interesting to consider the results of all this work some twenty- five years after the researches aimed at chess-playing programs began: unfortunately these results are discouraging. The best programs play chess of only 'experienced amateur' standard characteristic of county club players in England. Chess masters beat them easily.

More important, progress on constructing chess-playing programs has been made solely by heuristic methods. The programs seek to maximise in what may be called the foreseeable short term a complicated 'evalua- tion function'; this function, constructed entirely from human knowledge and skill, represents an evaluation of a position, depending on large numbers of different measurable features of it with different weights attached to them. What relatively modest success the programs have achieved is a measure primarily of human skill and experience in the concoction of this evaluation function. The computer's contribution is primarily rapidity in looking a few moves ahead and finding a line that produces a position change good on the basis of that evaluation. The 'intelligence' contribution is human; what the computer offers is its speed, reliability and biddability. By contrast, 'learning' programs are not considered applicable to computer chess at present.

To sum up, this evidence and all the rest studied by the present author on AI work within category B during the past twenty-five years is to some extent encouraging about programs written to perform in highly specialised problem domains, when the programming takes very full account of the results of human experience and human intelligence within the relevant domain, but is wholly discouraging about general- purpose programs seeking to mimic the problem-solving aspects of human CNS activity over a rather wide field. Such a general-purpose program, the coveted long-term goal of AI activity, seems as remote as ever.

In thus regretfully noting the remoteness of this goal we must not, eye-hand co-ordination' and 'scene analysis' capabilities that are much studied in category B represent only a small part of the features of the human CNS that give the human race its uniqueness. It is a truism that human beings who are very strong intellectually but weak in emotional drives and emotional relationships are singularly ineffective in the world at large. Valuable results flow from the integration of intellectual ability with the capacity to feel and to relate to other people; until this integra- tion happens problem solving is no good because there is no way of seeing which are the right problems. These remarks have been included to make clear that the over-optimistic category-B-centred view of Al not only fails to take the first fence but ignores the rest of the steeple-chase altogether. It will suffice, however, to judge the work on its own rules and its own aims in order to conclude that the attempt to construct a true Bridge between categories A and C is not succeeding.

Postscript

It is only fair to add at the end of this section on 'Past Disappointments hat some workers in the field would have agreed with the view just expressed until the appearance less than two years ago of an exception- ally good PhD thesis [1] on a computer program for use of natural language, since when they have felt resurgence of optimism about the coherence and viability of the concept of integrated Al researches. It is important to analyse how this reaction has come about and how far such resurgence is justified.

The thesis is exceptional in more than one way. The style in which most papers on Al research are written is depressingly turgid or jargon- dominated and almost makes the authors appear antagonistic to the special human gift for relating to, and communicating with, other people in an imaginative way (as if such authors appreciated only those human capabilities which they seek to mimic in robots!) By contrast, the thesis is a pleasure to read, the author's substantial research achievement and attractive personality being communicated infectiously by his style of writing. His gift for language and communication has without doubt contributed to making his researches widely known all over the world, as well as having contributed to the success of the actual linguistic analysis underlying the development of his program.

This analysis is the strongest of those mentioned above under category C as having developed far beyond the transformational- grammer approach of general linguistics into new methods for machine interpretation of natural-language sentences within a limited universe of discourse, that make very substantial use of stored knowledge concerning that universe. Full use of such knowledge is regarded by the author of the thesis as an essential ingredient of the success of his approach, to which however the penetration and originality of the analytical methods he introduced has also made a vital contribution.

Specifically, the universe of discourse is an 'abstract table-top world' and in the lengthy, and now rather famous, conversation' between the author and his program the program accepts, and is deemed to have carried out, certain commands to perform well defined block-stacking operations, while it queries commands that are impossible or ambiguous. The program deals similarly in answer to questions put to it regarding the present and past states of the table-top world. In constructing the program, two of the high-level programming languages referred to under category A above were used: one to program the events in the

[1] 'Understanding Natural Language' by Terry Winograd, MIT. Published in UK by Edinburgh University Press, 1972.

abstract table-top world and one to perform the linguistic analysis. The thesis well illustrates the value of these high-level languages.

The contribution of brilliant presentation and deep originality has made this thesis deservedly influential on workers in most areas of AI research. The biggest and clearest influence is on psycho-linguistics itself (work in category C), where the studies have been in large part responsible for a movement towards viewing the processes by which the human CNS responds to and uses language as Semantics-controlled (or knowledge-controlled) even in their fine structure.

