ARTIFICIAL INTELLIGENCE AND SOCIAL FORECASTING

Margaret A.Boden
University of Sussex

Cognitive Science Research Paper

Serial No. CSRP. 018

The University of Sussex,
Cognitive Studies Programme,
School of Social Sciencw,
Falmer, Brighton BN1 9QN

ARTIFICIAL INTELLIGENCE AND SOCIAL FORECASTING

Margaret A, Boden

University of Sussex, England

Abstract

Artificial intelligence (AI) is the study of how to write programs enabling computers to do things that would require intelligence if done by people, and it could engage with social forecasting in two ways. First, it is part of the overall social-technological context within which forecasters work. Commercial Al-programs will affect markets and life-styles; and advice-giving "expert" systems will raise novel legal, social, and psychological problems. Second, Al-programs might be used for making the social forecasts. Unlike the (essentially quantitative) computer models used for this purpose today, they could reason (and explain the selves) in verbal form. Writing an expert system requires clarification of the theories, assumptions, and "rule-of-thumb" inferences concerned. It would be easier to identify the inherent moral-political bias than it is in models comprising sets of differential equations.

ARTIFICIAL INTELLIGENCE AND SOCIAL FORECASTING*

Margaret A. Boden

University of Sussex, England

I: Introduction:

Once upon a time, a father and his son crashed their car into a lorry. The father was killed instantly. The son, gravely injured, was rushed to hospital. The emergency surgeon was called, and scrubbed up in preparation. But when the child was placed on the operating table, the surgeon exclaimed: "I can't operate on this boy: he's my son!" -- How can this be?

If you have solved that conundrum, you may like to ponder this one**: On two neighbouring soybean-plantations, the plants show symptoms of disease five years in succession (different every year). One plantation owner is rich, and can afford to go to the world-expert on soybean-disease and ask for his advice. The other owner cannot pay for the world-expert's opinion, and has to seek advice elsewhere. Yet each

---

* -- This paper was written as the keynote address to the interdisciplinary Canadian Workshop on Artificial Intelligence and Society, organized in 1982 by Prof. E. Zureik, Dept. of Sociology, Queen's University, Kingston, Ontario.

** -- The second puzzle is based on an idea in an article on expert systems by Paul Ellis, of Bradford University Management Centre (Forecasting and Policy Alternatives Group).

ear, the poor owner's plants respond to treatment whereas the rich wner's plants sicken and die. -- How can this be?

The first of these puzzles foxes most people, who assume that the urgeon must be a man (not a woman). Were social forecasters similarly o forget that women, too, can be surgeons, their forecasts would be ighly misleading. Employment patterns, educational policies, family tructures, and many other socially important matters have been nfluenced by the entry of women into professional occupations.

The second puzzle defeats virtually everyone. Short of postulating n answer to prayer, there seems to be no way of making sense of it. One ssumes that the advice must have been given by a person, whether human r divine, failing to allow that it may have come from a machine. owever, a computer program already exists which can give more reliable iagnoses of soybean diseases than any person, although this particular uzzle-story is a fiction [NOTE 1]. One could compose a similar puzzle ased on actual dialogue between a human doctor and an existing "expert" edical program. For example, most people would assume the two speakers" were human if they were to see a note of a fragment of onversation like this [Feigenbaum (1979)]:

HUMAN USER: "Why didn't you consider streptococcus as a possibility for ORGANISM-1 [previously diagnosed by the program as pseudomonas-aeruginosa]?

PROGRAM: "The following rule could have been used to determine that the identity of ORGANISM-1 was streptococcus: RULE 33.

But clause 2 ('the morphology of the organism is coccus') was already known to be false for ORGANISM-1, As this snippet of interaction suggests, certain sorts of medical diagnosis can be carried out by a program, which can even enter into some kind of explanatory dialogue if a human doctor questions its advice. Indeed, the program quoted above almost always gives better diagnoses (and drug-prescriptions) than the nonspecialist human doctor, though it gives the same advice as the topflight consultant-physician in

only (only?) 90% of cases.

The development of apparently intelligent programs like this one cannot fail to influence society, and ignoring them would produce no less misleading forecasts than would ignoring changes in the occupational status of women. Social forecasting has been dubbed "the art of anticipation," a label expressing insight as well as alliteration. As an intellectual activity, social forecasting is more art than science, theories in social science being both incomplete and imbued with moral-political values. As a potential influence on policy-making, it implies active anticipation rather than passive expectancy: it is meant to be used to help foster or avert various possible outcomes of social action.

