NOT FOR QUOTATION
WITHOUT PERMISSION
OF THE AUTHOR

PAPERS FOR TASK FORCE MEETING ON FUTURE AND IMPACTS OF ARTIFICIAL INTELLIGENCE 15-17 AUGUST 1983

Robert Trappl
Tibor Vasko
Eds.

August
CP-83-36

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
2361 Laxenburg, Austria

ACKNOWLEDGEMENT

Thanks should be extended to the Austrian Federal Ministry for Science and Research for sponsoring and supporting this project.

PREFACE

IIASA's Clearinghouse activity is oriented towards issues of interest among our National Member Organizations. Here, in the forefront, are the issues concerning the promise and impact of science and technology on society and economy in general, and some selected branches in particular.

Artificial Intelligence (AI) is one of the most promising research areas. There are many indications that the long predicted upswing of this discipline is finally in the making. A recent survey had Nobel-laureates predict that the most influence in the next century will be made by computers, AI, and robotics. Already, at present, "expert" systems are emerging and applied; natural language understanding systems developed; AI principles are used in robots, flexible automation, computer aided-design, etc. All this will have an, as yet, unspecified social and economic impact on the activity of human beings, both at work and leisure.

It certainly takes interdisciplinary and cross-culturally based studies to enhance the understanding of this complex phenomenon. This is the aim of our endeavors in the field which is in excess of our duty to pass useful knowledge to our constitutency. We think that IIASA, cooperating in this respect with the Austrian Society for Cybernetic Studies (ASCS), can develop some comparative advantage here.

This publication contains papers written by leading personalities, both East and West, in the field of artificial intelligence on the future and impact of this emerging discipline. We hope that the meeting, where the papers will be discussed, will not only identify important areas where the impact of artificial intelligence will be felt most directly, but also find the most rewarding issues for further research.

Robert Trappl and Tibor Vasko

CONTENTS

IMPACTS OF ARTIFICIAL INTELLIGENCE

Margaret A. Boden*
(UK)

1. DEVELOPMENT IN NEXT 10 YEARS

1.1 Core Research Areas and Likely Results

Several core research areas are likely to make solid progress within the next decade. Each of these is already being worked on in various countries, and progress does not depend upon the success of Japan's ambitious "Fifth Generation" project (though it might be accelerated by associated hardware and software developments).

One is low-level vision, based on techniques using parallel hardware and cooperative processing. Current "connectionist" research in this area differs in its approach from work on 2D pattern-recognition by "property-lists", and

*University of Sussex, England. I am grateful to the SSRC for support while writing this paper. I alone am responsible for the views expressed in it.

from the top-down "scene-analysis" of 3D scenes. Based on detailed studies of image-formation, it is able to extract from the ambient light information about 3D-features (such as shape, depth, texture, and surface-orientation) which in previous approaches could have been computed only, if at all, by way of high-level knowledge of the expected scene. Some of this work is being done in the context of human psychology and neurophysiology, some in a more technological context. Dedicated (massively parallel) machines are being designed for this research, and major advances depend upon such hardware.

A second area in which we can expect significant progress is robotics. This includes problems of movement control, trajectory planning, and visumotor coordination (and will take advantage of advances in low-level vision). As in the case of vision, some projects will rely on "artificial" means to ensure (such as light-stripes for automatic welding-machines, capable of recognizing different sorts of weld-joint and guiding the welder accordingly), while others will relate more closely to psychophysiological theories of motor control and visumotor coordination in living organisms.

Knowledge-based "expert" sustems will multiply enormously in the next decade, not least because there is considerable commercial interest in them. Different domains of human expertise may require different approaches to knowledge-engineering. In domains less fully covered by an explicit scientific theory, it may be easier to extract knowledge from human experts who are competent but who have not yet achieved ghe "intuitive" mastery of the domain which topflight experts enjoy. The latter give the right answer more often, but cannot easily introspect their reasoning processes, which happen very fast and are not consciously accessible. The former takes time to come to a decision, often consciously weighing distinct considerations against each other and verbally identifying areas of unclarity. Domains (such as medical radiology) which depend on the comparison and interpretation of complex visual images are especially difficult to automate, since low-level visual processes are not open to voluntary inspection or control. Indeed, experts often give highly misleading advice about how they may be carrying out the relevant comparisons (eye-movement studies show, for instance, that expert radiologists do not scan X-ray photographs in the way they say that they do). In tandem with the increasing experience of AI-trained knowledge-engineers, further psychological studies of the organization of knowledge in different domains should be useful.

Research on expert systems will also focus on the computational architecture required to deal with large, complex, knowledge-bases. Current systems are relatively simple and inflexible, and restricted to very narrow domains. They can be incrementally improved, but only up to a point. Eventually, the interactions between the increasing number of independently-added rules become too difficult to control, and the system's reliability and intelligibility are jeopardized. Current systems have no access to higher-level representations of the knowledge domain and their own problem-solving activity (see below). Special problems arise if a system has to work in real-time, where unexpected events can require quick switching from the current activity to some other. The next ten years will see some general work on powerful IKBS architectures (as well as the production of more examples of specific commercially useful systems), including parallel-processing devices.

Progress can be expected also in natural language processing, both of individual sentences and of texts. Key issues include syntactic parsing, the integration of syntax with semantics, and the understanding of connected text. Machine translation could in principle benefit from advances both in single-sentence parsing and in text-analysis.

Current work on parsing is motivated both by theoretical (linguistic) interests, and by the hope of improving the man-machine interface so as to make it possible for non-specialist users to communicate with programs in (some reasonable subset of) natural language. Where a program is used for some specific purpose, semantic factors can be more readily used to help in the parsing and disambiguation of queries and instructions input by the user. Verbal interchanges about lunar geology, or about airline reservations, are already reasonably "natural" because of the exploitation of semantic constraints, and further domain-specific semantics will be developed over the next decade. More generally applicable (theoretical) research will continue into the best point at which to use semantics in parsing: from the beginning of the sentence, or spreading out from the middle, or only after an initial parse of the entire sentence?

Text-analysis programs can already give a precis of most short news-stories about specific topics (such as earthquakes, hi-jackings, and road-accidents). But they rely on rigid, pre-programmed schemas, which provide the semantic skeleton of the types of stories concerned. Some recent research is aimed at enabling a text-analysis program to learn new schemas for itself, to integrate one schema with another so as to understand a story combining both, and to use a given schema to reason analogically in an unfamiliar context. A high degree of success cannot be expected within the next ten years, but our understanding of the relevant problems should be advanced.

A variety of educational applications are already receiving attention. Some are focussed on particular curricular subjects, and require both a model of the theory of that subject and a model of the student's knowledge of it (which varies in level and in organization, from person to person and from time to time). Others are less specific, and aim to use AI-based techniques to improve the pupil's attitude to intelligence in general. There is some evidence that both normal and handicapped students can attain greater self-confidence and intellectual achievement by experience with these specially-designed programming environments. Controlled research into the classroom effects of AI-based systems has recently been initiated, and this can be expected to bear fruit within the next decade.

An extremely important area, which is increasingly being studied because of recent hardware developments, concerns the computational properties of large parallel systems. At present, we understand very little of the potential and limitations of such systems. Some of the connectionist work mentioned above suggests that cooperative processing may have some highly surprising properties. For example, the number of individual processors required to make the "human" range of visual shape-discriminations appears to be markedly less than one would naturally assume. Again, making an connectionist system stochastic rather than deterministic *improves* its chance of finding an optimal solution. The computational properties of parallel machines will not be well understood for a long time, but experience with these new systems in the near future will doubtless lead to some advance.

Five topics studied recently in AI, and which will be further developed over the next ten years, are non-monotonic reasoning, naive physics, self-updating memory, creativity, and machine-learning. I shall refer to these difficult problems in the section on "Long-Range Research" below.

1.2 Impacts on Other Sciences and Technologies

The impacts of AI on other technologies will include many different examples of applications to individual problems. For example, an old factory chip is being designed using AI techniques of pattern recognition. Given advances in

VLSI, instruments and products of many different kinds will come to include chips whose design makes use of AI methods. Any commercial-industrial task that could benefit from even a limited degree of intelligence could in principle be performed better with the help of AI, so that the technological applications of AI will be extremely diverse.

Turning from technology to science, AI will influence other sciences in their general philosophical approach as well as their specific theoretical content. Indeed, psychology and (to a lesser degree) biology have already been affected by computational ideas. And, contrary to what most people assume, AI has had a humanizing effect in psychology. The behaviorists had outlawed reference to "mind" and "mental processes" as unscientific and mystifying, but AI--based as it is on the concept of ***representation*** --has made these concepts theoretically respectable again.

AI's influence will be especially strong in the psychology of vision and language, and (as noted above) it is likely that robotics will engage with the psychophysiology of movement. Psychological research will feed back into AI--for example, insofar as psychologists arrive at a better understanding of the organization of knowledge their work may be useful in designing computerized expert systems. Cooperative interdisciplinary research should be encouraged: the institutional separation of empirical psychology and AI or computer science has hindered fruitful collaboration between these groups.

1.3 Social Implications

Social impacts will be of various types. First, there will be effects on individuals and institutions brought about by specific applications of AI, such as expert systems for medical diagnosis, legal and financial advice, or educational help.

These programs will not merely provide a service (whose adequacy should be very carefully monitored), but will very likely change the social relations of the profession or institution concerned. For example, if general practitioners, or nurses, can use an AI program to aid in various aspects of patient-care, the social image of the specialist physician may be profoundly affected. (And legal responsibilities for medical decisions may be assigned in a way very different from today.) Likewise, legal programs may undermine the status of lawyers, and alter the nature of their work. In both cases, while the mystique of the human experts may be lessened, their opportunity for exercising their specifically human powers may be increased.

The general public might come to be less dependent on human experts than they are today. Reducing the power of professionals such as doctors, lawyers, and teachers, would certainly have advantages. But replacing human professional advice by computer programs is dangerous to the extent that AI-systems in public use are inadequate--and/or ill-understood. Systems that have taken several man-years to develop (and whose original programmers may be retired, or dead) are often very difficult to evaluate or alter, because even computer scientists do not fully understand how they work. (It follows that attention should be given to methods of perspicuous documentation, to help make clear what it is that a given program is actually doing, and how.)

A second type of social impact concerns general social trends brought about by applications of AI and IT. These include changes in the proportion of the workforce in service and leisure industries changes in the division of labor and sexual roles, and changes in general lifestyles and patterns of interaction.

For example, males will be increasingly freed to take up jobs in the "caring" professions (such as nursing, education, and social welfare). This could change the general evaluation of emotionality in the masculine role, an effect that could also be encouraged by men's having increased leisure time to spend with family and friends. Such an effect could be liberating and humanizing, leading to a moe convivial society than we have today.

But other potential consequences of AI point in the opposite direction. The widespread use of home terminals, for instance, threatens to have an isolating influence even more powerful than that of television. If people are encouraged to work, and to shop, from their sitting-rooms, there may be unfortunate psychological effects in terms of personal stress and loneliness. Community computer-centers could offset these effects to some extent, providing a social meeting-place outside the confines of the home and nuclear family. Some writers even predict that commercially avalable (and highly profitable) AI-systems will be heavily used not only in task-oriented ways, but as surrogates for human contact. On this view, the strong tendency to anthropomorphism which most of us share will result in patterns of interaction being skewed away from human beings, and towards quasi-human computer systems (with naturalistic "voices", and sometimes even "bodies"). Although such forecasts grossly underestimate the technological difficulties involved in building programmed "friends-off-the-shelf", they do suggest that human interactions could be impoverished to some degree in the future.

These contrasting examples show that widespread application of AI will have subtle, and varying, influences in society. Moreover, AI could foster a general view of humanity as either "mechanistic" or "non-mechanistic", depending on how it is interpreted by the public. The commonest interpretation is that AI presents us as "mere machines", with no free choice or moral responsibility. Since this image of man could have socially pernicious effects, people should be helped to understand that it is fundamentally mistaken. The education or computer literacy discussed below could help here. More generally, we should start thinking now about what the optimal social arrangements might be for a post-industrial society.

1.4 Economic Aspects

The economic impacts will be far-reaching. Traditional manufacturing and clerical-administrative jobs will be decimated. But new jobs will be created: some directly connected with new technology (like computer engineers and programmers), others made possible because people are freed to devote their time to services (caring professions, education, leisure). Whether there will be enough new jobs to compensate for the loss of old ones (as has always happened in the past, at least eventually) is however unclear, for AI can potentially apply to all jobs where personal human contact is not essential. New methods of work-sharing and income-distribution will have to be worked out (with income not necessarily being closely linked to jobs). Radical structural changes in society are likely, and the transition phase will not be easy.

2. LONG-RANGE AI RESEARCH

There will be "more of the same", in that the areas mentioned above will provide perplexing problems for many years to come. Especially hard problems include learning, high-level vision, naive physics, and abstract work in computational logic—including the development of a taxonomy of representations and computational processes, showing the potential and limitations of distinct types.