Many workers in category B (the Bridge Activity of Building Robots) have at the same time felt encouraged by this thesis: its program, after all, can properly be described as a Robot 'with whom the author converses', and Building this Robot has succeeded in its aim to an extent undreamt-of in the unrewarding world of eye-hand machines. The program seems furthermore to open up more general possibilities of conversing with Robots by means of natural language. There is even a suggestion of an ultimate link through to work in category A if these new studies could revive prospects for achievement of machine translation.

To such somewhat Over-generalised euphoria it is necessary to respond, however, with certain cautionary reservations beyond the banal comment that one swallow does not make a summer. Outside the psycho-linguistic area where the thesis has truly helped to establish a new direction of research, suggestions for possible developments in other areas that can properly be inferred from the studies are rather discouraging. Thus, the studies show how the complex problems involved in computer use of natural language are rendered far more complex by the need to interact in detail with systems for structuring and accessing the necessary knowledge base. For an extensive universe of discourse this could put such developments out of practical reach.

In practice, a large computer together with very sophisticated pro- gramming using subtle new programming-language developments was found just sufficient to make slow conversation possible on the very limited material represented by the abstract table-top world; material restricted enough, for example, to allow resolution of ambiguities in natural language sentences by classical theorem proving techniques. Extension of the methods used to a much wider universe of discourse would be opposed violently by the combinatorial explosion.

Accordingly, the present author's view of the definite (though not overwhelming) promise of work in categories A and C, and of the general failure of work in category B to establish effectively the unity of AI research as a whole, remains unmodified by careful study of one particular piece of work of a very remarkable character. This postscript to the section on 'Past Disappointments', explaining this, has been included because emotional response to such work is very natural and desirable but needs to be integrated properly with an intellectual appraisal of where its significance primarily lies. The thesis is, of course, a triumph of human intelligence, and human intelligence can respond to it most correctly by recognising its main contribution as being to aspects of how the use of language by the human race has developed and of what processes within the human CNS that use may involve.

3 Future Possibilities

However controversial may be an analysis of the past, a forward look towards the different possibilities that the future may hold in Some area must be more controversial still, especially when made in a report to a body whose decisions can have a very substantial influence over a certain part of that area's future (specifically, the British part). Contro- versy has its unattractive as well as its attractive features, but it cannot be avoided in a period when the abilities of Scientists jointly to arrive at wise decisions on research policy is publicly regarded as being on trial.

Research on Al in Some other countries may be funded by military agencies (AR PA in USA) or by other mission-orientated public bodies. With this type of funding it is common for Scientists to 'close their ranks' and avoid public disagreement among themselves, in the hope that the total funds available for science may thus be enhanced to an extent that may outweigh any harmful results of a distribution of those funds determined on the basis of insufficient scientific discussion. Such optimism would be unjustified in a poorer country such as Britain, while the alternative approach here advocated accords with the desire to 'keep our Al research civilian' expressed to the author by various British workers in the field. This Suggests that decisions within the UK should be taken only after carefully contrasting and comparing different informed views of the research field's future available to SRC. Thus, due weight should be given to the principle 'Heterarchy not Hierarchy' (an Al maxim of considerable soundness concerned with file Structures).

To the required debate this report's contribution consists not of any detailed costed recommendations, but of certain general considerations based on the analysis of the past given in Section 3 as well as some thoughts about the present and future now to be presented. After omission in this published version of all specific comments on British research work in the field, these consist essentially of an attempt to look to the field's scientific future in the world as a whole, subject to the proviso that any speculations beyond the end of this century are regarded as too uncertain to justify mentioning them or basing on them any present research decisions whatever.

It is assumed that more precise policy formulation and detailed decisions on projects will stem from the normal working of the machinery of the Science Research Council and its Boards and Committees, influenced to some extent by such special reports as may be available, including the 1972 Policy and Programme Review of the SRC Comput- ing Science Committee and also the 1972 report of a joint SRC/MRC panel on Neurobiology, as well as the present 'personal view' of Al as a research field.

The next twenty-five years

This 'personal view' which saw the past twenty-five years of Al research as having developed a 'bimodal' distribution of achievement, with some respectable (though not as yet lofty) peaks of achievement in categories A and C but relatively speaking a valley between them in category B, looks ahead to still greater bimodality, amounting practically to fission, arising during the next 25. Specifically it foresees, whether within category A or category C, certain research areas making very substantial further progress, coupled in each case with the forging of far stronger links to the immediate field of application than to the supposed bridge activity B. Rising confidence about the work's relevance within the associated field of application may add prestige and thence strength to such an area of research, while continued failures to make substantial progress towards stated aims within category B may cause progressive loss of prestige, from which a diminution of funding will ultimately follow even where scientific claims are not always subject to full scientific scrutiny. In due course the overriding significance of the links between each research area and its field of application will rupture the always fragile unity of the general concept of Al research.

