

## SOCIAL IMPLICATIONS OF INTELLIGENT MACHINES<sup>1</sup>

Margaret A. Boden

School of Social Sciences  
University of Sussex  
Falmer, Brighton BN1 9QN

The much-discussed issues of privacy, unemployment, leisure, centralization of political power, and military misuse of technology are raised by work in artificial intelligence no less than by applications exploiting the "brute force" of computers. But this paper focuses specifically on matters associated with the social use of intelligent machines in particular. Some current and predicted developments in machine intelligence are described and possible ill and good effects these may have on society outlined. Precautionary measures that might be taken in the writing and presentation of programs to forestall the social dangers implicit in this area of research are examined.

Key words: Artificial intelligence; social implications of computers.

### 1. Current Achievements and Future Developments

Computer hardware gets steadily cheaper: it is predicted that by the year 2000, a 65 k bit silicon chip capable of 20 million instructions a second will sell for one US dollar. But the social application of intelligent machines also demands advances in software, such as more powerful programming languages and improved organization and use of knowledge. Work in artificial intelligence has shown that the problems of organizing and accessing large data bases will not be quickly solved. The early optimism in the field has waned accordingly: already by 1962, the mediocrity of chess programs was being stressed by someone who in 1957 had predicted that a computer would be world chess champion within 10 years. And much current research aims to develop organizational principles whereby a knowledge-domain can be economically represented and appropriately addressed without triggering a combinatorial explosion[1].

Nevertheless, professionals participating in a multistage "Delphi" forecasting exercise have predicted that within the next 30 years (in many cases, within only 10 or 15 years), social applications of artificial intelligence will be widely, i.e. commercially, available[2]. In general, the prototype is expected 5 ± 2 years ahead of the commercial version. The dates I shall mention are

taken from this Delphi study (although my own view is that these predictions tend to underestimate the difficulties involved).

The applications forecast run from robot housecleaners, chauffeurs, and industrial workers, through programmed gamers and storytellers, to automatic teachers, physicians, legal justices, marriage counsellors, and literary critics. In all these cases, the emphasis is on reasoned and flexible judgment on the program's part, as opposed to the storage and regurgitation of isolated facts, or the repetitive performance of a fixed sequence of discriminations and movements.

For example, the computer diagnosticians of the 1980's will not simply store lists or symptom-diagnosis pairs, or prescribe treatment in a blindly dogmatic (and apparently "objective") fashion. One current prototype is the MYCIN system, an interactive program that simulates a medical consultant specializing in infectious diseases [3]. It engages in question-and-answer conversa-

<sup>1</sup> Based on Chapter 15 of Artificial Intelligence and Natural Man, New York: Basic Books and Hassocks Sussex: Harvester Press, 1977. From The Radio And Electronic Engineer, Vol. 47, No. 8,9. Reprinted by permission of that journal. © The Institution of Electronic and Radio Engineers, 1977

tions (lasting 20 minutes on average) with doctors needing specialist help, and in 75% of cases gives the same counsel as a human expert. The physician asks MYCIN for advice on the identification of micro-organisms and the prescription of antibiotics, and also for explanations of its advice expressed at the appropriate level of detail.

MYCIN's explanatory capacity enables physicians who disagree with specific aspects of the program's clinical rationale rationally to reject MYCIN's advice. It also helps nonspecialist doctors to learn more about the complexities of diagnosis and therapy in this class of diseases. And it allows human consultants to make general improvements in the program, by telling it about relevant knowledge that they realize in specific cases to be missing or inadequately stated.

A program like MYCIN involves artificial intelligence techniques. Quite apart from its (rather restricted) natural language understanding, its ability to explain itself on many different levels of detail implies a self-knowledge of its reasoning and goal-structure that are crucial to intelligent thinking. And its ability to learn by being told implies some mastery of the problem of making spontaneous inferences on the basis of input information, though admittedly only for a very limited area of discourse.

Many features of MYCIN would be embodied also in programmed legal arbiters (prototype predicted by 1988). These would not only search for relevant legal precedents in the judicial literature - a far from trivial task - but would also offer legal advice. Like MYCIN, legal arbiters will preferably be used to augment human judgment, rather than replace it. Accordingly, like MYCIN, they should when appropriate offer several (reasoned) alternative judgments, not just the one of which they are most confident. MYCIN's assessment of degrees of confidence is not a mere statistical probability measure. It takes into account psychological factors about the evidential relations of beliefs, factors that philosophers of science have considered in regard to "confirmation theory." Legal programs, too, would have to incorporate more subtle concepts of confidence and evidence than Bayesian probability, in order to avoid judicial absurdities of various types[4].