I referred earlier to short-term research on IKBS architecture. But the deep problems involved in the organization and control of large knowledge-bases will not be solved within a decade. This is so quite independently of the fact that parallel machines may support forms of inference radically different from those implemented today.

For instance, expert systems are at present unable to explain their reasoning except by "backwards-chaining": giving a resume of the chain of inferences (rules) which led up to their conclusion. They cannot relate their conclusion to the domain in general, nor rely on an overview of the problem to assess the relative theoretical reliability of different hypotheses (probabilities are of couse built into rules, but are assessed for each rule individually or in relation to a small number of other specific rules). Nor can they monitor and adjust the structure of their own problem-solving, for they have no high-level representation of it. They are unable, too, to integrate different knowledge-domains, and to use concepts and patterns of inference taken from one domain to reason (analogically) in another. Nor can current systems explain their conclusions differently to different users, taking account of the specific user's knowledge. The user can ask for a "deeper" explanation (a more detaoled inference-resume), but the program has no user-model in terms of which to adjust its explanations to the human's particular range and level of knowledge. For this reason also, the pattern of interaction between user and system is at present very limited. The user cannot offer his own conclusions for comment and criticism, for example, as students can do with human teachers.

All of these abilities which current programs lack will need a richer understanding of the structure and content of different knowedge-domains. Some of the projects mentioned elsewhere (such as research on naive physics) are highly relevant to many domains, and psychological research into human reasoning processes could be useful also. This is just one illustration of the fact that theoretical and empirical research may be needed for radical improvements in technological applications.

The need for a model of the user's knowledge also delays advances in educational programs such as those mentioned above. In principle, computer-assisted-instruction based on AI techniques could be highly flexible, and subtly attuned to the student's particular strengths and weaknesses. But this requires that the program be equipped with a representation of the rich content and inferential organization of human knowledge in the relevant domains (which in turn requires psychological understanding of a high degree). To achieve this will be a difficult task, for the long range rather than tomorrow.

A special case of human knowledge is "naive physics", one's everyday knowledge of the properties and behavior of different sorts of physical substances, and the nature of the causal relations between them. This knowledge enters into vision and motor control, and also into natural language. For example, a language-using program would have to understand the differences in meaning between verbs such as *pour*, *flow*, *spill*, *drop*, and the like, if it were to give instructions or understand texts about activities dealing with liquids. Similarly, a robot capable of seeing that a container was just about to spill its contents onto the object below, and of adjusting its movements accordingly, would need some representation of the behavior of fluids. (It might of course be programmed to halt movement if it saw an unexpected patch appearing at the rim of the container, but that is a different matter). Very little work has been done on these issues so far, and they are likely to provide a challenge for many years.

Another topic that is likely to receive much attention in the future is truth-maintenance using non-monotonic increasing. Traditional logical systems are monotonic, in the sense that propositions are proved one and for all: if a proposition has been inferred as true (or as false) on one occasion, its truth-value cannot change thereafter. But in commonsense reasoning, a proposition may be taken as true for very good reasons, but later found (or inferred) to be false. AI systems dealing with complex problems involving incomplete knowledge similarly require non-monotonic reasoning, and new canons of inference are needed to control such knowledge-systems, and to prevent them from falling into absurdities.

The development of self-updating computer memories is closely related to the issue of non-monotonic reasoning. But in addition to allowing changing truth-values, such a memory-system needs to be able to make inferences of many different sorts on being told "one" new fact. Human beings do this every day. For example, if one is told that an acquaintance is a supporter of a particular political group, one's internal representation of that person may "unthinkingly" alter in many different ways (and one's attitudes and future behavior regarding the person will be influenced accordingly). An intelligent program, presented with new information ought to be able to do the same kind of thing. Although some preliminary AI-work has been done on this problem, it is not yet well-understood.

"Computer-aided design" is typically thought of as involving the graphical display of precise three-dimensional specifications of various products (from machine-tools, through cars, to buildings), taking into account a wide range of values of many parameters. But a recent form of computer-aided design involves suggestion rather than specification, in the sense that the design-program originates novel ideas--ideas which are not merely quantitatively different from previous specifications. For example, heuristic programs are already being used to suggest novel experiments (described at the intramolecular level) in genetic engineering, or to help design new sorts of three-dimensional silicon-chips. These programs were developed in tandem with a closely similar system that originates interesting mathematical ideas from a basis of elementary set-theory.

The potential of systems like these should be further explored. The computer-modelling of creative thinking will require long-term research, especially with respect to domains whose crucial concepts cannot be so readily defined as the concepts of molecular biology, chip-circuitry, or set-theory. We need a better understanding of how conceptual structures (and the inference-patterns associated with them) can be explored, represented (on varying levels of abstraction), compared, and indexed. Highly interconnected processing networks may turn out to be useful for mediating "unexpected" conceptual associations. But association is not enough: associations need to be evaluated and controlled, and integrated into previously existing cognitive structures. As yet we have little idea how to do this so as to model creative thinking.

Machine-learning is a pressing problem for the future. If a program cannot learn for itself, its development is limited by the time and ability of the programmer to provide it with new information and ideas. The system should be able to induce regularities for itself. Some progress is being made in enabling computer-systems to learn about specific aspects of particular domains. But "open-ended" learning, where what is to be learnt is not defined beforehand, is especially intractable.

Some connectionist workers have presented a general "learing algorithm", claiming that the input of large numbers of instances (of visual scenes, for example) could enable a connectionist system to learn to recognize the structure of the input-class, irrespective of what that structure is. However, to say that something can be done in principle is not to provide a practical usable way of doing it. These claims cannot be further explored until suitable hardware is available (dedicated machinery is currently being designed).

In general, the properties of parallel computation will be a focus of research in the long term as well as in the next few years. It remains to be seen whether the Japanese hopes concerning VLSI and PROLOG (thought to be especially suited to a parallel architecture) will be achieved. But massively parallel hardware will increasingly become available, and will enable AI to progress in ways that are impossible using traditional types of machine.

3. WHICH AREAS WOULD YOU SUPPORT MOSTLY, AND WHY?

Areas can be supported for their intrinsic scientific interest and/or for their social usefulness. Low-level vision and robotics include work qualifying on both counts, and solid progress is likely within the next ten years. Naive physics is less well developed, but is likely to be important not only for advanced robotics but for language-understanding too.

Research in computational linguistics and speech-understanding merits support for its practical uses and theoretical interest. User-friendly programming environments and man-machine interfaces require natural-language "front-ends". Although these do not need to handle ever linguistic subtlety, so can ignore many problems that are theoretically interesting, there is still much room for improvement.

Support for IKBS should encourage basic research into general issues of system-architecture and non-monotonic reasoning, rather than leading to the proliferation of the relatively simplistic systems available today. This is a long-term project, but essential if AI systems are to be widely used in decision-making contexts.

More research is needed on the educational applications of AI. A few groups have already started to study the effects of giving children (of various ages) access to the "LOGO" programming environment in the classroom. Some experience is also being gained in using LOGO to help gravely handicapped children. As noted above, preliminary results suggest that this programming environment helps both normal and handicapped children to express and develop their intelligence, emotional relations, and self-confidence. As with new educational methods in general, it may be the enthusiasm and commitment of the pioneers involved which is crucial. Carefully controlled studies in a range of schools, involving a range of teachers, are needed to evaluate the claims that have been made in this context.

Psychological research into the organization and use of knowledge in different domains could contribute usefully to appliations of AI. As mentioned above, both educational and "expert" AI-programs will need an internal model of the student-user to enable them to interact in flexibly appropriate ways.

The general problem of computation in parallel systems has been referred to several times already. It is clearly an important area. For a few years yet, we can expect exploration rather than exploitation. But this exploration of the potential and limitations of such systems is essential.

Funds should also be made available for combatting the ignorance and sensationalism that attends AI today. Research on friendly programming environments, and on interactive "programmer's apprentices", should be supported. This involves not only work on natural-language interfaces, but also psychological studies of how people learn to program and (what is not the same thing) how they carry out and interpret an interaction with a quasi-intelligent program. It may be that certain words or phrases, and certain ways of structuring the interaction, help users to appreciate the specific limitations of the program they are using, and remind them that they are interacting not with a person but with an artefact. Some universities have already begun to develop programming environments and exercises designed primarily to awaken naive users to the potential ***and the limitations*** of AI-programs, and the general educational value of such experiences should be explored.

One might ask why widespread ignorance about AI *matters.* Part of the answer is obvious: in a society where most jobs involve access to computerized facilities making use of AI techniques, individuals without any understanding of AI will be at a disadvantage (and the more of them there are, the more social unrest is likely). But there is another important consideration, which can be illustrated by an advertisement recently shown widely on British television.

The advertisement showed six people sitting at six computers, each sold by a different manufacturer. The "voice-over" message said something to this effect: "We provided details of the performance and cost of six different computers to the six computers themselves, and asked them to choose the best. The X chose the X (I shall not advertise the firm further by giving its name here)--and so did all the others. It makes you think that a person ought to choose the X too."

This type of persuasion is pernicious, for it deliberately obscures the fact that each machine was running the same choosing-program, which someone had to write in the first place (the "someone" in question being, of course, an employee of firm X). People who do not understand what a proram is--who do not realize that not only its data, but also its inferential or evaluative processes, are in principle open to challenge--may indeed be gulled into believing that "If computers choose something, then we should choose it too." If the choice merely concerns the purchase of one commodity rather than another, this is perhaps not too worrying. But if it concerns more socially or politically relevant problems, such mystification could be most unfortunate.

Sensationalism feeds on ignorance, and many descriptions of artificial intelligence in the media, and in "popular" books about the subject, are sensationalist in nature. Whether proclaiming the "wonders" or the "dangers" of AI, they are not only uninformative but highly misleading--and socially dangerous to boot. They suggest that things can be done, or will be done tomorrow, which in fact will be feasible only (if ever) after decades of research (including the "long-range research" mentioned above). And they underplay the extent of human responsibility for these systems, much as the X-advertisement does.

Unfortunately, these sensational reports are sometimes encouraged by ill-judged remarks from the AI community itself. A recent hour-long BBC-TV science program began and ended with a quote from a senior computer scientist at MIT, gleefully forecasting that the intelligent machines of the future would worry about all the really important problems for us (***for*** us, not ***with*** us). As he put it (with apparent satisfaction): if we ever managed to teach chimps to speak, we wouldn't talk to them for long--for they would want to talk only about bananas; super-intelligent machines will be similarly bored by people, for we won't be capable of understanding the thoughts of the machines. His conclusion was that the super-intelligent AI-systems will justifiably ignore us, leaving us simply to play among ourselves.

Humanity has of course been advised before to neglect the difficult moral and philosophical questions, to live life on the principle that "***Il faut cultiver son jardin***". But that was said in a rather more ironic spirit. Enthusiasts evaluating AI's contribution to society would do well to emulate the common sense, if not the scepticism, of Voltaire.

REFERENCES

[1] Ballard, D.H., and C.M. Brown, eds., 1982 *Computer Vision.* Englewood Cliffs, NJ: Prentice-Hall.

[2] Boden, M.A., 1977 *Artificial Intelligence and Natural Man.* Brighton: Harvester.

[3] Boden, M.A., "Artificial Intelligence and Biological Reductionism", in M. Ho and P. Saunders, eds., *Beyond Neo- Darwinism.* London: Academic, in press.

[4] Boden, M.A., "Educational Implications of Artificial Intelligence", in W. Maxwell, ed., *Thinking: The New Frontier.* Pittsburgh: Franklin Press, in press.

[5] Bond, A.H., ed., 1981 *Machine Intelligence* (Infotech State of the Art Report, Series 9, No. 3). Maidenhead: Pergamon.

[6] Brady, J.M., ed., 1981 *Artificial Intelligence* (Special Issue on Vision).

[7] Davis, R., and D.B. Lenat, eds., 1982 *Knowledge- Based Systems in Artificial Intelligence*. New York: McGraw-Hill.

[8] Erman, L.D., P.E. London, and S.F. Fickas, 1981 "The Design and an Example Use of Hearsay III", *Proc. Seventh Int. Joint Conf. on AI*, Vancouver BC.

[9] Feigenbaum, E.A., and A. Barr, 1981 *The Handbook of Artificial Intelligence*, 3 vols. London: Pitman.

[10] Feigenbaum, E.A., and P. McCorduck, 1983 *The Fifth Generation*. New York: Addison-Wesley.

[11] Hayes, P.J., 1979 "The Naive Physics Manifesto", in D. Michie, ed., *Expert Systems in the Micro- Electronic Age*. Edinburgh: Edinburgh Univ. Press, pp. 242-270.

[12] Hayes-Roth, F., D.A. Waterman, and D.B. Lenat, eds., 1983 *Building Expert Systems*. Reading, MA: Addison-Wesley.