Artificial intelligence (AI) might engage with this anticipatory art in two ways. First, and inevitably, AI will constitute one dimension of the overall cultural and technological context within which forecasters work. It is no longer confined to academia, or scornfully dismissable as "one of the many flights of fancy that Californians are prone to". Many of the best AI-researchers are now employed by commercial laboratories, and over half a dozen companies have been formed in the past three years by university-based AI specialists intent and probable effects of nuclear energy or genetic engineering, so they will increasingly have to consider the social uses and effects of AI-technology (including the changes it may make in our ways of thinking about people). Second, and more problematically, AI-programs might come to be used to help in the making of -- or even to make -- the social forecasts themselves.

II: AI as Part of Society:

Some sociopolitical dimensions of computer technology have long been recognized, including such important issues as mass unemployment, privacy, family structure and sex-roles, leisure, centralization of political power, and the military misuse of technology. These problems are significantly compounded by advances in telecommunications, which make the concept of the "global village" more like prediction than fantasy (though it can exist only to the extent that a common culture is shared: the agents of multinationals already inhabit a global village, but random English and Chinese citizens provided overnight with telecommunications would not). All these socially crucial issues are raised by work in artificial intelligence no less than by other uses of computers, so general discussions of computerization and telecommunications have a bearing on the effects of AI [Forester (1980); Pelton (1981)]. But our special interest here is in the social implications of intelligent machines in particular, and these include a number of issues not usually raised in discussions of computers and/or microprocessors in general.

AI programs differ significantly from most computer systems their performance and the social uses to which they can be put, but also their relevance to the ways in which we view ourselves and social relations in general. Most current applications of computer science exploit the "brute force" of the computer, its ability to carry out large numbers of computations in a short time. Moreover, these computations are rigidly predictable, and usually very simple in type. But AI programs enable computers to do things that would require intelligence if done by people, and their computations are less predictable and more complex accordingly CBoden (1977)D.

AI programs involve a significant degree of computational flexibility. Rather than merely doing lengthy sums, or rapidly accessing pre-tabulated data, they generate reasoned judgments or sensible guesses. These often involve their going beyond the information explicitly mentioned in the input-problem concerned — for it is a mark of intelligent thinking that it can take things for granted, and/or make sensible guesses, when some of the potentially relevant information is missing.

Typically, AI systems can draw inferences of various sorts, according to circumstance — many of which will be non-deductive inferences. The information they work with is sometimes mathematical notation, but more often it is verbal text, visual input, auditory speech or music, or tactile data. Often, they can make use of information whose truth is known to be uncertain, and attach qualifications to their current (and sometimes their previous) judgments. They can construct and search over a large "space" of potential problem-solutions. They can match a specific (and perhaps unfamiLiar) incut with some aeneraL Dattern- so as to fit it into an by the acquisition of new data or concepts, or even by the self-modification of their own rules for reasoning. And some can explain their reasoning to human users at whatever level of detail is most appropriate to the person concerned.

Very few current uses of the micro-chip involve AI techniques (some widely marketed chess-programs are an exception). This is due to the relatively undeveloped state of AI rather than to any longterm technological impossibility. AI programs are usually run on large computers, and only a few are already in public use. These few are restricted to very limited domains, but within those domains they do a useful job. For example, some of the advances in AI-vision are now on the market for use in robotics. Although these programs do not enable a robot to see a cat or a king (except possibly as an uninterpretable blur), they may enable it to pick distinct machine-parts out of a bin for assembly -- where these parts do *not* (as with current manufacturing robots) have to be placed in precise positions and orientations on specially-moulded palettes.

Experts engaged in a Delphi-study some years ago forecast a very wide range of commercial applications by the end of the century [Firschein *et al.* (1973)]. One may have reservations about the timescale they envisage (I believe them to have underestimated the difficulties to be overcome [Boden (1977), ch. 15]), but there is no question that varied applications are imminent. AI projects currently being researched include machine-translators; automatic picture-interpreters, speech-interpreters, and newspaper-precis writers; tutorial programs hopefully far superior to traditional "teaching-machines"; programmed planning systems and expert consultants; selfenvironments that make programming easier for novice users; an intelligent robots that can see, hear, touch, and plan as well as move.