Flexible planning, and an intelligently structured representation of knowledge, is used by the Computer-Based Consultant, a system designed to give on-the-job advice about how to assemble a machine to novice mechanics having varied levels of expertise[5]. The program uses the specific queries posed by the human novice as cues directing it to answer at one level or another. Thus, it tells one mechanic simply to "Replace the pump," but advises a less experienced person first to "Remove the 4 mounting bolts at the base of the pump using a 3/8-inch open-end wrench." When a human failure occurs because of an unexpected happening (including but not restricted to previous mistakes on the novice's part), the Consultant can question the mechanic in an intelligent fashion so as to locate the difficulty. In most cases, this will not be a simple matter of asking "What's the trouble?" since the novice usually does not know just what has gone wrong, and may claim to have

followed the program's advice to the letter. Because the program's representation of the semantics of this domain includes detailed and intelligently structured knowledge, in the majority of cases it is able to spot the trouble and work out a way of putting the human worker back on the right track.

Future automatic tutors will be more sophisticated than today's computer-aided instruction systems, even those that allow differential branching of the "syllabus" according to the student's mistakes and queries. Like MYCIN and the Computer-Based Consultant, they will be able to initiate and answer questions at various levels of detail, according to the pupil's range of expertise. They will concentrate on asking probing questions enabling them to model the student's understanding of the topic, and will devise an individually tailored tutorial strategy designed to build on this understanding in fruitful ways. Such programs presuppose an intelligently structured representation of the knowledge concerned, in its various aspects and degrees. Commercial availability is forecast for 1988, and exploratory programs already exist[6].

Public access to a powerful computer would be required for many of the predicted social applications of artificial intelligence. John McCarthy accordingly has forecast the widespread use of video-teletype home information terminals, linked to a national network of time-sharing computers[7].

A speech-understanding program would allow the user to speak requests instead of typing them[8]. At present, this can be reliably done (for instance, by the HEARSAY system[9] for playing voice-chess) only by using a very small and deliberately distinct vocabulary, with unnatural pauses between the words. This is because (as in the interpretation of visual scenes) the continuous stream of normal speech has to be sensibly segmented into individual words before it can be understood - but, as artificial intelligence research has shown, it has first to be understood in order to be segmented. Even the individual sounds cannot be distinguished on purely auditory grounds, but only by reference to the wider linguistic context. However, commercial uses are forecast for 1983.

According to some forecasts, the stationary home terminal will in 1995 be supplemented by a domestic robot. It should not be necessary to clear the floor so that the robot can vacuum it: the robot itself will supposedly be able to recognize the objects littering the room and put them in their proper place - or, by default, in one corner. Current achievements suggest that a robot might fairly soon be able to recognize a waste-paper bin wrongly placed by the window, or even high up on the table, so as to reposition it by the desk[10]. But only if the room is otherwise unnaturally tidy, and only if the bin, window, table, and desk are of uniform type.

Industrial agriculture robots (forecast by 1980) are less fanciful. These would not have a fixed sensorimotor capability, like the automated machines of today. Instead, they would be flexible in operation, and could learn new tasks. They would learn not by being laboriously reprogrammed, but by being shown examples of (for instance) new

machine parts and machine tools, and by being given an outline sketch of the desired procedure which they would then elaborate in detail. The sort of planning facilities used by the Computer-Based-Consultant, or programs like BUILD, would be crucial here. BUILD plans how to assemble brick-structures which may require "creative" steps such as using a temporary scaffold, counterweight, or support[11]. Features like potential stability and steadiness of movement are continually taken into account in deciding what to do, and if a safer method can be found of constructing the desired building, then that method will be used. If it turns out that the "safer" method is not safe after all, BUILD can alter course accordingly, meanwhile losing none of the information gained during the aborted attempt. Slight local difficulties can be recognized and overcome without affecting the general strategy, but graver obstacles prompt radical replanning.

There are already some mobile carrier-robots, and programmed hand-eye systems that can assemble simple machines from components they learn to recognize by being shown examples or that (like MIT's COPY-DEMO[12]) can build visually demonstrated structures out of a "warehouse" of familiar parts. Moreover, a robot linked to the Computer-Based-Consultant would be able to assemble a pump even if someone else had left the job half-finished, or had wrongly positioned some of the parts. This flexible behaviour is very different from that of current industry's computer-controlled devices, such as the Unimate, which require everything to be in the right place at the right time.