[13] Hinton, G.E., 1981 "Shape Representation in Parallel Systems", *Proc. Seventh Int. Joint Conf. on AI*. Vancouver BC.

[14] Hinton, G.E., 1983 "Optimal Perceptual Inference", *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*. Washington, D.C., June.

[15] Hinton, G.E., and J.A. Anderson, eds., 1981 *Parallel Models of Associative Memory*. Hillsdale, NJ: Erlbaum.

[16] Jenkins, C., and B. Sherman, 1981 *The Leisure Shock*. London: Eyre Methuen.

[17] Lehnert, W.G., and M.H. Ringle, eds., 1982 *Strategies for Natural Language Processing*. Hillsdale, NJ: Erlbaum.

[18] Lenat, D.B., 1982, 1983, "The Nature of Heuristics (three papers)", *Artificial Intelligence*, 19 and 20.

[19] McCarthy, J., 1980 "Circumscription--A Form of Non-Monotonic Reasoning", *Artificial Intelligence*, 13.

[20] McDermott, D., and J. Doyle, 1980 "Non-Monotonic Logic", *Artificial Intelligence*, 13.

[21] Marr, D., 1982 *Vision*. San Francisco: Freeman.

[22] Nilsson, N.J., 1980 *Principles of Artificial Intelligence*, Palo Alto: Tioga.

[23] Papert, S., 1980 *Mindstorms: Children, Computers, and Powerful Ideas*. Brighton: Harvester.

[24] Paul, R.P., 1981 *Robot Manipulators: Mathematics, Programming and Control*. Cambridge, MA: MIT.

[25] Rich, E., 1983 *Artificial Intelligence*. New York: McGraw-Hill.

[26] Schank, R.C., and C.K. Riesbeck, eds., 1981 *Inside Computer Understanding: Five Programs Plus Miniatures*. Hillsdale, NJ: Erlbaum.

[27] Selfridge, O., M. Arbib, and E. Rissland, eds., *Adaptive Control in Ill-Defined Systems*, in press.

[28] Sleeman, D.H., and J.S. Brown, eds., 1982 *Intelligent Tutoring Systems*. London: Academic.

[29] Stefik, M., J. Aikins, R. Balzer, J. Benoit, L. Birnhaum, F. Hayes-Roth, and E.D. Sacerdoti, 1982 "The Organization of Expert Systems", *Artificial Intelligence*, 18, pp. 135-173.

[30] Waterman, D.A., and F. Hayes-Roth, eds., 1978 *Pattern- Directed Inference Systems*. New York: Academic.

[31] Winograd, T., 1983 *Language as a Cognitive Process: Syntax.* Reading, MA: Addison-Wesley, (chapter on computer models).

[32] Winston, P.H., and R.H. Brown, eds., 1979 *Artificial Intelligence: An MIT Perspective*, 2 vols., Cambridge, MA: MIT Press.

[33] *Proceedings of the International Joint Conference on Artificial Intelligence* (biennial, odd-numbered years).

[34] *Artificial Intelligence* (quarterly journal).

[35] *Proceedings of American Association for Artificial Intelligence* (annual).

[36] *Proceedings of the European Society for the Study of Artificial Intelligence and Simulation of Behaviour: AISB* (biennial, even-numbered years).

[37] *International Journal of Robotics Research* (quarterly journal).

THE INTERSECTION OF AI AND EDUCATION:

Can we bridge the gap and overcome the crisis by investing the limited resources in thought-intensive endeavors?

Stefano A. Cerri*
(Italy)

INTRODUCTION

In this paper we intend to submit to the attention of the workshop's participants the issue concerning what could be done by the people, groups, and countries who do not have access to large human and financial resources to contribute to and benefit from high level scientific and technological developments.

We claim that the field of Artificial Intelligence can prove to be promising in giving unpredictably rewarding results provided some (relatively limited) resources are invested into R&D projects with the following characteristics: medium-long term time span, thought and expertise-intensive goals.

*Faculty of Information Sciences, University of Pisa, Corgo Italia, Pisa, Italy.

The optimism of this statement accompanies the fundamental pessimism arising from other considerations--for example, the conviction that only the US and Japan are at present able to envisage an expansion of their production because they seem to have access to the high technologies needed to cope with the keen competition in the modern international market. This implies that even countries in Western Europe (and obviously all those less industrialized) will play an ever more subordinate role in the next years.

Such a picture should not appeal even to the American and Japanese authorities and we are ready to present many reasons to show why such a situation is not convenient to anyone.

So, let us start from the hypothesis that a more even distribution of the technological--hence also economic and social--developments in the world is beneficial for everyone. We will call in the following: 'even distribution' and the associated 'fair growth' (without further definitions) "our purpose".

One of the most important developments is in Information Technologies. Artifical Intelligence, and its related methods, techniques, and products are among the most promising branches of these technologies (cfr: the ESPRIT project in the EEC, the ALVEY project in the UK, the Japanese 5th Generation project and the US recent investments in the fields).

Provided one agrees to the cited strategic importance of AI, the question arises if it is conceivable to start any R&D endeavor in a technological field, with limited financial resources and expertise, and hope to compete in a reasonable time with countries, groups and people who have access to bigger fundings and a well established scientific school.

We do not have the ambition of giving a definite answer to this question. We wish, instead, to raise the problem and approach a tentative solution by stimulating a discussion on issues related to what we believe to be a good testbed, nl. the three intersections of AI and Education: AI education, the automatic acquisition of knowledge and Intelligent Tutoring Systems.

The line of reasoning will be as follows:

Economic Impact

> Assuming that the development of AI R&D projects is strategically important for the development of Information Technologies, and therefore has a great economic impact, are there aspects of AI R&D which make it peculiar within other R&D projects, so that--tentatively--the conclusion can be drawn that AI R&D is potentially suited to "our purpose"?

Social Impact

> Assuming that the growth of Information Technologies implies a growth of their use, that both are somehow unavoidable and have certainly a nontrivial social impact, are there aspects of AI R&D which make it peculiar so that--tentatively--the conclusion can be drawn that this social impact can be positive?

INFORMATION TECHNOLOGIES: COMPUTER SCIENCE vs ARTIFICIAL INTELLIGENCE

There are many issues which can be considered to discriminate the traditional Computer Science and the AI approach to Informatics, see for example, Newell [1].

In this paper we cannot discuss explicitly Newell's considerations, which we consider in any case to be illuminating in understanding both the history and the nature of AI as compared to traditional Computer Science.

Shortly, whereas traditional CS is mainly concerned with numeric and well defined problems to be solved algorithmically by constructing efficient computing systems, AI is interested in symbolic, ill-defined problems which are hard to be formalized, but can be solved by designing adequate representations of the knowledge available, possibly in a hierarchy of representation levels; and nondeterministic search strategies and reasoning systems.

The fundamental question to be put to such AI systems is not whether they are consistent or complete or efficient, but whether they represent an adequate solution to the original problem.

In a later stage, the verification of the consistency and completeness of the system is needed to ensure the correctness of the solutions. Finally, efficiency considerations allow one to consider the system as (economically) viable.

The shift from the efficiency to the adequateness of the solution as a main goal has some immediate consequences:

(a) it reverses the traditional approach to Information Technology: one is required to design "special purpose" software and hardware starting from the concrete needs and ending with a system;

(b) it requires considerable efforts and expertise that should come from at least two sources: the expert in the application domain and the expert in the formalization of the knowledge needed to solve a problem in that domain;

(c) it can have a considerable impact not simply on the applications in whatsoever domain, but also in the production of computer technology, provided the solution found can be shown to be valid in other domains.

Let me skip all discussions about the reason why I believe that knowledge framework and the knowledge-engineering approach are good candidates for solving many of these problems which are and will be relevant in the next few years, because I believe these issues will be elaborated by many other contributors.

Instead, I shall concentrate on the need for high-level competence of a NEW TYPE in the construction of knowledge based systems.

NEEDS FOR SPECIFIC KNOW-HOW IN AI

Whereas one can expect a traditional computer scientist to be capable of designing an efficient algorithm for the solution of a well-formalized problem, it is hard to imagine that the same skills will be sufficient to solve problems for which there is no formalization available, because the solution is only to be induced from the behavior (possibly unconscious) of an expert in the object domain. So the traditional computer scientist does not have IN PRINCIPLE an adequate education (mastering of concepts and skills and ability to communicate) for the AI tasks.

Thus, EDUCATION is central to any attempt in this challenge. We wish to emphasize this point because our experience in European countries has been negative and we believe that a different attitude, both from public and private institutions and companies, could modify sensibly the situation in a relatively short time.

There are at least five main reasons why AI education has not deserved much attention in Europe:

(1) AI was not a field "formal enough" to be a candidate of academic interest;*

*As far as we know, Cambridge and Oxford, for instance, do not have a chair even of Computer Science!

(2) for all relatively new fields, the latency in Europe is higher than that in the US. In the best cases, AI education is offered as a 'side' course in CS in the last years of the undergraduate curriculum, while the conceptual content can and should be assimilated at a much earlier stage, even possibly in high schools;*

(3) there was little or no interest from the consumers, i.e., the Industries;**

(4) even when there has been interest from the industries in some fields of science or technology, normally the reaction of the Academic sector has been very slow: the Industrial-Academic cooperation has always been poor;

(5) being AI interdisciplinary (CS, Psychology, Linguistics), no traditional department has been able to catalyze AI (educational) activities, but each specific department has been strong enough to inhibit the emergence of the new, competitive discipline.

Let us now consider AI education as a prerequisite for successful AI R&D projects. We have claimed that AI requires a new type of know-how. We are going to suggest now that Europe on the one hand, and the US and Japan on the other, possibly for different reasons, should both sponsor activities on AI Education and AI fundamental research, i.e., projects for building up fundamental, specific know-how in Artificial Intelligence.

Concerning the Europeans, a good reason can be represented by the motto that it is not easy to build up know-how, but it is cheaper than importing it.

---

*An exception—conforming the rule—is given by the Department of Artificial Intelligence in Edinburgh.

**I read the first advertisement in an Italian newspaper caling for AI experts only a few months ago.

In fact, we believe that the technological gap between the US/Japan and, say, the (West and East) European countries, cannot be bridged by studying at present prototypes embodying technologies which are or will soon be available in products on the commercial market. Europeans should enter a new context, the one where they have the maximum chance to invest fruitfully, i.e., thought intensive R&D. Obviously, building ON the state of the art technology.

Furthermore, we think that more investments on Education and fundamental research in AI can be fruitful also in the US and Japan, because the acceleration of the process of know-how formation in a time of crisis can be vital, in a medium term, for a economic expansion.

Actually, new, economically promising fields can be opened while solving important issues such as the growing need for selected information exchange as an adequate base for decision making (cfr. in offices, or the explosion of scientific information frm technical reports); the access to natural resources (cfr. oil); the need for a growth of agricultural products; the need for a democratic control of the weapon control, etc.; which are all fields potentially taking advantages from basic AI advancements.

In the US and Japan, research and development in AI must address relevant issues in a systematic way; these are thought-intensive endeavors and not simple exploitation of existing tools (cfr. Schank, 1983 [2]). It is regrettable that most if not all the American PhDs in AI tend to move from the research to the industrial world.

All these considerations rely on economic evaluations which are far beyond the scope of this paper. However, we can estimate some costs for building and maintaining an AI R&D laboratory. Roughly, we need about two-four years education for the average trainee, we pay him/her at a rate which could be about 30-50% higher than the rate of another scientist/technician--

because the field is not yet stabilized—and we spend an average of $50 K per man in special hardware-software resources. We believe these to be costs affordable for everyone interested in high technology and sensitive to the arguments we have presented here.

IS AI PECULIAR?

Artificial Intelligence is in its infancy like a framework for future developments in high-level technology. So it is a matter of feeling to take one or another position concerning what AI really is and what its impact will be.

An easy position would be taken by stating that, since AI is new and promising, everyone should hurry so as not to miss the expected rewarding results.

However, our task in this arena is to make provocative statements in order to stimulate reactions.

We claim that Artificial Intelligence is peculiar within other disciplines, and that its characteristics justify an optimistic view about the expected economic and social impact of AI with respect to what we previously called "our purpose".

Many scientists complained that results in AI have not yet been CUMULATIVE, i.e., a completed "intelligent" program or achievement can hardly be utilized for a new project.

This aspect of Artificial Intelligence is not temporary: we believe that it is inherent. In fact, the formalization of a piece of knowledge or reasoning is a (relatively) new enterprise every time. If it were possible to build a completely additive theory of Artificial Intelligence, we would have built a theory of thought, which is very improbable in the coming centuries.

So, each new problem requires one to think hard, to attempt formalizing the knowledge at hand, and structuring his/her formalization process. Notions such as heuristics, imprecise, approximate, or contradictory knowledge, and the like do not expect a general, logic-deductive theory which embodies every solution to every problem.