Although most AI research-projects are not yet ready for publi use, and many are (as yet) rather less impressive than they may sounc they do have capabilities -- and commercial potential -- that mar people would find surprising. For instance, the newspaper-precis writer which can be attached to the UPI wires and left to run unsupervized, able to pick up only those stories relating to the specific topics knows about -- such as terrorist bombings and kidnappings, earthquakes bank-robberies, car crashes, and so on [DeJong (1979)]. But it gives usable precis of over 80% of these. Giving such a precis ofte involves drawing inferences to facts that were not actually stated the original passage. Thus a newspaper-story about a car crash ma mention the driver's name at the outset, and later on say that a ambulance from the local hospital arrived soon after the incident without explicitly stating that the driver was taken in the ambulance the hospital. Nevertheless, the precis-writing program can tell us tha "Joe Bloggs was taken to hospital." If one were interested in only or type of press-cutting (for example: terrorist bombings, or insuranc frauds), such a program could be useful.

Among the most publicly visible (though not the mos computationally sophisticated) AI programs are a variety of "exper systems" [Michie (1979)]. Their function is to provide knowledgeabl advice to human users struggling with a problem of some kind. They ca be used to suggest solutions to problems, solutions that are (shoul be?) checked by humans for "obvious" errors. Also, they can be used t teach novices about a problem-domain, or to assist human experts b reasoning, and noting trade-offs between various solutions suggested by themselves or by the human expert.

They communicate their advice in (a restricted subset of) natural language, so the user does not need to learn to program. Depending on the problem (how to pay less taxes, or how to mend one's car), they can be thought of either as automatic professional consultants, or as dynamic question-and-answer manuals. They can make use of theoretically justified inferences when these are available, but can use mere hunches when they are not. These systems do not merely offer an answer to a question, point-blank. They can assess the degree of confidence to be attached to the various suggested answers, and they can exhibit their background reasoning on several levels of detail if asked to do so. This makes it easier for the human to come to an informed judgment on whether or not to accept their advice, in cases where there is some reason to doubt it.

Each expert system is, in essence, a collection of IF-THEN rules, which determine what action the program will perform given such-and-such a situation. A "situation" may be defined in terms either of one condition (or its negation) only, or of a conjunction or disjunction of conditions. The information about the problem-at-hand that is fed in by the human user forms a crucial subset of the relevant conditions. So too do the theoretically based inferences, and even the rule-of-thumb hunches, relied on by the human expert who provided the relevant knowledge when the program was written in the first place. The "action" may be (among other things) a diagnostic judgment, a prescription to the human to do something specific, an explanation of some previously-given advice, or a suggestion that the human find the answer to a particular or

Examples include systems expert in: diagnosis and drug-prescription for infectious meningitis, pulmonary disease, and glaucoma; geological prospecting (for minerals and oil); the chemistry of some drugs used in contraceptives; design of electrical circuits and computer systems (including microchips); planning of gene-splicing experiments; and diagnosis of soybean diseases. In addition to developing specific systems, the AI-researchers concerned are producing a "core" program, that could be used as the nucleus of indefinitely many expert systems provided that the relevant expertise is added to it [Buchanan (1982)].

Some expert systems are already in the public domain, and their performance, in certain limited areas, is superior to that of all but the very best human experts. Indeed, the soybean program's 99% success-rate outdoes the world-expert's 83% (the program benefits not only from his knowledge initially written into it but also from self-generated inductive rules) [Chilansky et al. (1976)].

Several medical programs are being used experimentally in research-hospitals, both for making diagnoses and/or prescriptions and for teaching medical students about diagnosis and prescription in the relevant area. An automated specialist in pulmonary diseases is in routine use at the Pacific Medical Center in San Francisco, and 85% of its reports are countersigned by doctors without change. The meningitis program developed at Stanford Medical School [Shortliffe (1976)] gives the same advice as human specialists do in nearly 90% of cases (one would like to know what proportion of the other 10% are obviously "crazy", as opposed to being the sort of judgment -- right or wrong -- that a human medic might make). At least one oil-prospecting company (Schlumberger) is already using an expert geological program written by still handles the really tricky cases. The development of the program was very expensive, but since it costs $15m dollars to drill an oil-well (which may turn out to be dry), oil-companies will pay large sums for reliable advice on where to drill. Schlumberger now employs dozens of researchers in three AI laboratories, at an estimated cost of over $5m a year.