## 2. Possible Effects on Concepts of Self and Society

The potential social influence of artificial intelligence is ambiguous, for two reasons. First, specific applications (as of any science) may be used for good or evil ends, and may have unsuspected side effects. Second, much depends on the background human context, including the general public's implicit philosophical assumptions or "image of man." I shall first discuss some possible bad effects, and then sketch potential good effects of intelligent machines.

There is a widespread suspicion of artificial intelligence among the general public, based on the common philosophical assumption that regarding man as a computational mechanism entails denying human subjectivity, individuality, and moral freedom. This suspicion has been voiced by critics of urban industrialism such as Theodore Roszak and Herbert Marcuse[13], and by psychologists in touch with laymen seeking practical help. For example, this complaint is from Rollo May, a therapeutic and counselling psychologist[14]:

"I take very seriously ... the dehumanizing dangers in our tendency in modern science to make man over into the image of the machine, into the image of the techniques by which we study him... A central core of modern man's "neurosis" is the undermining of his experience of himself as responsible, the sapping of his willing and decision."

May's point is that choices made without confidence in their possible relevance to eventual action are unlikely to be effective - or perhaps

even to be made at all. For if a man's self-image represents himself to himself as an autonomous purposive creature capable of pursuing certain ends, then it can be used to generate choices and guide his action accordingly. Even in machines, as is evident in planning programs such as BUILD, the internal representation of the possible modes of action that are available to the system can be crucial in directing performance. But if a "depersonalization" of the self-image occurs so that the self is no longer seen as a truly purposive system, then relatively inhuman, "Pathological" behaviour can be expected in consequence.

This type of degenerate self-model is encouraged by artificial intelligence in general, given the popular (though mistaken) philosophical assumption of total incompatibility between mechanism and humanism. What May calls "the undermining of one's experience of oneself as responsible" may therefore be exacerbated by the development of clever programs.

As well as having immediate consequences in one's personal life, this sapping of willpower can have widespread social implications. For example, the political institution of participatory democracy assumes an ascription of responsibility to individuals which fits ill with the dehumanized image remarked by May. Consequently, providing citizens with home terminals, with the partial aim of enabling them to vote and express their political views without leaving their fireside[15], might subtly undermine their sense of civic responsibility. This is the reverse of the intended effect, which is to encourage individual citizens to engage more fully in democratic government, thereby lessening the common feeling of alienated helplessness with respect to the governmental process. (Of course, to feel less a cog is not necessarily to be any less a cog: the home terminal might function as a subtle form of social control, damping down dissent by contenting people with an illusory sense of political participation.)

While any intelligent program may have a dehumanizing effect on people who see an unbridgeable metaphysical gulf between themselves and machines, as many people do, some would be especially open to this criticism. For instance, the prediction that by the year 2000 automatic interviewers will be used to aid the diagnosis and treatment of psychiatric patients is a suggestion that many would spurn. Thus Joseph Weizenbaum, the creator of the early conversational program ELIZA[16], has bitterly denounced the "obscene" idea of employing programs in clinical situations[17].

In medical contexts dealing with basically physical illness, artificial intelligence may be more welcome. The MYCIN system has not yet been tried out in a clinical situation, though the authors report some resistance on the part of clinicians to the idea of using the program. When MYCIN has achieved a 90% match with human experts, it will be introduced experimentally into a hospital; only then will its authors be able to see who uses it, how often, and what effect it has on the prescribing practices of doctors and the clinical status of their patients. But a much simpler program for the diagnosis of peptic ulcers has been used on patients, who often claim to prefer this

diagnostic method[18]. Ironically, they describe the machine (with which they communicate by teletype) as more friendly, polite, relaxing, and comprehensible than the average physician.

These chastening observations about the superiority of the personal habits of programs over those of human doctors can doubtless be expected also with regard to automatic lawyers, bureaucrats, and teachers. While perhaps appreciated in isolated interactions of a tedious, technical, or embarrassing nature, this imperturbable mode might come to be consciously spurned in human relations in general, with a consequent increased emphasis on emotional spontaneity. Alternatively, the opposite effect might result: the blandness of one's guest-computer might come to be emulated in one's own mode of expression - or that of one's children. (The possible ill effect on children was cited in the Delphi survey as a disadvantage of the domestic robot.)