The fact that languages such as mathematics or logic are helpful in explaining how the knowledge manipulation done by AI programs actually works, does not contradict with the fact that the selection of knowledge and knowledge processing mechanisms is NOT a purely logical process. Just like in Mathematics, the intuition of a theorem is not as logical as the proof is. About this controversy, I adopt Newell's [3] point of view on Nilsson's [4] statements about logic in AI.

The inductive, experimental character of the construction of Expert Systems--the embodiment of AI methods and tools--by formalizing the expertise of humans and transforming it into running systems, makes AI difficult for a novice but accessible to everyone who has been trained in its methods and tools.

Though AI seems conceptually difficult for a novice we do not expect--surprisingly--to have to spend much effort in this training, because the structure of reasoning in AI is much more similar to the human everyday thought than it is the case, for instance, of traditional mathematics or, even of contra-evident disciplines such as physics or chemistry. Moreover, "intelligence is not all or nothing"--so we can grade our ambitions.

We gave two one-year (90 hours) courses to about 100 CS students, and at the end of the courses, we could consider having formed the participants to be potential AI workers; part of them have been so enthusiastic and worked so hard on their own that we do not hesitate to hope that they can produce high level research proucts in a short time and become tutors of other novices.

We therefore contend that though results in AI are not cumulative, AI education and training are very cumulative. If one wishes some general reason for this phenomenon, we can suggest that this might be a consequence of the fact that AI education cannot be purely formal, instead it can rely much on common sense intution and concrete engineering practices.

Nilsson [4] states that "AI research should be more concerned with the general form and properties of representational languages and methods that it is with the content being described by these languages", but also that "notable exceptions involve 'common-sense' knowledge about the everyday world and metaknowledge".

The argument seems to us somehow inconsistent. If one accepts Nilsson's two "exceptional" fields of AI research as appropriate, one includes (from the second) also what Nilsson states to be the main concern of AI research: the content of research on metaknowledge seems to us to coincide with the general form and properties of representational languages and methods. The first "exceptional" field--the study of common-sense knowledge and reasoning--can be seen as the whole AI concern when applied to some external problem; while AI applied to knowledge and reasoning is AI applied to AI, i.e., the study of metaknowledge.

We conclude that AI is a basically experimental discipline; its theoretical framework is built by applying the experimental methods and tools to AI itself.

This experimental, applied, practical aspect of AI might also be beneficial for a better cooperation between Industries and the Academic sector. The rigorous formal, logical deductivism of traditional science and engineering education (particularly in Europe) has damaged the creative, inductive processes of construction of experimental sciences and technological artifacts.

There is another aspect which is rather peculiar to AI and can be fruitful for "our purpose".

The cooperation of the two knowledge sources for the construction of Expert Systems requires not simply knowledge in the object domain and in AI, but the creative skill of combining the two in a unique system which is only satisfactory if it is adequate to the purpose. This skill is missing in many industrial and academic contexts. Once gained, it can be used in many applications.

It requires the only resources which might be available even in low-budget projects, i.e., intelligence and hard intellectual work.

We have presented above some convictions about the need for AI Education.

Many qualified people have given significant arguments to make the study of learning and automatic knowledge acquisition in Expert Systems a central RESEARCH issue for all further developments in AI (cfr. Schank [2]).

We think that the major components of Intelligent Tutoring Systems could represent key endeavors in AI DEVELOPMENT, i.e., making complex AI systems based on the state of the art know-how. We explicitly refer here to such issues such as the analysis of correct and incorrect problem-solving behavior from human protocols; natural knowledge (linguistic, graphic, and possibly pictorial) understanding, representation and production, cooperative mixed-initiative behavior in a "mutual (incomplete) knowledge" framework, integrated (hardware-software) knowledge compilation, etc. When building an "intelligent" tutor one can begin with a traditional CAL*-like system and add progressively "intelligent" modules to make the interaction more "friendly" and effective. We believe that ITS are good testbeds for AI development projects, because they

*Computer Assisted Learning.

can address a wide variety of object fields--including AI itself--and they can easily embody a spectrum of levels of "intelligence", without loosing necessarily their effectiveness. We believe that the development of ITS in the seventies has been one of the major sources of ideas, methods and tools for AI in general.

Finally, we integrate the three fields: AI Education, Learning and ITS, with the remark that the systematic organization of prescriptions for Expert System design needed to render the results cumulative, is based on a creative skill which is normally attributed to a good teacher; this can be embodied in an Intelligent System teaching AI notions, and this system could possibly be able, in the future, to learn from its own experience.

CONCLUSIONS

We have tried to motivate the optimistic view of AI as an activity with a positive economic and social impact even for low-budget countries or groups by examining if AI can be considered peculiar within other disciplines, in relation especially to its ECONOMIC impacts. In a discussion about the three intersections of AI and Education: AI Education, research on Learning and the development of Intelligent Tutoring Systems, we have claimed and tentatively motivated some characteristics of AI.

These were the "lack" of CUMULATIVENESS of the results; the need of COOPERATION between experts in the problem domain and knowledge engineers; the EXPERIMENTAL nature of AI and the COSTS of an AI project.

The discussion on the SOCIAL impact of AI seems to have been of little interest to us.

Instead, we will now try to show that the arguments presented to describe the peculiarities of AI as a vehicle for economic development--if accepted--can be convincing starting points for a discussion on the social impact of AI.

First

As we said in the introduction, the social importance of AI can be considered to be a direct consequence of the strategic importance of AI in the development of advanced technologies.

However, this does not mean that whatever technological growth is socially acceptable.

Second

One of the main concerns of AI is the construction of "user friendly" systems. Actually, the shift from the 'efficiency' to the 'adequacy' of the problem solution was always parallel, in the history of AI, to the shift of concerns from the 'computer' to the 'man' in interactive systems. As now one can accept that the computer will become an everyday tool for everyone's activities, we can deduce easily that AI and its methods will be important in the "Information Society" of the future.

While this seems to be the major aspect of the social impact of AI, we believe that the next two considerations deserve even more attention than the previous ones. We refer to the above discussion on the nature of AI.

Third

As AI is concerned with the formalization of common-sense reasoning in fields not yet formalized, the development of AI can play a central role for the scientific organization of knowledge (or theory formation) in traditional "humanistic" disciplines (e.g., linguistics and psychology) but also in some aspects of disciplines--such as law or medicine--where problem solving and decision taking was left to the "clinic eye" of the expert. Down to the apparently "lower" aspects of the oganization of the work in the office, the access to information allowed by the development of "user friendly" systems and the power of AI systems for problem solving and decision making can have a radical impact on the society not only because these systems propose solutions previously unexplored, but also because the solutions proposed are a formalization of the problems themselves.

However these applications of AI to real problems should not be considered a direct expansion of the application of the traditional information technology. The relevant difference, we believe, is reflected by the following consideration.

Fourth

Because AI is also concerned with AI itself as an application field (the meta-level) the formalization, generalization and abstraction power of AI reflected in AI systems allows one to think to a technological framework where, possibly for the first time, the cumulation of results, theoretically hard to be justified, is obtained simply as a consequence of practice. In this view, work on meta-knowldge and meta-reasoning have a social impact on the parameters defining the social impact of AI systems applied to the solution of "external" problems (e.g., the speed of introduction of new technologies in different yet "analogous" domains).

One can argue that this characteristic was present also in traditional CS. The difference is that the paradigm of AI seems to have attacked in a unique framework--knowledge--all the levels of the representation and processing of information which were previously left to a large set of specific formalisms and languages.

We do not know if the AI metaphor can also be considered good for modeling human behavior. Certainly, it is the most advanced paradigm now available for describing it.

So, we conclude by presenting our last provocative belief that networks of "intelligent" computers will be able, in the future, to simulate (autonomously?) that part of human SOCIAL behavior--through the centuries--usually called intellectual development.

REFERENCES

[1] Newell, A., 1982 'Intellectual Issues in the History of Artificial Intelligence', Dept. Comp. Science Report, CMU, Pittsburg. To appear in: Machlup & Mansfield (eds.), The Study of Information: Interdisciplinary Message, New York: Wiley, 1983.

[2] Schank, R.C., 1983 'The Current State of AI: One Man's Opinion', *AI Magazine*, IV, 1.

[3] Newell, A., 1981 'The Knowledge Level', *AI Magazine*, Summer.

[4] See, for instance, Nilsson, N.J., 1981 'Artificial Intelligence: Engineering, Science or Slogan?', *AI Magazine*, Winter.

ARTIFICIAL INTELLIGENCE: A LESSON IN HUMAN SELF-UNDERSTANDING

Ivan M. Havel*
(Czechoslovakia)

"Substance of history consists in the experiences in which man gains the understanding of his humanity and together with it the understanding of its limits." (Voegelin, New Science of Politics.)

I

The middle of the twentieth century, with its specific technological, cultural, and intellectual climate, gave birth to a new scientific discipline: artificial intelligence (henceforth, AI). The newborn grows, relatively fast comparing to other scientific disciplines, and approaches already its adolescence. We become to be seriously concerned about its future, and ask ourselves whether we should take it in hand and how.

*Engelsovo nb. 78, Prague, Czechoslovakia.

AI has its characteristic inner development [1] but it has also its outer life: those actual or potential (and apparent or hidden) impacts on society, technology, our life-style, and last but not least our way of thinking.

For the purposes of this study I found it useful to discriminate between three kinds of impacts that AI, like many other human activities, may cause.

Impacts of the first kind are the ***intended*** outcomes of AI, i.e., those which actually motivated concrete research projects (I have on mind particularly the engineering trends in AI). They are or will be realized in various products like automatic consultants, advisors, tutors, computer psychiatrists, marriage counsellors, sophisticated knowledge bases, picture processors, identification and recognition systems, intelligent industrial and domestic robots, etc. Impacts of this kind were the subject of some earlier assessments [2] and they are also a favorite subject of popular newspaper articles and pocket books.

As *impacts of the second kind* we classify those which are *unintended.* They may be foreseeable by people with stronger imagination but they may also come quite unexpected. They are more likely to be valued negatively, as something beyond our control, but this is not a rule (similarly it is not a rule that impacts of the first kind are valued positively).

The more obvious cases of impacts of the second kind are possible side effects and/or misuses of the intended products of AI (example: speech-understanding systems used for orwellian purposes like surveillance of private conversations). There are, however, some deeper and less apparent impacts which also were not among the preconceived aims of AI. They take plce in the 'ideosphere' (Hofstadter's term) and involve shifts in viewpoints. Thus, e.g., one of the main maxims of AI, that "Machines may be smart", tends to wake up the old myth of l'homme machine–with all its negative consequences for human self-respect.

> "If the public believes--rightly or wrongly--that science regards people as 'nothing but clockwork', then clockwork-people we may tend to become." (Margaret Boden [3].)

Impacts of both kinds have been widely disputed within and outside the AI community. Joseph Weizenbaum, for instance, is the main representant of the insiders who have warned against irreversible dehumanizing or unethical effects of what he calls 'instrumental reason' [4].

Here I want to point to still another category of impacts, characteristic for AI, which I consider somewhat more fundamental than the previous ones. I call them *impacts of the third kind.*

II

Impacts of the third kind manifest themselves through *reflection.* That is, through full conscious awareness of the new situation in which man appears due to the existence of AI and due to lessons it offers to him.

The main intellectual contribution of AI is that it elaborates a computer metaphor for the mind. Thus, as a challenging theme, the mind and together with it the mysterious self enters into the focus of scientific interest. This gives reasons for hopes that AI will overcome those much-talked-about dehumanizing effects we may be afraid of.

Let us illustrate this point by a passage from Boden's book:

> "Contrary to common opinion... the prime metaphysical significance of artificial intelligence is that it can *counteract* the subtly dehumanizing influence of natural science, of which so many cultural critics have complained. It does this by showing, in a scientifically acceptable manner, how it is possible for psychological beings to be grounded in a material world and yet be properly distinguished from 'mere matter'. Far from showing that human beings are 'nothing but machines', it confirms our insistence that we are essentially subjective creates living through our own mental constructions of reality (among which science itself is one). In addition, for those of us who are interested, it offers an illuminating theoretical metaphor for the

> mind that allows psychological questions to be posed with greater clarity than before. The more widely these points are recognized, both within and outside the profession, the less of a threat will artificial intelligence present to humane conceptions of self and society." ([3], p.473.)

In virtue of the orientation to human self and due to inherent value of any reflecting activity impacts of the third kind are fundamentally positive. Moreover, abstract as they seem, they are more significant that "lower" kinds of impacts because they directly influence the social and intellectual climate--the very climate that gave birth to AI.

III

Let us expound some points touched above, as well as some further aspects of impacts of the third kind.

1. AI elaborates a *computer metaphor for the mind.*

There is a subtle but important difference between the metaphor and the model. Viewing the computer as a *model* of something involves the reductionist standpoint: the model gives the criteria of explanation, it forces us into its specific language. The modeled entity is virtually identified with the model; any difference is disregarded.