Companies and institutions that cannot afford such privately-funded developments may nevertheless be able to benefit from expert systems, given the governmental interest surfacing in various countries. The Japanese are giving massive financial support to comparable research, as a crucial element of their "Fifth Generation" computer technology (for which the state-funding alone is $45m over the next three years). Even the U.K., making an exception in these recessionary times, has recently promised appreciable amounts of money for research on expert systems. Inevitably, then, many more such systems will be made available over the coming years. Despite their (current) inability to handle the very varied, and seemingly ill-structured, knowledge-bases that human experts can deal with, they will be widely used by government, business, and private individuals. (Future advance in AI can be expected to produce more sophisticated computer-experts, even if these are radically different in kind from current examples.)

Put to commercial use, Al-programs will appreciably affect not only markets, but also personal and communal life-styles. Expert systems, for instance, will raise legal, social, and psychological problems of an unfamiliar kind. Whether they are used to make decisions or merely to provide expert advice to (probably less expert) decision-makers, the status of the human expert [illegible] expertise may be seen ambiguously as helpful or exploitative -- much as human technicians are.

To take just one example: how will patients (and nurses) react to physicians (and nurses) whose advice is based largely on a computer program? Given the way in which most people currently think of computers -- namely, as "brute force" calculators rather than as flexibly intelligent systems -- we may expect a good deal of resistance here. Even patients who realize the potential flexibility of computerized inferences might feel less close to or trustful of their human doctor -- thus undermining the therapeutic effect of a good doctor-patient relationship, or comforting bedside manner. A legal expert or tax-consultant would probably be more generally acceptable, since in these areas a human relationship is not seen as so important.

One might predict that such resistance would be minimal, or very likely soon overcome, on the ground that it is irrational to object to the use of an AI-program if one does not object to a doctor's using a textbook (both, after all, were written by humans in the first place). But people start from different viewpoints, according to only some of which is a particular attitude irrational, and social forecasters should try to take this fact into account. This is why, as I shall argue later, AI in general could be either de-humanizing or re-humanizing, depending on the background philosophical assumptions held by members of society.

A medical expert, whether human or machine, cannot be expected to get it right every time. Intelligence typically involves flexible thinking in the face of uncertainty, and one of its prime characteristics is the ability to deal sensibly with incomplete or with it infallibly (indeed, only an intelligent system *could* make mistakes). So AI programs are not only more powerful and surprising than computerized filing-cabinets or wage-clerks, but are also potentially more fallible. Intelligent seeing-machines, for instance, would be prone to visual illusions much as we are [Gregory (1967)], and similar remarks apply to intelligent programs expert in medical problems. Only a purely deductive program, which never had to guess, or make reasonable assumptions that could in principle be false, would be immune to error. (It might of course be incompetently programmed, so as to contain false information, unreliable inference-rules, and/or "bugs" making it do things it was not meant to do; but that is a different matter.)

Given the possibility of these various sorts of malfunction (not to mention hardware faults), physicians and hospitals who use an expert program in diagnosis and treatment risk legal complications of a new kind -- especially if they live in a country where medical litigation is a popular sport. Who can be sued, and for what? If the doctor wrote his own private program, doubtless he would have to pay damages in the event of mishap. But the programmer can hardly be held responsible in general, since AI programs are often the work of an ever-changing team, many of whose most important members may already be dead. Suppose that the Distillers Company had not only marketed thalidomide, but had distributed (worldwide?) a programmed medical expert which advised using this drug for sleeplessness during pregnancy, perhaps with a 90% confidence-measure? Would -- or should -- they be more heavily liable than they already were on account of their over-confident promotional literature?

Clearly the legal implications of computer-based professional very basic social perceptions. A recent article in <u>Futures</u> attempts to lay the groundwork for a legal definition of AI and "robot criminality" [Lehman-Wilzig (1981)]. The author starts from the precedents concerning legal responsibility for manufactured products, wild animals, domestic pets, slaves, the mentally disabled, infants, agents and servants, and autonomous persons. These laws contain a number of pitfalls threatening those who rely on AI advice. For instance, a master is legally at the mercy of his servant, since "the master is jointly and severally liable for any tort committed by his servant while acting in the course of his employment . . . based, not on the fiction that he had impliedly commanded his servant to do what he did, but on the safer and simpler ground that it was done in the scope or course of his employment or authority" [<u>ibid</u>.]. This covers the situation where a human expert (such as a chef) carries out actions which the non-expert master could not have commanded in detail. General practitioners hoping to use programmed professional expertise:-- beware!