In general, one has to consider the dehumanizing effects of people's becoming decreasingly dependent upon human contact for satisfaction of their needs. Many who today can do their jobs only by going to a particular place of work, might tomorrow be able to stay at home and communicate with their clients and co-workers via the home terminal. The socially isolating influence of television is as nothing to the alienation and loneliness that might result from over-enthusiastic reliance on the home terminal and associated gadgetry.

Let us now ask whether social applications of artificial intelligence might have any welcome effects on the way people view themselves and other people.

Computational models of intelligence are in fact markedly more human than the behaviourist models of mankind that have been widely accepted for years, because the computational approach can endorse the humanist's stress on the idiosyncrasy of people's subjective world-views and on the directive role of the self-image. A program constructs its interpretation of the input by way of its particular epistemological scheme, or set of inner models of the world, and the same input may thus be 'experienced' very differently by different programs. If this commonly unsuspected 'humanizing' feature of artificial intelligence can be brought home to the general public, then many of the ill effects I have hypothesized will be allayed.

Educational methods based on the pedagogical philosophy of LOGO-turtles[19] might change ways of thinking about 'failure.' Instead of the passively defeatist 'I'm not good at this,' the child would say 'How can I make myself better at it?' This attitude is encouraged by the computational way of thinking about thinking, with its emphasis on the creative interrelation of many different procedures, and on the unintended effects of specifiable bugs in basically well-conceived attempts to achieve one's goal. By contrast, constructive self-criticism is not encouraged by a conception of intelligence that views it as the product of a number of mysterious monolithic 'talents' or abilities, which one either has or lacks, willy-nilly. (An increase in the use of programming in schools may also help to avert the growth of a socially divi-

sive 'computer elite,' a small group of people whose members are the only ones to understand computation and so the only ones not to feel alienated in the computerized society of the future.)

Publicly available programs are unlikely to be rendered 'emotional,' even assuming this to be in principle possible, since there would be little point in doing so. If people still felt a need to draw a line between themselves and 'machines,' their valuation of the emotional life would probably increase, with corresponding effects on cultural mores. Incorporated in the self-image of most Westerners is the Protestant Ethic that only hard work is a really serious activity. And 'work' is implicitly defined as paid, as done in one's employer's time rather than in one's own 'leisure' time, and at the employer's behest rather than for one's own purposes. Consequently, massive unemployment could be more soul-destroying than the most repetitive of factory jobs, causing men destructively to see themselves as social parasites. This is less likely to happen to women, who enjoy an internalized acceptance of emotional values that enables them to derive greater fulfillment from personal relations and the expression of emotions outside the immediate family circle. If automation increases men's opportunities for human interaction, with friends as well as family, we may expect radical changes in the social definitions of sexual roles (for emotionality at present is seen as secondary to the masculine role). These changes would be due primarily to economic shifts in working hours and the sexual division of labor, but they could be reinforced by a general increase in evaluation of the emotional life deriving from the 'emotionless' nature of programs.

### 3. Precautionary Measures in Writing and Presenting Programs

Sometimes things will go wrong, so that a program needs to be adjusted. Programmers who wish to know what is going on, and what needs to be stopped should anything go amiss, must take steps beforehand to allow for this. Their programs should be intelligible and explicit, so that 'what is going on' is not buried in the code or implicitly embodied in procedures whose aim and effect is obscure.

Programs should be generously commented, so that what a procedure is supposed to be doing (and why) is readily visible. The importance of this for aiding debugging (whether by a human or an automatic programmer) is evident from the program-writing program, HACKER[20]. HACKER writes programs to perform tasks such as stacking bricks in specified ways, and is able to correct its own mistakes - and to avoid similar mistakes in future tasks of a generally similar nature - by way of its understanding of the purposive structure of task and program like. HACKER uses the intention-coding comments it attaches to each line of its 'first draft' programs in amending these programs later. Only because it has such a good idea of what it is trying to do, and how it is trying to do it, is it able self-critically to modify its own procedures.

Intelligibility and explicitness are to some extent opposed, and programs get less readable as

they approach the machine code level. What counts as 'machine code' is likely to become rather more intelligible to human beings, since instructions that now have to be programmed may be 'hardwired' into the electronics of the machine.