The metaphor, on the other hand, always reminds us of the difference: in fact, it is the tension between the identity and the difference which gives the metaphor its creative power. Metaphors are catalytic tools which, on principle of analogy, lead to new ideas, to a new understanding. A metaphor initiates exploration, whereas a model closes it. What is characteristic of metaphors is that they are not despotic, they never exclude other metaphors, even mutually incompatible ones.

We should bear in mind this distinction between modeling and metaphorical transfer when talking about computers, in particular. Take, for instance, the following Weizenbaum's statement:

> "...the computer is a powerful new metaphor for helping us to understand many aspects of the world, but... it enslaves the mind that has no other metaphors and few other resources to call on." ([4], p.277.)

Obviously, the second part of the sentence talks about the computer as a model rather than as a metaphor.

2. Through AI *the self enters into the focus of scientific interest.*

The revival of interest in self, in human subjectivity, can be documented already by the highly suggestive titles of books like "The Mind's I" [5] or "The Brain and Its Self" [6].

AI is primarily a constructive discipline. It tries to construct something that would, at least by its performance, remind, replace, or transcend human thinking, or in the case of robotics, remind, replace, or transcend humans as such. Whoever plans to launch such a project is inevitably confronted with questions like "What is thinking?" or "What is man?". As a matter of fact, himself a man, he may ask these questions in another, self-referential way: "What is my thinking?", "Who am I?". Now there is an important difference between these two ways of putting the questions: the former is external, using the more common *outer view*, while the latter is internal, relying on one's *inner view*. This is how *the self* enters into the picture.*

Is subjectivity realizable on the principles of AI? As I see it, the answer should be no. While constructing any artifact requires a prior *external* descrip-

*An exciting intellectual path from AI through the phenomenon of self-reference to the riddle of the self is one of the strands in Hofstadter's braid [7].

tion, our self is experienced only by the mind's *inner* eye. Our subjective self is more a view than a thing. Incidentally, this difference points to certain limitations in writing computer programs that would realize genuine intellectual activities. Even a particular and rule-obeying activity, like playing chess, if considered in its depth, involves the subjective intentional self as an unseparable constituent.

Successful in its aims or not, AI research awakeness our desire for greater self-understanding. This very desire is, I believe, the most positive impact of AI.

The above-cited passage from Boden's book illustrates the point. Its essence is that AI can counteract dehumanization and that this counteracting ability may be based on showing the difference between psychological beings and 'mere matter'. On the other hand, the claim that AI actually shows *how* the psychological being transcends matter is somewhat overstated. This "how" most likely refers to Boden's claim that

> "the crucial notion in understanding how subjectivity can be grounded in objective causal mechanism is the concept of an internal model or representation" ([3], p.428.)

Maintaining internal model of the world, perhaps even of hypthothetical worlds and of the system's own capabilities, is one thing and possessing inner *perspective* is quite another thing. What has AI shown up to now is just that computers, in a sense, transcend clockworks.*

*It is worth mentioning that AI is not the only, nor the first scientifically based discipline which directs attention to the subjective self. One example is the new physics: "If the new physics has led us anywhere, it is back to ourselves, which, of course, is the only place we could go" ([8], p.114). Another example is the biological-cybernetic approach of Ruyer [9].

3. Computer metaphor of the mind gives a *new flavor to the old issue of free will.*

Self-determined action [3], authentic choice [4], or intentionality [10] appears to be something else than just an opposite to determinism. Surely, computers can be programmed to make deterministic decisions as well as random steps. But free choice is something quite different than randomized decision making. It is a creative act guided by reason grounded in authentic individual experience. Yet this act is essentially unpredictable. The world picture based just on chance and necessity [11] is no longer sufficient for it leaves no room for intentionality.

4. There is a particularly interesting *lesson to be learned from computational linguistics.*

Attempts to teach a machine to understand and talk in natural language has revealed that talking is something more than just exchanging coded messages. Language acts are deeds that affect things and events in the world in an analogous way as physical actions. Through language man enters into a complicated arena where the ideosphere pervades the biosphere.

A common opinion in AI is that a large knowledge base might solve the problem of natural language understanding by computer. This perhaps holds for an ad hoc restricted language with rigid semantics. However, talking in human language is hardly feasible without actual human experience in the real world.

5. AI leads to *new methodological approaches.*

Current reseach in AI is an unusual combination of constructive, empirical, and theoretical methods. A new powerful source of ideas in AI is also introspection (e.g., the intuitive basis for Minsky's frame systems).

AI gave a concrete form to heuristic reasoning and reasoning by analogy; the creative role of metaphors was mentioned above. Related to metaphors is K. Pribram's concept of "abduction" [12]: it is a creative transfer of a whole complex of concepts from one scientific field to another (example: the holographic brain. Finally, we should not forget paradoxes as a rather unconventional tool which helps to get a deeper insight where other tools fail.

6. The influence of AI through *literary fiction as well as non- fiction* belongs to impacts of the third kind, too.

Besides widely distributed science fiction and standard popularizing literature (giving often a rather simplified and/or exaggerated picture of AI) I would like particularly point to a new kind of literature best exemplified by the well-known Hofstadter's book Gödel, Escher, Bach or by the anthology The Mind's I [7],[5]. It is a highly influential and stimulating literature which, besides its entertainment value (or perhaps by virtue of it), has profound philosophical and scientific significance. Maybe, impacts of AI induced by this type of literature will be more recognizable than impacts associated with concrete applications of AI.

7. Impacts of third kind *manifest themselves through reflection.*

Every human activity can be subjected to reflection. Reflected may be both the activity itself (its motivations, purposes, means, etc.) and the acting subject. What is specific of AI is that these both reflections are, in a sense, already its constituents: man is, as we have seen, its prototype and among the goals of AI is, after all, to test its own limitations.

Any proper reflection has to reflect also the reflecting subject as well as the act of reflection itself--a self-referential loop that leaves little room for fallacies or ill effects. An argument for not being so afraid of the consequences of AI.

Let us turn now to the specific questions recommended for the IIASA Task Force Meeting. (Incidentally, the very fact that such a meeting has been summoned us an example of an AI impact of the third kind.)

IV

1. DEVELOPMENT DURING THE NEXT TEN YEARS

1.1 What are the most likely core research areas in AI and what results do you expect?

I expect that large knowledge systems (the present core research area in AI) will *not* grow ever larger. The known "pig's principle" (when something is good, more is better) certainly does not hold here. Perhaps more attention will be given to methods of converting data into knowledge. *Semantic search* (as opposed to combinatorial search) or *knowledge retrieval* may be the names of new areas of interest. The 'knowledge space' may be endowed with a topography, in which events like 'changing one's viewpoint may have a programmable analogy.

New conceptions of knowledge organization in the computer may be expected. For instance, using the visual field as a metaphor suggests that the system should have at every instant a 'concentration point' which, however, cannot be deprived of its surrounding environment, i.e., if the context which gives it the proper meaning. A shift of topic is then represented by a transfer of the concentration point.

1.2 What impacts of AI on other sciences and on technology do you expect?

In view of what has been said about the third kind of impacts of AI we can expect in the near future a substantial effect on psychology and the related fields (in fact, the effect is visible already now). I do not mean dependence of psychology on computer modelling (if so, then rather on the computer metaphor) but a more fundamental change: psychology will no longer be a science about some 'objects' (called 'subjects'just habitually) but a well-founded effort for deeper self-understanding of man. Neuropsychology (a lower level approach) and behavioral pscychology (a higher level 'outer' approach) need a third companion (a higher level 'inner' approach) which would use introspection as its main source of knowledge.

Through its impacts of the third kind AI may influence also other social sciences and humanities, perhaps even more than natural and technical sciences.

A few words about mathematics. It seems that the current growth of interest of mathematicians and logicians in computer science will continue and that AI will provide them with a number of challenging topics for fundamental research. For instance, certain intuitively defined qualities relevant to AI can be set up in opposition to analogous qualities which already have mathematical explication: heuristic v. algorithmic, vague v. accurate, small (large) v. finite (infinite), analogous v. homomorphic, insightful v. deductive, self-referential v. recurisvie, paradoxical v. contradictory, etc.

1.3 What social impacts do you expect?

In the near future impacts of the first and second kinds will be more likely associated with computer technology than specifically with AI.

As for the impacts of the third kind, we should be able to feel them soon as slow changes in the intellectual climate. However the way they may affect our daily life is difficult to assess.

1.4 What economic impacts do you expect?

Regarding the use of computers in management I think that what we can learn from AI about the difference between programmable decision making and free authentic choice can spare us quite a few desillusions.

2. IF YOU WOULD SPECULATE ABOUT THE LONG RANGE AI RESEARCH RESULTS AND IMPACTS, WHAT COULD THEY BE?

I already expressed my scepticism concerning the possibility of fabricating autonomous mind. A more promising long range trend in AI seems to be using the computer as an on-line extension of the human brain (the occasionally used term 'symbiotic AI' is improper because symbiosis involves *two* organisms in mutual interdependence). After all, in order to help somebody it is not necessary to replace him.

To speculate about long range social impacts—and choosing the optimistic end of the scale—there is a chance that impacts of the third kind, particularly the revival of interest in human self, may help to restore the respect for human individuality, privacy, and freedom of thought.

3. WHICH RESEARCH AREAS OF AI WOULD YOU AS PART OF THE AI COMMUNITY SUPPORT MOSTLY, AND WHY?

Impacts of the third kind seem to be most valuable and least dangerous. Therefore I would favor those areas of AI which promise to yield new ways of human self-understanding. It seems that theoretical investigations are in this respect more promising, in general, than ad hoc programming.

V

One final remark on precautionary measures.

Discussions concerning scientists' ethics often circle around the issue whether certain research areas with potentially harmful applications should be abandoned and by whom (by individuals, scientific community, or by governments). I doubt the problem is properly posed. Much earlier than such an issue can occur in a concrete form, the social climate have had to be already pathological, namely by producing the very *idea* of the harmful activity in question. Neither the inventor nor the manufacturer of a microphone is fully responsible for its possible misuse as a bugging device: the idea of breaking one's privacy appeared earlier (incidentally, there had been times when such an idea would have been absurd!).

Eliminating symptoms is not the most efficient therapy. In case of possible harms caused by science and technology it is the climate itself that should be diagnosed and cured first.

The growing concern of the AI community and of general public about potential influence of AI (to be clear: already the existence of the concern, not some particular measures it may yield) is an indication that the contemporary climate has still some abilities to look after itself.

REFERENCES

[1] Newell, A., 1983 "Intellectual Issues in the History of Artificial Intelligence" to appear in: F. Machlup and U. Mansfield (eds), *The Study of Information: Interdisciplinary Messages*. New York: John Wiley and Sons.

[2] Firschein, O. et al., 1973 "Forecasting and Assessing the Impact of Artificial Intelligence on Society", in IJCAI-3.

[3] Boden, M., 1977 "Artificial Intelligence and Natural Man". New York: Basic Books.

[4] Weizenbaum, J., 1976 "Computer Power and Human Reason". San Francisco: W.H. Freeman.

[5] Hofstadter, D.R., and D.C. Dennet (eds)., 1981 "The Mind's I. Fantasies and Reflections on Self and Soul". New York: Basic Books.

[6] Popper, K., and J. Eccles., 1978 "The Self and Its Brain". Springer Verlag.

[7] Hofstadter, D.R., Gödel, Escher, Bach., 1979 "An Eternal Golden Braid". New York: Basic Books.

[8] Zukav, G., 1979 "The Dancing Wu Li Masters: An Overview of the New Physics". New York: Bantam Books.

[9] Ruyer, R., 1965 "Parodoxes de la conscience et les limites de l'automatisme". Paris: Albin Michel.

[10] Dennet, D.C., 1978 "Intentional Systems". In: Brainstorms. Montgomery: Bradford Books, 1978.

[11] Monod, J., 1971 "Chance and Necessity". New York: Random House.

[12] Pribram, K.H., 1981 "The Brain, the Telephone, the Thermostat, The Computer, and the Hologram. *Cognition and Brain Theory*. Vol.4, No.2, pp.105-122.

OUTLINE—REMARKS REGARDING SOME ASPECTS OF AI-RESEARCH

Friedhart Klix*
(GDR)

Artificial Intelligence presently covers a wide range of research activities. Because it has been multidisciplinary since its origin there is a wide variety in the opinions regarding this field, reflecting the different personal backgrounds with which the judgments of scientists are linked. This text is written from a psychologist's point of view.

My own research activity is twofold:

(1) I am engaged in memory-research, mainly, the way concepts are represented in human memory and how are they interrelated. The basic idea is that there are two different types of knowledge: event-related knowledge, available by priming or search processes, and property-related knowledge available by algorithm-like automatized

*President, International Union of Psychological Associations, Department of Psychology, Humboldt University, Berlin, German Democratic Republic.

inferences, mainly established by comparison and decision processes. Many experiments have been designed and performed, and most data suggest that the discrimination between types of knowledge is significant. Computer programs now being derived serve as hypothetical models of the conceptual recognition processes, with programs that can be used as tools in order to derive new hypotheses, to be tested by further experiments.