For example, in the technological applications within category A the work will become increasingly specialised, in accordance with the common experience that Al researches are successful to a degree closely correlated with the degree of use made of detailed knowledge of the problem domain. Techniques for Advanced Automation can now be expected to move forward fastest where research workers concentrate upon practical problems, acquiring for the purpose good detailed knowledge of the technological and economic contexts of the problems chosen. Benefit to both sides will flow from closer integration with control engineers, who have deep experience not only of the practical but also of the financial and sociological questions arising in automation, and can conversely learn much that is of value to them from experts in advanced computational theory. That theory itself (including the development of new programming languages) can in the meantime forge ahead through a combination of the spin-off from Advanced Automation developments and of the activities (serving far more than Al research) of general computing-science laboratories.

Recently the Japanese government announced a 40M research programme oriented very closely along these lines, aimed at the automation of factory assembly processes, which among the areas within category A mentioned in section 2 may well be one of the most promising. Another clue to how Advanced Automation developments may proceed comes from recent experience in the sister field of Com- puter Aided Design (CAD): generalised researches dominated early CAD work, but later on several specific industries developed form of CAD very specialised towards their own problems. This example is mentioned mainly as an analogy to how Advanced Automation may come to experience a similar degree of fragmentation; however, we should also remember from section 2 that CAD is itself an area where advanced computational theory may be found to have a role, especially in the integration of subsystem designs. Close relationships between Al theorists and CAD workers (such as here and there have already come about) are one prerequisite for those developments.

A similar outward-looking trend is expected in the mathematical and scientific applications of researches within category A; in mathematics, for example, from utilisation of far more detailed observation of 'how mathematicians actually prove theorems'! The structuring and utilisation of scientific data bases is another area where good results depend on detailed study of the data's special characteristics. The one part of that field with which the present author has been closely involved, as Chairman of the Steering Committee for the Experimental Cartography Unit of NERC since its inception, affords a good example of this: the structuring of geographically located data is found to demand quite specialised techniques, closely related to the cartographic character of the output. Another good example from the scientific sphere is the chemical-structure work described in section 3.

It is not to be expected that, in all the areas within category A listed in section 2, striking successes will be reached during the next twenty-five years. The view here proposed is rather that the chances of success in any one area will be greatly improved through close integration of the researches with the field of application.

Substantial advances are at the same time expected within Category C, where success will again be related to how closely the work is linked to the fundamental associated disciplines of psychology and neurobiology. Computer-based CNS studies can from experimental psychology gain greatly through more substantial use of the extensive data available, eg on reaction times, on pattern-recognition abilities, and on the types of errors made in different tasks. From modern neuro- biology they can derive a valuable appreciation of the detailed evidence to the effect that the CNS uses 'specialised hardware' very economically to perform significant and important tasks. Computer-based studies have a role to play in analysing how some of this hardware may function, and conversely can derive a salutary reminder that simulation of the extraordinary self-organising capabilities evolved by the human CNS may actually be unattainable through ingenious software developments.

Conversely, psychology and neurobiology will benefit to an extent closely related to how far Computer-based CNS researches behave as if they felt integrated within one or both of those fields. Psychologists and neurobiologists may especially gain increased appreciation of the value of computers for theorising about complex systems and for making sense of complex masses of data, while integration at a more funda- mental level can be expected to follow. With the resulting growth in understanding of the human CNS, respect for it may, perhaps, grow to an extent that will reduce the ebullience characteristic of past predictions of Al possibilities.

In the meantime, the intended Bridge Activities within category B may well have been found increasingly disappointing as achievements from Building Robots of the more generalise types fail to reach their more grandiose aims. On the other hand, some robots designed primarily as Computer models for Comparison with experiments on how the human CNS performs linguistic or problem-solving tasks will become integrated with work in category C, while others aimed at practical tasks, related eg to engineering assembly, will become integrated in Category A.

These processes are expected to bring about, at a slow but increasing rate, the fission of the field of Al research predicted at the beginning of this brief attempt at looking into the future. That attempt may now be Concluded with the observation that such a broad-brush view, un- specific in matters of detail, is possibly all that can properly be attempted over time-spans as long as twenty-five years.