To economize on computation time, the next best thing to hardwiring is compiling. It would in general be advantageous to have the possibility of switching from compiled to interpreted mode if necessary, so as intelligently to guide the giving of detailed instructions to the machine in light of current circumstances, and the programmer should specify the sorts of contexts in which this switch might be advisable. The learning program HACKER, when in doubt (and when running a new program for the first time), can switch into a slower 'careful' mode in which every step is carefully examined before it is taken.

It is easier to see what is going on if sub-routines are written so that they can easily be 'got out' from the program as a whole. This principle of modular programming is exemplified in programs like BUILD. The understanding and improvement of BUILD are facilitated by the clear distinction between the intercommunicating 'expert' sub-routines. Modular programming will be required also in writing programs that cannot possibly be spied on or maliciously altered. For unless it is small, a program cannot be proved to perform exactly the functions required by the designer and no other functions whatever. The only way to make unauthorized access absolutely unfeasible would be to build the system around individually proved modules, or 'security kernels.' Modular programming also helps to counter the conservatism inherent in widespread applications of very complex systems, for if faults can be adjusted without necessitating widespread tinkering with the system then programmers will be better able to face the task of improving a large-scale system.

If programmers are to be able to see what is going on, they should not use programming techniques which - while making for readable programs - render the control structure obscure. For example, use of the programming language PLANNER tends to produce readable programs in which one cannot tell what is going on by examining the code[21]. PLANNER programs are legible because it is a 'goal-directed' language in which routines can be indirectly invoked by way of general 'goal-patterns' matching the index of the routines in question. Specific advice to try one routine before another can be included by the programmer. But the language embodies an automatic backtrack facility which tags on to the end of the specific advice (if any) the instruction to USE ANYTHING that might work: that is, any routine whose index-pattern matches the goal currently being sought. This strategy might lead to some nasty situations: a distraught parent in squalid surroundings may try anything to stop the baby screaming - and succeed appallingly well. A human has to be distraught to disregard the side-effects of 'effective' measures, but a program may not even know about them.

Moreover, in the PLANNER automatic backtrack situation, the program simply tries out the first method that seems apt, since it has no way of comparing all possibilities beforehand. The knowledge

that hitting may hurt the baby could be buried inside the relevant procedure, to be found only when this way of making the baby quiet was run. Only if the programmer had specifically included the advice never to hit the baby, would a PLANNER program refuse even to consider it. The programming language CONNIVER, (which was developed largely in response to the difficulties involved in PLANNER backtracking), by contrast does allow for the potential choices to be listed for higher-level criticism[22]. (BUILD uses this facility in deciding on the safest way to build a brick-house.) If a CONNIVER program knew that hitting hurts, and was able to access this knowledge when needed, it would be able to control itself long enough to find an alternative way of quieting the baby, or to decide to abandon this goal as not legally achievable.

This example makes it clear that a flexibly intelligent control structure is only useful if the program embodies sensible criteria of what effects are 'undesirable'. Some 'sensible' criteria are culture-specific, and one may expect much moral-political disagreement about what precautions regarding artificial intelligence are worthwhile. Isaac Asimov's well-known 'Three Laws of Robotics' each assume that we know (and have communicated to the robot's program) what is to be counted as 'harm'.

Should we, for instance, regard it as harmful to lie? (We commonly say it is immoral, and utilitarians argue that it is socially harmful.) But if so, we may be landed with a teletyping program that is forbidden to tell tactful white lies. One of the reasons people often have for lying is precisely to avoid harming others. The more 'personal' the program, the more likely that lying (or a tell-tale silence?) might be in order. Some people would even claim that for social-political reasons--such as preventing dissent, disorder, or panic--certain 'impersonal facts' should be kept from the general public. On this view, should all home terminal programs (particularly those notionally contributing to participatory government) be kept ignorant of them also, or should they be able intelligently to take account of them while 'protectively' keeping them to themselves, if necessary lying in order to do so?

The ethical ambiguity of lying is only one of many difficult cases. One may therefore experience some reserve about recommendations of a new profession of value-impact forecasters: ethical experts armed with scientific tools for making cost-benefit judgments[23]. Roszak has sourly commented [24]: 'What these ethical engineers will know of "value" (Old Style: the meaning of life) may of course be only a computer simulation of a statistical illusion gleaned from questionnaires whose unreality crudely approximates a moral imbecile's conception of an ethical decision.'