(2) The second approach is strongly related to the first one. The question is whether there are individual differences in the availability of inferred knowledge, what the differences are, and how they may be simulated by computer models. It has been demonstrated that the individual differences are strictly related to what we intuitively call human intelligence. Mathematically gifted adults have been shown to differ in the way they perform complex recognition tasks (relation-detection among complex patterns and analogy-detections between them), and how their performance differs from that of highly intelligent subjects in a more generalized sense. Computer programs which synthesize the different strategies of different groups of subjects are able to reconstruct the probable mode of solution-generation. They are in a specific sense artificial, and they claim also in a particular sense the denomination "intelligent".

Before this background of personal experience I would like to stress the following AI-research streams in the near future:

(1) There is some delay in language understanding by computers (related to the most promising beginning which started about 15 years ago). I am convinced that research on human memory activity may give some impacts so that a new start on language (=text)--understanding

by computers may begin after 1985. This means that language understanding remains a core area in AI.

(2) There is some evidence that a convergence of the variety of programming languages to two basically different types may happen: one type, most suitable for language-processing, expert-systems, file-organization and manipulation, and related service capabilities; and a second type, most suitable for higher-order numeric and especially non-numeric operations or transformations as they are related to the transformation of algebraic and/or equivalent geometrical expressions.

(3) The influences on other sciences will be important where there are problems of a higher complexity. Man-computer interaction in the sense of a hybrid-intelligence will be able to tackle high degrees of complexity and make those degrees transparent which overtax the capability of a single human nervous system. One example may be the weather-forecasting, another one: organizational management of large scale systems.

(4) Social impacts have to be expected with regard to so-called man computer-interaction. A specific elaboration which is particularly related to this topic is appended.

(5) I have no specific idea on economic impacts.

(6) In the long term AI research will give rather large impacts on the capabilities of robots. I think that more and more intelligent recognition and decision devices will be used in order to enlarge the flexibility of the (expensive) robots. In addition some groups of "instinct-machines" with hardware wired recognition and decision device will "survive" for specific groups of tasks.

(7) In my case I welcome AI research work that makes clear the different types and modes of "intelligent" programs, the purposes for which they may be developed, and how they may be beneficially used in different applications.

Circular II

Second Information on an International Network on Psychological Man-Computer-Interaction-Research (MACINTER)

(1) This is to inform you of progress made with regard to the idea of an international network for coordinating and stimulating man-computer interaction research, mainly from the psychological point of view. With Circular I (the "Proposal") you are informed on possible main areas of common interest. Most of you have sent in your answers, so that we now have an idea of what the real main interests are and how they are distributed among the foundation members (i.e., the groups represented in Edinburgh). Before continuing this point we would like to inform you of a number of further remarkable points.

(2) In connection with the admission of IUPsyS to the International Council of Scientific Unions the President of ICSU stressed in his address to the Assembly that there is an interest in contributions of Experimental Psychology, especially concerning matters of man-machine resp. man-computer interaction research. On the other hand, this is closely related to aims and projects listed in the Medium-Term-Plan, drawn up by UNESCO's Social Science Sector and related to the activities of ISSC members.

All in all, we have good reasons to assume that our network might receive substantial support from the institutions mentioned. The topic of MACINTER covers social-scientific, natural-scientific as well as technological areas. interdisciplinary par excellence, and presents a great challenge to psychology, too.

(3) To get back to the proposals which have been made by the members of the starting group. On the whole, there was no reply which indicated that the proosed topics were of no interest to the particular group. Nevertheless, there were differences in the evaluation of the different subtopics. Taking great care not to neglect too many of the specific interests, we consider it appropriate to divide the whole range of interests into three major sections, namely:

(1) Socio-organizatonal problems.

(2) Psycho-technological problems and

(3) Basic and applied research aspects of relevance to (1) and (2).

Regarding (1), the widest evaluation has been given to the following subtopics:

1.1 Social consequences of comurerization for man-ower, skill and aging;

1.2 Decentralization of decision processes;

1.3 Task taxonomy, training procedures (programs) and efficiency criteria, social conditions for task-transferability to developing countries;

1.4 User involvement in the development of computer application.

Regarding (2), widest evaluation has been given to the following subtopics:

2.1 Computer-aided design;

2.2 Dialog design (user related criteria for programming languages, and dialog programming used as problem-solving techniques); and

2.3 Adaptivity and learning in expert systems.

Regarding (3), the following aspects are additionally stressed as topics of basic or applied research:

3.1 The measurement of mental load and stress (including psycho-physiological techniques);

3.2 The mental representation of computer structure and functions; and

3.3 Knowledge transactions in the design of digital systems and the managing of "hybrid-inteligence".

We believe that these proposals are a remarkable step forward. They do not contain independent, but interrelated topics, and indicate the interdisciplinary relevance of this type of psychological research.

(4) We have gained the impression that most of the groups that have sent in their answers are intersted in a meeting to be held for exchanging their ideas, making further proosals and--most important of all-informing about their own research results and planning on the basis of examples or demonstrations.

So we would like to invite you to a first network-seminar in Berlin/GDR. Concerning the timing we propose the last week in June 1984 or the first week in July (from June 25 to 29, or from July 2 to 6).

An alternative to this proposal might be in October 1984. The contents of the meeting should be: first day and a half: topics 1.1-1.3; second day and a half: topics 2.1-2.3; fourth day: topic 3; and fifth day: further steps and plans for research coordination and communication.

(5) Please, consider this information as an invitation, and as a call for papers, and to make additional proposals and to actively cooperate in making the first network-seminar fruitful. We are trying to get some funds to pay for costs of stay in those cases where colleagues do not receive sufficient support from their own institutions.

Winfried Hacker
(Technical University Dresden);

Friedhart Klix
(Humboldt University of Berlin).

OUTLINE—SOCIAL AND ECONOMIC IMPACTS OF ARTIFICIAL INTELLIGENCE

Makoto Nagao*
(Japan)

1. INTRODUCTION

Artificial Intelligence has begun to influence various aspects of social activities, especially social intelligent systems and industrial robots. The true impact will come in the future, but there are already both great expectations for and serious fears surrounding artificial intelligence. We have to assess of the impact of artificial intelligence upon society from all conceivable points of view. At the same time we must keep in mind that the utilization of modern science and technology has made a great contribution to the living standards and culture of human beings.

*Department of Electrical Engineering, Kyoto University, Sakyo-ku, Kyoto, Japan.

2. AI IN JAPANESE INDUSTRY

Japanese industries have achieved certain extremes in efficient and economical manufacturing systems, and in product quality. They have high level potential for accepting, the concept of artificial intelligence, for utilizing the technology of artificial intelligence in their manufacturing activities, and for creating a new era for industrial society. But in another way, they have no other ways to improve and innovate their industrial activities than the use of artificial intelligence.

The incorporation of artificial intelligence in the key points of production processes contributes to increased productivity, enhanced production quality, improved labor environments, energy savings, and the suppression of labor-cost increases. Furthermore, artificial intelligence should realize efficient production of a variety of objects in small quantities, create varieties of new products, and enhance the function and reliability of existing products. Other advantages should derive from the higher added value of products, and lower production costs.

The utilization of artificial intelligence, especially of robotics techniques, is expected to elevate the standard of small and medium enterprises in particular. Many small and medium enterprises suffer from a lack of skilled labor. The use of equipment such as high-performance, low- cost, and high-operability robotic machines that incorporate artificial intelligence helps solve the problem of labor shortages while increasing productivity. It will lead to the modernization of small and medium enterprises.

3. AI IN FUTURE INFORMATION SOCIETY

Japan is entering a post-industrial information society. Everybody wants to have benefits from information. The information systems which will be extended throughout Japan sometime in the near future must be able to accommodate the demands of all the people. Therefore the information systems must be "people-friendly", and need the artificial intellligence approaches to achieve the quality which satisifes every persons' demands. The system should be mature enough to accept natural language (speech, handwritten/typewritten characters, wordprocessors, etc.), to provide any information people want, and to be able to both protect privacy and prevent criminal actions in the systems.

4. SOME POSSIBLE FIGURES

Many of the ordinary tasks will be accounted by some sort of intelligent machine. For example:

(1) Intellectual systems

- (a) computer-aided instruction (CAI) in elementary and middle school educations, technical educations at university level, and technical trainings at industries;
- (b) consultation systems for medical diagnosis, shopping, cooking, travel, and so on;
- (c) information retrieval systems of all kinds such as daily news, encyclopedic information, and library information.

(2) Physical systems

- (a) robots for housekeeping, hospital care, guide for the blind, fire extinguishing, deep-sea exploration, and so on;

(b) industrial robots of various kinds;

(c) inspections and testings of every kind.

5. UNEMPLOYMENT PROBLEMS

Recent advances of microcomputers and microelectronics in Japan that have created many new industrial fields, created many new jobs, and improved efficiencies, have caused small unemployment. There have been no particular labor problems caused by the introduction of factory and office automation systems in Japan, and it is believed that no serious problems will arise in the near future from the introduction of artificial intelligence technology to industries and society in general.

Many more will be created by the development of intelligent systems, and robotics systems of all kinds, than the unemployment caused by the introduction of such systems. Artificial intelligence applications will create new market products, which will also create new jobs.

6. NECESSITY OF CONTINUOUS EDUCATION

Introduction of many highly intelligent systems to every part of society may divide the quality of labor into two extremes: one which requires intelligence higher than the robotic intelligent systems can attain, and another, which is simply more intelligent than the systems. The advanced information society will also have the option of dividing itself into two extremes: one which makes different kinds of profits by utilizing *all* the information available from the systems; and the other, which has no ability to enjoy such intelligent systems, and falls into some lower level.

This phenomenon, if it becomes a reality, is very serious for future society. We have to think about some means to elevate the intelligence levels of the people to accommodate the intelligent systems for their individual purposes. The gradual and smooth shift of employment structures to adapt to the future information society is essential, and can only be achieved by the continuous education of the general public.

7. CONCLUSION

This kind of intelligent information society will not be realized in a short time. Robots cannot become so intelligent as to displace many human jobs in the near future. Robots will be introduced first in simple low level jobs in certain controlled conditions. We have time enough to think about and discuss the future potential of artificial intelligence and how it can be dedicated to human welfare and peace. But to reach such goals artificial intelligence must be advanced much more, especially as it becomes:

(a) a strong inference mechanism;

(b) a powerful tool for the representaton of knowledge;

(c) a means of increasing flexibility in decision making as an organic, human-like system;

(d) a means of unifying language, image, speech, and other sensory information through knowledge support; and

(e) a means of development a strong software mechanism.

REFERENCES

[1] Nagao, M., 1983 "Exploration of Knowledge and the Algorithm in Human Brain", presented to the Fourth Convocation of Engineering Academies, Stockholm, May 30-June 1.

[2] Report of the Working Group on Technology, Growth, and Employment. Declaration of the Seventh Heads of State and Government and Representative of the European Communities, Chateau of Versailles, June 4, 5, and 6, 1982.

[3] Japan Information Processing Development Center (JIPDEC): Report of the Investigation of the Effects of Microelectronics on Employments, January 1980 (in Japanese).

OUTLINE—ARTIFICIAL INTELLIGENCE: ITS IMPACTS ON HUMAN OCCUPATIONS AND DISTRIBUTION OF INCOME

Nils J. Nilsson*
(USA)

Artificial Intelligence (AI) will have many profound societal effects. It promises potential benefits (and may also pose risks) in education, defense, business, law, and science. In this talk we shall explore how AI is likely to affect employment and the distribution of income. Some people think that the automation of work by machines and the resulting unemployment present grave threats. Yet, the majority of people probably would rather use their time for activities other than their present jobs, and thus they ought to greet the work-eliminating consequences of AI enthusiastically. We propose to illuminate this paradox by presenting arguments in favor of four rather controversial propositions.

*President, American Association for AI (AAAI), Artificial Intelligence Center, SRI International, Menlo Park, California, USA.

(1) AI (together with other developments in computer science) will make it possible for machines to do all (or very nearly all) of the world's work, both physical and mental, within the next 50 years or so. Although the automation of work might lead to an expanding economy, it is not necessarily true that this expansion will lead to a net increase in the number of jobs, because any new jobs also will be performed by machines. Even today, a significantly large percentage of workers are not really needed to produce the world's goods and services. Thus, the world's current unemployment problems are already systemic and getting worse. With goods and services provided without human work, most people would be unemployed.

(2) Being unemployed has traditionally meant receiving no wages. As a result, until appropriate new income distribution policies are formulated to fit the changed technological conditions, most people would be poor in the midst of unsurpassed wealth (in terms of potentially available goods and services). Fortunately, there exists at least one (and possibly several alternative) stable economic system(s) of equitably distributing the products and services of machines to humans. Such systems will separate income from human employment. (We note that "make-work" schemes are simply one way of distributing income; they have the objectionable characteristic of requiring people to perform unnecessary, and usually undesirable work.)