One authority on jurisprudence has even gone so far as to maintain that "the master is held liable for any intentional tort committed by the servant where its purpose, however misguided, is wholly or in part to further the master's business", being prepared to except from liability only those actions committed by the servant during "frolic and detour" [<u>ibid</u>.]. Whether computer-systems can truly be said to have intentions, the capacity to engage in frolic, or even rights [Willick, 1983] may thus be questions of more than merely academic interest.

This reference to philosophical puzzles bearing on highly practical matters reminds us that an important social aspect of AI is its effect on the way people think about themselves, and about society. Outside the novels, such as Burgess's _A Clockwork Orange_ or Vonnegut's _Player Piano_, Futurology and science fiction have much in common, and forecasters should consider the subtle changes in social attitudes that various technologies may encourage. Political commitments and patterns of interpersonal relations can be significantly affected by such attitudes. Forecasters often consider the effects of changes in people's representations of themselves and society, as for instance when they assess the spread and influence of religions, or Marxist ideas. Whether AI is more likely to be dehumanizing or, instead, positively _rehumanizing_ is not yet clear. A plausible scenario can be sketched for either outcome.

Most people assume that there is an absolute distinction between men and machines, so that mechanism in general and computer-based psychological analogies in particular are essentially incompatible with concepts of purpose, morality, and freedom. To the extent that people continue to believe this, the increased public visibility of apparently intelligent programs might subtly undermine their view of themselves as morally responsible beings. Even before the development of AI, the psychotherapist Rollo May complained of the insidious influence of behaviourism (and the natural scientific tradition in general), in "sapping the willing and decision" of his patients and weakening their sense of moral autonomy [May (1961)].

This dehumanizing effect would be even greater if programs were used in contexts previously thought of as human _par excellence_, and it is not surprising that the well-intentioned suggestion that computer programs be introduced into psychiatric diagnosis and therapy has been bitterly denounced as "obscene" [Weizenbaum (1976)]. Again, the counsellor will send a frisson of horror through many people's bones. Certainly, a program's questions might help achieve a marriage-counsellor's primary aim, which is to enable spouses to make their problems and perceptions explicit, so aiding calmer reflection (without the program's being able to offer any advice, whether good or bad). But for many people, the knowledge that it was a program that had asked the questions might devalue the personal insights gained in the process, and a general sense of alienation might be encouraged by the decision to use the non-human system in the first place.

On the other hand, as I have argued elsewhere [Boden (1977)], the psychological implications of AI can properly be interpreted as humane rather than dehumanizing. For this approach allows for the real influence of mental representations (variously functioning as concepts, beliefs, preferences, desires, ideals . . .), and thus provides a welcome rebuttal of the anti-mentalist bias of behaviourism. AI-ideas have already influenced theoretical psychology, encouraging a view of thinking as a constructive, self-reflective activity.

Applied to educational psychology, for example, the computational account of intelligence has inspired educational practices wherein the child's passively defeatist "I'm no good at this" gives way to the more optimistically constructive "How can I make myself better at it?" [Papert (1980); Boden (1982)]. Applied in small pilot-studies. these educational ideas have apparently produced significant improvements in the self-confidence and academic achievement not only of normal children, but also of children needing remedial maths-teaching and grossly handicapped (palsied or autistic) children. They are attracting interest in institutions such as the Inner London Education Authority, they live up to their promise, they could help foster changes in educational philosophy that could have widespread social effects.

III: AI as Applied Social Science (Computer Forecasting):

We have considered some dimensions of AI-induced social change that forecasters should take into account. But what of the use of AI, in the near or far future, to make the social forecasts themselves? Could there be an expert system for social forecasting, an AI-program to rival Meadows or Kahn? And, even if it is in principle feasible, is such a system desirable?

Computer models of society have, of course, already been used by forecasters. But these are not like the medical or geological programs mentioned above, which can identify the grounds and possible weak points of their diagnoses, and express their reasons in a subset of natural language. On the contrary, current social-forecasting models cannot express semantic information,nor make conceptual inferences in anything like the way we normally do ourselves. Today's forecasting programs are essentially "number-crunchers". They deal with quantitative rather than qualitative information, being defined in terms of mathematical parameters and differential equations. Consequently, they suffer from several serious disadvantages.