Rozzak regards the use of psychological terms in psychological contexts, and artificial intelligence research in general, as inescapably dehumanizing for the culture that admits them. He would urge programmers not merely to refuse to write obviously exploitative programs, but to re-examine the philosophical assumptions of their whole enterprise. In view of the discussion in Section 2,

this surely is not too much to ask. Psychological theories in general are not purely descriptive, but are largely constitutive of social reality, and computational theories are no exception. If the public believes--rightly or wrongly--that science regards people as 'nothing but clockwork', then clockwork-people we may tend to become. This is why computer scientists should stress the basic philosophical compatibility of 'humanist' and computational views.

In this connection, perhaps programs for public use should include explicit reminders of some of the differences between computers and people. This is done (for utilitarian reasons) in the peptic ulcer program[18] previously described: it continually reminds patients that they are on-line to a machine. The reminder helps to avoid mystification of the patients, stressing that they are merely 'filling in a form' by teletype rather than engaging in flexible communication, still less participating in a human relationship. 'Plausibility-tricks' are sometimes included in language-using programs. For instance, the CAI programs currently used in schools take care to address the children from time to time by their first names, so as to put them at their ease. In view of the dehumanizing potential of computer applications (quite apart from the possible results of an over-generous misunderstanding by the person), it may be that the limits of individual programs should be made as clear as possible, to users, and plausibility tricks used sparingly--if at all.

On one point, however, even Roszak presumably would agree that the similarity between people and programs should be deliberately stressed. One of his objections to the social use of computers in an advisory capacity is that the machine may be seen by the public (including the politicians) as purely objective, and therefore not to be argued with. Accordingly, the inescapable subjectivity of a program's judgments should be made clear to its users. We saw that this is done implicitly in the MYCIN medical consultant: MYCIN not only offers alternative judgments when it perceives several possibilities of diagnosis or treatment, but gives its reasons for each so that physicians can rationally reject its advice if they see fit. In general, it would be worthwhile somehow to remind the user that programs function within their own subjective cognitive worlds, just as we do. Any differences between the two are matter for epistemological debate, not for servile capitulation on the part of the person.

#### 4. Conclusions

Increasingly intelligent machines are likely to become socially available, whether for use by the general public or by political and administrative institutions.

Many of the potential ill-effects depend on the common (though mistaken) view that mechanism is incompatible with 'human' qualities. If the public assume that science offers an image of man that is irreconcilable with humanism, they must either deny their humanity--with socially destructive results--or else forfeit a scientific understanding of mankind.

So in addition to writing programs in a sensible fashion (so that the control structure is perspicuous and alteration is facilitated), professionals involved in artificial intelligence should take pains to see that this choice is not regarded by the public as unavoidable. For example, they should avoid 'plausibility tricks' in programs that might lead to unnecessary mystification of the users, and they should point out that a program's data and inference-procedures may always be questioned in principle, just as a person's can be.

Recognition of these issues by the profession considered as a social institution (as opposed to isolated individuals within it) will be necessary if they are to be adequately faced. Insofar as the profession succeeds in reassuring the public that mechanism (of sufficient complexity) is in principle capable of generating the distinctively human characteristics of subjectivity, purpose, and choice, the increasing social use of intelligent machines will present less of a threat to humane conceptions of self and society.

#### References and Bibliography

1. The 'chess champion' forecast was made and withdrawn by H.A. Simon, one of the programmers of the General Problem Solver (GPS). For a discussion of the organization of large amounts of knowledge, see, e.g., Minsky, M.L., 'A framework for representing knowledge', in 'The Psychology of Computer Vision', ed. Winston, P.H., pp. 221-77 (McGraw-Hill, New York, 1975).
2. Firschein, O., Fischler, M.A., Coles, L.S. and Tenenbaum, J.M., 'Forecasting and assessing the impact of artificial intelligence on society', Third Int. Joint Conf. on Artificial Intelligence, pp. 105-20, 1973.
3. Shortliffe, E.H., Axline, S.G., Buchanan, B.G., Merigan, T.C. and Cohen, N.S., 'An artificial intelligence program to advise physicians regarding anti-microbial therapy', Computers and Biomedical Research, 6, pp. 544-60, 1973; Shortliffe, E.H., Davis, R., Axline, S.G., Buchanan, B.G., Green, C.C. and Cohen, N.S., 'Computer-based consultations in clinical therapeutics: explanation and rule acquisition capabilities of the MYCIN System', Computers and Biomedical Research, 8, pp. 303-20, 1975.
4. Shortliffe, E.H. and Buchanan, B.G., 'A model of inexact reasoning in medicine', Mathematical Biosciences, 23, pp. 351-79, 1975.
5. Hart, P.E., 'Progress on a computer based consultant', Fourth Int. Joint Conf. on Artificial Intelligence, pp. 831-41, 1975.
6. For current attempts to base CAI on the flexible representation of knowledge, see: Collins, A. 'Process in acquiring knowledge', in Anderson, R.C., Spiro, R.J., and Montague, W. E., eds., 'Schooling and the Acquisition of Knowledge' (Erlbaum, Hillsdale, N.J., in press); Seely-Brown, J. and Burton, R.R., 'Multiple representations of knowledge for