(3) In addition to providing a means of distributing income, jobs have traditionally given "meaning" to peoples' lives. They have provided a social focus and context as well as an opportunity for individuals to gain recognition. In the environment of massive unemployment, most people would miss these advantages. However, there is ample evi-

dence that it is possible for humans to live rich, fulfilling, and rewarding lives even though they are not under pressure to perform work in order to secure an income.

(4) Probably the most difficult challenge presented by the possibility of automating work is that of smoothly transforming our current economic system to one that separates income from employment without inflicting either economic or psychological hardship on individuals living through the transition. (In a certain sense, we are already in the transition.) We believe that enlightened leaders and an informed citizenry can pursue policies that minimize human distress during the years ahead.

SOCIAL AND ECONOMIC IMPACTS OF ARTIFICIAL INTELLIGENCE

Roger C. Schank* and Stephen Slade
(USA)

1. AI IN THE 1990'S

What are the most likely core research areas in AI and what results do you expect?

There are many branches of AI and each has its own research agenda. However, there are several general problems which affect efforts in the main areas of AI: natural language, vision, robotics, speech recognition, and expert systems. The first is learning. It is widely accepted that AI programs require very large amounts of knowledge in order to be effective. Once we know how that knowledge is to be represented, in general, we must then fill the computer with specific facts. These facts about the world are innumerable and ever changing. A smart program must be a learning program. It must be able to

*Computer Science Department, Yale University, PO Box 2158, Yale Station, New Haven, Connecticut, USA.

respond as a person does to new information and situations. It must adapt to changes in its environment. It must be able to grow on its own.

Another aspect of the same problem of the large amount of knowledge in these programs is both theoretical and technological. The theoretical problem is *How can that knowledge be organized*? The main current computer systems that have tremendous amounts of information are very large data bases. These systems tend to be fairly homogeneous and inflexible. By contrast, AI knowledge bases tend to be heterogeneous and variable. What is the proper model then for making these AI programs orders of magnitude larger? This leads to the technological question of *How can we actually build such large, complex programs*? Our machines are getting more powerful and allow us to build larger programs. It is still not certain the machines of tomorrow will be able to keep up with a geometric growth in AI programming requirements. There will always be a need for more powerful machines for AI research. The problems of knowledge representation, acquisition, and organization will continue to be at the core of AI for decades to come.

We have taken the view in our own research that large, complex computer programs should be psychologically motivated. That is, the computer programs should be models of human cognitive behavior. There are several reasons for this point of view. First, people provide us with an existence proof for intelligence. We know that people can understand language, solve complex problems, and learn through experience. Therefore, we know that such behavior is *possible*. There may be other ways of developing intelligent computer programs, but we wish to try to model human cognition in as direct a way as possible.

Second, since our programs are psychological models, we can take advantage of psychological experiments and methodology in formulating our theories. We can appeal to psychological evidence in developing our programs, and, in turn, our programs and theories can lead psychologists to try new experiments.

Third, and most important, we want our programs to be psychologically correct so that we can learn more about the human mind. As computer scientists, we are well aware of the intrinsic importance of intelligent machines, but we must aver that an understanding of human intelligence is a far greater goal than the creation of smart computers. Our work in modeling human cognitive processes has focussed attention on many facets of cognition including language ability, learning, memory, motivation, emotions, interpersonal relations, beliefs, and planning [Schank and Abelson 1977, Schank 1982].

In the next decade we would expect that the two central problems of learning and memory organization will be explored in greater depth. It is very difficult to predict exactly what the results might be, since the answers are obviously unknown. However, these two areas should have the greatest impact of the field as a whole, since they touch on practically all areas of AI.

2. AI AND FUTURE TECHNOLOGY

What impacts of AI on other sciences and on technology do you expect?

The initial effect of AI on other scientific enterprises has been the use of AI's tools and methods. One can already witness the application of LISP software engineering techniques and hardware to other areas, such as CAD/CAM, text editors, and compilers. At Yale, we have ongoing research efforts in all three, which are quite separate from our AI research program.

Just as AI can no longer claim sole ownership of LISP, it can neither be proprietary with its basic paradigm of process models of human cognition.

Numerous other fields, including psychology, philosophy, linguistics, and anthropology, are investigating AI theories and models of cognitive processes. This new interdisciplinary focus on cognitive behavior is called cognitive science. We see this as a very exciting research area, ancillary and complementary to AI.

Expert systems offer another way of relating to other fields. The AI methodology of modelling expert behavior can be applied to other scientific endeavors. For now, the achievements in other fields by AI programs are rather modest, when compared to human performance. We would expect though that future programs--especially those that can learn from experience--will be able to solve problems that might have baffled the human expert.

At Yale, we are developing several such expert systems that are meant to adapt to new situations based on prior experience. A fundamental assumption of this research is that the basis of expertise is experience itself. Our computer programs build a large memory of previous cases in a specific domain on which they can draw when a new problem arises. The previous cases are the source of general rules which can handle the normal situations. This is comparable to the common rule-based model of expert systems. However, when a novel situation occurs, the program should be able to analyze the new event in terms of previous events that shared features. Thus, even though no particular rule was invoked, the program would still be able to reason about the problem at another level.

In the next 50 years, AI may well come to be viewed as a type of metascience, like mathematics. That is, AI will be used as a tool directed toward problems in numerous scientific and technical domains. A program that can reason and reformulate its knowledge according to experience will be a powerful tool in any application.

There is a belief among many people in AI that computers will be able to become far more intelligent than people. This is a debatable proposition, but there is one aspect to the argument which may pertain here. Computers can be made to devote all their resources to a given task. They are not subject to the usual array of human distractions such as hunger, thirst, boredom, or exhaustion. Furthermore, if you have developed a program to explore a certain problem, it is very straightforward to run that same program on 100 or 1000 computers at once--until one of them comes up with the answer. It will most likely happen that a computer program will make significant discoveries in medicine or economics or mathematics. In the next century, an AI program might very well be awarded a Nobel prize. A person will have written the program, but the program will be the one to have made the award-winning discovery.

3. AI AND SOCIETY

What social impacts do you expect?

In the next 10 years, there will probably not be a significant social effect from AI. Surely, the proliferation of home computers will continue and the public will become more aware of computers. However, the home computers will not be able to support the large, complex AI programs. People will begin to realize the great difficulty in making a computer that can reason as a person can.

People will have much greater exposure to computers, both at home and at work, and there will be a greater need to make computers more accessible. The design of the man-machine interface will be a vehicle for AI technology in general and natural language processing in particular. The computer should communicate effectively with the user, understand the intentions of the user, and learn about particular users from repeated exposure.

Thus, the relationship between man and computers will continue to evolve over the next 10 years and AI will be able to offer ways to normalize that relationship by making it easier for people to work with computers. There will be a greater public understanding of the problems of AI and also the possible applications.

One possible social impact of AI is unemployment. Over the next ten years it is not likely that this will be much of a concern. Viewed as a resource allocation problem, the use of AI programs will most likely be for applications that are too expensive or dangerous to use people. AI programs will be fairly expensive (compared to word processing and other personal computer applications) due to the hardware requirements and the fact that there are not many people trained in AI.

In the long term, AI programs will take care of tasks that are better suited for computers, allowing humans to devote their time to personal services. For example, computers will play a major role in education in the coming years and the role of the teacher will change considerably. We would expect that the teacher in the next 10 or 20 years will be able to concentrate less on preparing lessons and correcting homework, and spend more time dealing with the interpersonal and social aspects of school. AI will play a growing role in education. The current educational software is at a very primitive state. These programs often consist of little more than an electronic book in which the child types RETURN to see the next page. One active area of AI research is ICAI (Intelligent Computer Assisted Instruction). An ICAI program builds a model of the student and tries to understand the misconceptions the student may have about the problem. Natural language abilities will clearly be of great use to the educational programs of the future.

4. AI AND ECONOMICS

What economic impacts to you expect?

AI programs require quite large amounts of computer power. As hardware becomes less expensive, more powerful, and widely available, AI programs can finally be economically feasible in the marketplace.

As mentioned above, AI plays a growing role in the design of software systems--particularly the user interface. Given the increasing power of hardware, these commercial systems will become ever more common and widespread. The demand for these intelligent systems will be tremendous. In ten years, an AI interface will be the *sine qua non* of any major software system.

There is a problem with this scenario though. There will not be enough trained people available to build the systems. Even today, there is a critical shortage of manpower in AI. Only a limited number of people in the world today are actively engaged in AI endeavors. Expansion of the field will certainly occur, but it is likely that the average quality of the work will diminish.

It may seem odd to some people to state this problem. Aren't there thousands and thousands of computer programmers in the world? We see job training centers popping up to turn unemployed steel workers into programmers. We recognize the demand for programmers and apparently there are many efforts underway to meet this demand.

However, the case is quite different in AI. An AI programmer requires considerably more training and ability than an applications or systems programmer. The problems are not well formulated. The answers are not absolute. The usual traning for an AI programmer (who should already be a proficient programmer) involves a period of apprenticeship at a research lab. This is vaguely comparable to an internship or residency in medical education. The problem though is that most medical schools in the country graduate more doctors each year than the total number of AI researchers trained in the entire world.

Thus, the amount of progress in the field and the application of results is severely limited by the number of researchers. This shortage of trained researchers and practitioners will have a noticeable damping effect on the economic impacts of AI.

5. AI IN THE 21ST CENTURY

If you would speculate about the long range AI research results and impacts, what could they be?

Imagine the world 100 years from now. There will be robots that can understand speech, read the newspaper (if there are newspapers), recite Shakespeare, and assemble a bicycle. These robots will be close to R2D2 or C3PO than to any current machine. They will be our servants.

This is the science fiction vision. It is probably true, but it is surely a small part of the picture. Saying that AI research will culminate primarily in robot slaves is like saying that the main result of discovery of electromagnetism is color television.

Clearly there is a technological aspect of AI that drives much of the research. However, there is a very important scientific aspect of AI which involves the study of the mind. What is the organization of human memory? How do people learn? How do people make decisions? How do people think?

It is the answers to these admittedly grand questions that hold the greatest potential benefits to mankind. AI as a science may be able to point the way to discoveries in other areas. It may be possible to apply such advanced AI programs to a variety great human problems--medical, educational, political, economic, and social.

For a case in point, imagine what a political candidate might expect from AI in 100 years. There could be program which could advise the leader on the course of action to take in a given situation. You could think of this program as a political expert system. The program could devise campaign strategy, draft speeches, send out press releases, prepare position papers, conduct polls (via electronic mail), develop media commercials, and analyze the opponent's record.

Now let us take it one more step. A computer program that could display such an encyclopedic knowledge of the campaign and politics should be a candidate *itself*. While we do not wish to suggest that people will end up voting for computers, it is reasonable to predict that computers, with the help of AI programs, will play a growing role in the way in which our governments are run. For example, an AI expert system on preparing the budget will most likely evolve over the next century. Such a program would not be a mere spread sheet calculation program, but would analyze the line items of the budget from a political perspective. A country's budget is the national agenda, an allocation of the country's resources, and a reflection of the political values of the country. An AI program to prepare the budget should know what the goals of the leadership are how those goals can best be enacted through the budget. The program should know about the political feasibility of passing various provisions and what impact they will have on society.

This political advisor is but one of many types of applications of AI that could have a much greater impact on people's lives than a robot which brings you your coffee every morning. The horizons for profound applications of AI are wide.

6. AI TODAY

Which research areas of AI would you as part of the AI community support mostly, and why?

As indicated above, the areas of learning and knowledge organization are of central importance. These touch on many of the sub-fields within AI including natural language processing, vision, speech recognition, expert systems, and robotics. We have also mentioned the reasons for developing a psychological methodology for modelling cognitive phenomena.

One main requirement for productive and useful AI research is to work on large, real world problems. There is a broad consensus that no major breakthrough in AI will be found in a small, concise program or in toy domains. AI programs should be explorations in complexity and should tackle the world as it exists.

One extension to this argument is that AI researchers should look at problems that span more than one sub-field. Thus, there should be efforts in integrating expert systems and natural language, or speech recognition and vision. Researchers tend to shy away from these large problems saying that there are still many small problems that have yet to be solved. That may be, but by looking at larger problems, we may gain important insights into the nature of intelligence and cognition.

REFERENCES

[1] Schank, R., 1982 "Dynamic Memory: A theory of learning in computers and people". Cambridge University Press.

[2] Schank, R. and R. Abelson., 1977 "Scripts, Plans, Goals and Understanding". Lawrence Erlbaum Associates, Hillsdale, New Jersey.