The use of numbers, and of broadly arithmetical computations over them, tends to encourage a spurious precision in forecasting. What, for instance, is one to make of Kahn's forecast that the 22nd century will see an annual growth rate of "2.3 percent _or so_"? [Kahn _et al._ (1976)] Such pseudo-precision is culturally oppressive, even though most social

For many policy-makers and members of the public (to whom Kahn's book was addressed) do not: they are likely to credit futurology with the judgmental authority of arithmetic.

This pseudo-precision is harmful also because it is commonly coupled with a reluctance to consider uncertainties about the future. Since some of these could be influenced by conscious political choice, the anticipatory, policy-engendering aspect of forecasting is thereby dangerously flawed [Freeman & Jahoda (1978), ch. 11]. Even a forecast that two centuries hence we shall see a growth-rate of 2.3% GNP tacitly assumes that our current concept of (accumulative) GNP is the best measure of national welfare, and will still be used to define "progress" in the future. But some forecasters have suggested that measures of growth should include other (distributive) aspects, so as to reflect the extent to which members of a society share in employment, consumption, or life-expectancy [Kuznets (1972)]. Optimization policies based on growth-rates so defined would lead to very different decisons on resource allocation and the like.

This point highlights a grave disadvantage of number-crunching models. Any relevant information in the mind of the social-scientist modeller must be expressed numerically if it is to be included in the model at all. Inevitably, this inhibits consideration of any dimension that cannot be numerically measured. Among such dimensions are social-political values, whether positive or negative. If these cannot be quantified they are unrepresented in models demanding numerical precision. Even if they are somehow quantified, so as to enter into the model, their representation as numerical parameters hides them from the view of all but the methodological sophisticate.

This is because the reasoning — or rather, the calculation — behind the conclusions generated by these models is commonly opaque to all but the mathematical initiate. This makes it very difficult for politicians and the general public to produce sensible critiques of the inferences and/or assumptions involved, or even to understand such critiques when produced by other social scientists. Computer-based models, of one sort or another, will inevitably be used in forecasting: fifteen years ago, Kahn was already claiming that "decisions are becoming necessary that are too large, complex, important, uncertain or comprehensive to be left to mere mortals whether private or public[11] [Kahn & Wiener (1967)3. It is crucial, therefore, that "mere mortals[11] should be helped to understand what these non-human representational systems are doing.

Al-based programs, by contrast, can deal not only with quantitative data but also with semantic (non-numerical) information, expressed in some approximation of natural language. As we have seen, expert systems are capable of giving reasoned judgments, and of explaining their reasoning so as to invite challenge, in verbal form. This makes them, in principle, less opaque to the nonspecialist user than are purely mathematical models (though this is not to say that every example is written so as to minimize user-opacity).

Moreover, programming an expert system requires clarification of the theories, assumptions, and "rule-of-thumb" inferences used by the human experts concerned. The system concerned with organic chemistry, for instance, forced specialist chemists to give an explicit theoretical statement of various chemical intuitions — such as that amines are

reasoned (though

coherent theoretical base, as medical systems in general are, at least enable implicit assumptions or intuitive hunches to be made explicit, and so laid open to theoretical or empirical investigation. In either case, the result is a welcome clarification of the reasoning employed by the human expert.

Such clarification is, at least in principle, available to the social scientist and the forecaster no less than to the chemist, geologist, or medic. Forecasters vary in the extent to which they rely on theoretical inferences based in economics, sociology, or social psychology. Some virtually ignore these theoretical disciplines, relying primarily on empirical correlations between observables (such as demographic changes, city-size, or investment-rates). However, the social sciences are notoriously ill-defined and controversial, and societies themselves are awesomely complex. I suspect that the issues involved in social forecasting are too varied and ill-defined for useful expression in any of the current generation of AI-experts. Rather than trying to extend these computationally limited systems to the social domain, it might be better to await further advance in AI techniques for handling large, conceptually diffuse, knowledge-bases (though as AI programs get more subtle, their intelligibility is likely to decrease). Either way, the attempt to think about forecasting in AI-terms could hardly fail to achieve some useful clarification.