- tutorial reasoning', in Bobrow, D.G. and Collins, A., eds., 'Representation and Understanding: Studies in Cognitive Science', pp. 311-50 (Academic Press, New York, 1975).
7. McCarthy, J., 'The home information terminal', 'Man and Computer', Proc. Int. Conf., Bordeaux 1970, pp. 48-57 (Karger, Basel, 1972).
  8. Underwood, M.J., "Machines that understand speech," The Radio and Electronic Engineer, 47, No. 8/9, pp. 368-376, August/September 1977.
  9. Reddy, D.R., Erman, L.D., Fennell, R.D. and Neely, R.B., "The HEARSAY speech understanding system," Proc. Third Int. Joint Conf. on Artificial Intelligence, pp. 185-93, 1973; Reddy, D.R. and Newell, A., "Knowledge and its representation in a speech understanding system," in "Knowledge and Cognition" ed. Gregg, L.W., pp. 253-286 (Erlbaum, Baltimore, 1974).
  10. Tenenbaum, J.M. and Weyl, S., "A region-analysis subsystem for interactive scene analysis," Fourth Int. Joint Conf. on Artificial Intelligence, pp. 682-7, 1975.
  11. Fahlman, S.E., "A planning system for robot construction tasks," Artificial Intelligence, 5, pp. 1-50, 1974.
  12. Winston, P.H., "The MIT robot," in "Machine Intelligence 7," eds. Meltzer, B. and Michie, D., pp. 465-80 (Edinburgh University Press, 1972).
  13. Roszak, T., "Where the Wasteland Ends: Politics and Transcendence in Post Industrial Society" (Doubleday, New York, 1972); Marcuse, H., "One Dimensional Man: The Ideology of Industrial Society" (Routledge, Kegan Paul, London, 1964).
  14. May, R., ed., "Existential Psychology," p. 20 (Random House, New York, 1961).
  15. The technical aspects of "tele-voting" were discussed in 1959 by Dr. Vladimir K. Zworykin in the Institution's 4th Clerk Maxwell Memorial Lecture: "The Human Aspect of Engineering Progress," J. Brit. Instn. Radio Engrs, 19, No. 9, pp. 529-44 (see pp. 541-4) September 1959.
  16. Weizenbaum, J., "Contextual understanding by computers," Comm. Ass. Computing Machinery, 10, pp. 474-80, 1967.
  17. Weizenbaum, J., "Computer Power and Human Reason: From Judgment to Calculation" (Freeman, San Francisco, 1976).
  18. Card, W.I., Nicholson, M., Crean, G.P., Watkinson, G., Evans, C.R., Wilson, J. and Russell, D., "A comparison of doctor and computer interrogation of patients," Int. J. Bio-medical Computing, 5, pp. 175-87, 1974.
  19. Papert, S., "Teaching children to be mathematicians versus teaching about mathematics," Int. J. Educ. Sci. Technol., 3, pp. 249-62, 1972.
  20. Sussman, G.J., "A Computer Model of Skill Acquisition" (American Elsevier, New York, 1975).
  21. Hewitt, C., "PLANNER: a language for proving theorems in robots," First Int. Joint Conf. on Artificial Intelligence, pp. 295-301, 1969; Hewitt, C., "Procedural embedding of knowledge in PLANNER," Second Int. Joint Conf. on Artificial Intelligence, pp. 167-84, 1971.
  22. Sussman, G.J. and McDermott, D.V., "Why Con-niving is Better than Planning," AI Memo 255a, MIT AI Lab., 1972.
  23. Toffler, A., "Future Shock" (Random House, New York, 1970).
  24. Roszak, T., loc. cit., p. 69.