AI–SUBJECTIVE VIEWS, FUTURE, IMPACTS

Tibor Vamos*
(Hungary)

"If your intellect lets you down we approach the experience which is weaker device and not so lofty..." (Montaigne: On the experience, Essais III, Ch.15.)

1.1

First of all I would be happy if we could agree on a less gorgeous but more solid notation of our topic. I know well that AI is now an accepted and "well-sounding title--as cybernetics was 20-30 years ago--but it is giving rise to sci-fi expectations especially with people to whom we would like to sell our achievements and ideas, and introducing charlatans to a field which is a real R&D area for hard work, the typical 1% inspiration and 99% perspiration, discipline and self-controlled self-estimation. Misleading is even more dangerous in a period

*President, International Federation for Automatic Control (IFAC), Computer and Automation Institute, Hungarian Academy of Sciences, Kende utca 13-17, Budapest, Hungary.

when for the first time moderate successes, limits of recent possibilities have been reached by people devoted to this problem solving, machine intelligence, etc., are more attractive from this point of view which does not intend to attract cheap propaganda.

This short note on philosophical approach indicates my further remarks. I would like to select those areas where relevant results are obtained, i.e., solved or at least reached realistic problems. In this context I emphasize not only the primacy of practical applications (and not demonstrations) but the measure or problem complexity, too. I am sure that this is really the kernel of our estimation: a solver which enables or helps to solve probelms of real-life complexity, those ones which are unsolvable or hardly solvable by the conventional straightforward computational methods.

As any scientific discipline, our field grows in two directions: methods (methodologies)--a disciplinary way and applications. If they are not related (not necessarily simultaneously and directly), they loose their relevance.

According to my opinion and the above considerations two major fields emerged in the last period as successful, relevant and promising ones: vision and knowledge engineering (this latter includes the term expert systems and somehow the problem of understanding).

Vision is now a practical device on the market (its hardware and software), several products are available for different purposes. The most important ones are:

(1) picture processing

- -- special applications for remote sensing, i.e., meteorology, agriculture, natural resources, archeology, etc. (devices intended to automated massacres are beyond my scope);

- -- biomedical applications (especially recognition of morphological characteristics, i.e., shape, texture);
- -- inspection especially in technological processes (roughly the same tasks);

(2) robot vision

- -- object recognition (2D, 3D, color, partially invisible, shaded, overlapping, etc.);
- -- situation recognition (range finding, objects in movement, relations of several objects).

Here we ought to make some remarks. It is very typical and surprising that starting research projects have very ambitious (highly intelligent) general goals, but as the project approaches realization, it restricts its intentions to a rather pragmatic and economic solution. Most vision applications would not be considered as real AI results by some extravagant AI philosophers. These applications use several simple edge detection, region growing and selecting techniques combined with the computation of easily computable geometric characteristics (area/perimeter, momenta, connectivity, etc.). Easy teaching methods, standard object modeling, calculation of texture features (intersections, correlations, etc.) are the major tools.

Even the search procedures are mostly very simple for decision in case of ambiguities. The alternatives are well-known in advance, a learning period evaluates the most relevant features of the special application cases, some statistical or fuzzy parameters help in branching and accelerating the search, if the machine decision is not sufficiently convincing (beyond some level of certainty) the human operator must intervene--a real process should be reliable. Most of these applications evolved to feasibility due to microprocessor-based work stations, i.e., by solutions which are beyond the 10, maximum 20,000$'s range, where camera, other peripherals, software are included.

A fast further advance may be predicted. As the applications proliferate, cheaper, strictly dedicated and limited recognition and inspection systems are expected to replace many monotonous, not really human workplaces in inspection, selection, simple evaluation, transfer, assembly and other industrial operations (painting, welding, etc.). Vision replaces many more expensive tools: pallet fixture promotes the flexibility of manufacturing processes. Most probably in a very near future (5-15 years) machine vision will be as natural ingredient of work stations as microprocessors are today; a big industry with a wide variety of modular products will be the basis. Applications will be extended to commerce and services. Some people dream of a household robot in the 90s; I, however, being personally engaged in doing such work regularly at home, am not so optimistic.

What kind of research is in store for us to achieve these goals? Computer industry does its role in providing hardware--the present day microprocessors are more than sufficient for most applications, very fast parallel processors for picture processing (especially tracking moving targets and such which are not so easily recognizable) are on the horizon. Optical industry mated with semiconductors promises new devices for input. Areas where I would stimulate efforts are the following:

-- parallel algorithms of recognition;

-- character recognition, cheaper and more powerful optical character readers;

-- a more systematic attack on morphology. This would mean a real "AI"-like approach: to transcribed classical morphology descriptions of biologists, geologists, etc., into a formal metalanguage and transfer this procedure into recognition, discrimination;

-- more 3D, knowledge-based help for guidance of moving objects (transfer machinery) in industrial environment and for robot-assembly and other kind of complex manipulation, which needs flexible adaptation;

-- a computational link between design on displays and automated operations (machinery, robot operations). This task leads towards knowledge-engineering but it has a relevant visual interface, the understanding of images composed by various methods (computer graphics, 2D-3D transformations, photos, etc.).

Many tasks provoke improved methods of handling pictorial data. Some people think that DBM is the same for any kind of data. I do not believe in the efficiency of unified methods in this respect the requirements are especially high.

Vision was selected from other kinds of perception. This is the most relevant aid of human activity and the most advanced, too. The other important field is voice recognition. We rather clearly see the distinct borders of complexity magnitude: a recognizer of a few tens of words is commercial and an easy task for any expert engineer. The next order of magnitude is rather expensive but feasible (few hundred words). One order of magnitude higher appears already to be a Himalaya. The reasons are now clear and discussed by many authors--by those who could reach some altitude records by terrible efforts and by those who simply looked at the basic problems of complexity. My personal impression is that in the next decade we shall see a lot of very practical applications of strict, uncontexted (or simply contexted) military commands like spoken languages, a new tool which help people in communication with computer-based devices and leave their hands free for action, too, but a feedback and a check-confirmation operation will be needed in ever sensitive case. The day of the replacement of a good stenographer-typist is still far distant. This means that the main emphasis of activity should be on a low context target of maximally some hundred words. (One speaker--multispeaker) At many work (and other personal) stations this will trigger the next revolution: the mass replacement of keyboards but not of human writing.

The second area emphasized was knowledge-engineering (expert, knowledge-based systems, related topics). A tutorial paper was attached which reflects my ideas on the topic and another paper will be available in the next weeks, before the first meeting, with a special focus on applications in flexible manufacturing.

The spirit of these papers reflects my views: this new technology of including human expertise into computer-aided systems will not be a replacement of human activity but a higher level man-machine interface. Citing the conclusion of the first paper: "We cannot predict what the future perspective of the expert system is. They are not going to replace genuine human intellectual activity but are promising helpers in coming man-machine systems. Most probably a merging of recent trends in information systems and expert systems provides a new stage of information technology. Relational and distributed database, user-friendly man-machine communication, information service-networks will provide a practical application spectrum without the gorgeous nomenclatures which are nowadays popular in selling science and products."

1.2

Most of the further questions found a natural answer under 1.1. The corollaries are sometimes automatically emerging from the statements which reflect the views and prejudices of the author. Recognition will be a relevant step in any kind of automation. In process control this role was played by the process instrumentation (thermometers, pressure, composition and other sensors, transducers), but for most activities this link failed, it needed really a part of human intelligence (perception) but was never creative, really human. The mass-processing of inputs can revolutionarize also the possibilities of other sciences a storage, processing and retrieval of immense quantity of samples (medical, historical, archeological documents, etc.). It can help in mass screening of population, in exploration of mineral resources, etc.

Concerning expert systems I can only repeat my last sentence of the previous paragraph. This is a new and substantial help in self-rethinking of existing knowledge in various fields, in distribution of this knowledge all over the world, extending the global human knowledge base to a much broader community and by a real-time access way. It will be a great help (if properly used) in education, training, in any kind of maintenance (also biomedical), in improving services, it will be a logical further development of computer and communication revolution, if mankind will realize that modern technology can be used also for the people and not mainly against the people!

This relates also to the social impacts. I had several opportunities to express my opinion on more details on the problems of unemployment related to automation. Unemployment is an incurable disease aggrieving poor countries which cannot provide working places, nor afford appropriate training and retraining of people. The politico-economic decision whether they want to rearrange human activities as a consequence of progress in technology, or not, rests on the rich. The new needs for more social care, infrastructure, services are far from saturation. It is really a problem of economic strength and economic strengh is mainly based on the level of technology--automation.

The other relevant social impact is whether those new trends in automation which are characterized by our topic can be helpful in increasing democracy, individual freedom or on the contrary, they will be new weapons of Big Brothers against the individual and new means of uniformization, manipulation of masses. This is an open question and cannot be answered by our science. We can only tell that both opportunities are feasible. Because many papers, fictions are written on the disastrous alternative I refer to the optimistic one. I have to mention that these views were outlined at a General Assembly meeting of our Academy and popularized in my country as adopted ideas. As a third reference this paper is also enclosed. I quote here a part concerned:

"**State and Citizen**

Our entire administrative structure is affected by what we developed somewhere with the appearance of the written word, in an era when the clerks were literate and the people not. The development of information science, telecommunication networks, data banks and information systems makes possible and imperative an administration which is entirely different from the existing one. On the one hand this will be a demand of our being able to work competitively, and on the other hand, it renders a new possibility for improving the quality of life, developing the socialist democracy, and increasing the attraction of our society.

In connection with computerized information dangers have been stressed for a too long time which make possible for the centralized, bureaucratic state to control its citizens. Much less emphasis was laid on the opposite outlook, i.e., the extension of the rights of citizens, the transformation of public administration mainly into a real service, after the initially centralizing direction of computerization a much more forceful decentralization. With the development of nation- and world-wide information systems centralization and decentralization become modified to an extent where it is easy to imagine in today's sense countries without capitals and international systems functioning without centers, which are not hierarchic but cooperative as hoped for by the pioneers of socialism.

All this is not the distant dream of futurologists or sci-fi novelists but the possibilities of the next one, two, in their remotest contexts three decades, for which preparation is already today rather late than timely. We are very close to the situation when in the affairs and daily conduct of life of every citizen the actual statutes, regulations, those decisions of the governmental administration which are interesting for him, are permanently and openly displayed in any moment of the day on the screen of his home television set. *To the several thousand years old mythical curtain of the alienated state and its clerks, offices can finally be put an end.*

Thus the decisions and regulations can be made accessible and understandable to everybody; questions can be marked off exactly which can be automated on the basis of criteria of judgment accessible and understandable to anybody, do not require arrangements, exclude forcefully trickery, influencing and granting of privileges; decisions can be separated that demand human judgment and behind which consequently the person or body taking the responsibility appears in a direct form and connection. Just a socialist production and distribution can be realized in fact only at a high level of material production, computerized information opens up possibilities to socialist administration.

This is how production-social structure-organization-administration can arrive at a new synthesis. For us to be able to hold out in the world competition in the next decades, this link of chain is decisive.

We are facing the possibility of a society of a qualitatively different composition and of a qualitatively differently administered society. The change has commenced, and we cannot be passive concerning it."

We have to look at the problem which was rather neglected until now. In our efforts for humanizing the work we focused our attention to the possibilities of avoiding tedious, monotonous jobs and by that way increased the creative content of the work. We expected a fairly unlimited extension of human intellectual activity for people who did not do that before. This expectation is mostly realistic: we see how mankind stepped out of illiteracy, how fast children can master computers, how many decent skills can be learned if they are taught in appropriate way, at an appropriate age, in an appropriate environment. On the other hand there is a relevant proportion of population with retarded development in intellectual abilities due to prenatal or perinatal reasons, they can be integrated into the society only by low level useful activities. May be this will be a major new task both for medical people and sociologists, the problem of specially handicapped in a new society. May I mention that I am involved now in a project which attacks one part of this problem--just using expert systems for help!

1.3

Reference was made also to economy. I mentioned the evolving of future industries (e.g., vision), new professions are emerging (e.g., knowledge engineer). Most important is the fact that wealth and welfare stem from higher technology, we have seen the fragility of fortunes based on lucky natural resources without the culture of technology and the relative stability of the highly advanced countries in spite of all unpleasant symptoms of recent economic recession. Our field is just one of the most promising, influencing and promoting other disciplines, advancing technologies such that its impact on economy ought to be mostly positive. There is a real warning in connection with introducing its results in advanced countries: the gap can be even wider, which is undesirable for both partners. A proper distribution of these results can help developing countries, especially in training and in service and maintenance of more advanced equipment and systems. An open commercial availability of computer-accessible expertise can promote international trade, standardization, the application of new equipment and by that a higher level of international trade. As software became a major commercial item, a more sophisticated version than a network-service available expertise will be even more--if (and I emphasize once more, never enough) properly used!

The answers of questions 2 and 3 are included in the first part of this position paper and in the attached material. The paper on cooperative systems which is published as an IIASA Collaborative Paper is an addition to this topic.