Theoretical clarification, in the social sciences, can also lead to greater evaluative perspicacity. As has often been pointed out, but is too often forgotten, theories in social science typically involve hidden ideological commitments and tacit moral-political evaluations. Social forecasting in particular engages with social evaluation, since it the various alternative scenarios. But social forecasters do not always make these evaluative dimensions as clear as they might, so that social-political preferences masquerade as objective scientific fact and thus avoid challenge [Freeman & Jahoda (1978)]. Of course, an "expert system" for forecasting could be written with no explicit evaluative component. But, because of the verbal form of the inference-rules, it would be easier for the human author to express moral-political bias in such a forecasting system than in a model comprising a set of differential equations.

This would be useful because bias that is expressed can be challenged, discussed, and — if we so decide — accepted. It is for us to assess the relative merits of alternative policies, to use anticipation in our interests of acceptance or avoidance. What we must certainly not do (*pace* Kahn) is allow non-human systems to make our moral-political decisions for us.

Even if AI-programs are not used to do the forecasting, they will very probably be used to infer the "facts" (the sensible guesses) on the basis of which forecasts are made. So forecasters will need some form of training that engenders a reasonable degree of computer literacy. Indeed, other members of society — not least the politicians — will share this need; managers, for example, will need to be appropriately educated if they are to use AI tools in their professional work. Notoriously, criticism of computer-based social forecasts (such as *The Limits to Growth*) has been inhibited by misguided trust in the "objectivity" of computers. This shows the importance of educating people to realize that computer-programs are *not* objective or infallible — *but* representational systems subject to the same sorts of

(1981)]). If individuals are not to be mystified and oppressed by the largely AI-based computer culture that lies ahead, widespread computer literacy is essential. This is not primarily a matter of people's being able to program, but rather of their being aware of the sorts of potential and limitations that Al-systems have.

Specific steps might deliberately be taken by programmers when writing programs for public use, to help naive users to avoid certain sorts of common misconception about these matters Eibid.3» And people using a particular Al-program in commercial, political, or administrative institutions should be helped to recognize the limits of the system concerned. Otherwise, they may put more trust in its pronouncements than is warranted. (Non-AI programs are less worrying, since they are more obviously limited, or "stupid[11].) Any program — however initially impressive — has blindnesses and weak spots, not all of which are shared by every other. These particular bounds on the program's computational powers should be taken into account when it is used. For instance, what specific sorts of weakness should constrain our confidence in a given program for medical diagnosis? And is each other diagnostic program subject to identical constraints? If cigarette-manufacturers can be forced to print a Government Health Warning on every packet, surely retailers (and writers) of complex programs should be encouraged to make evident the bounds of sense within which any given program functions? If this is done, then forecasters who rely on AI will be more likely to have a sense of what the relevant systems can and cannot do. (Self-modifying programs raise difficulties here, since they can extend their bounds of sense.)

However, the extent to which warninas can be "written in[11] to projects of various sorts will be essential too. Some current educational developments suggest ways in which the requisite kind of computer literacy might best be fostered. For instance, my colleagues in the Cognitive Studies Programme at the University of Sussex have developed teaching-materials in the form of interactive programs, for use by nonspecialist, arts-based, first-year undergraduates. These enable our students, on the very first day of their programming experience, to realize two crucial points. First, that an apparently intelligent program will have unsuspected limitations; and second, that some of these may be overcome by amateur programming, whereas others are more recalcitrant. A realization of these points is a protection against regarding the AI program as an objective, godlike, unchallengeable -- and unalterable -- system. In addition, we try to counter fears of dehumanization and loss of autonomy by pointing out to students that computational ideas in psychology do not denigrate the mind, but instead allow for a recognition of its power and richness.

On the national level, projects such as the BBC's nascent microcomputer system (which has already been vastly oversubscribed, relative to expectations) will bring computer power into the living-room, and may help increase general awareness of the potential of AI. The mere provision of a computer for a family, or a school, will not necessarily do this, since AI is different from much computer technology. It requires special sorts of programming languages, and these need to be presented in such a way as to enable beginners to do interesting things. The project at Sussex mentioned above has addressed this problem of how to develop user-friendly computing environments (which will help change people's attitudes to computers and

The Swedish government has recently sponsored a public-information "mediation exercise" on computer technology. Of the six booklets of argument about different aspects of computerization, one was focussed on the possibility and desirability of AI, and another on the likely effects of computer-use on life-styles and social relations. These have been provided to schools, colleges, and the press, and widely distributed among the public at large. The aim is to provide a basis for open and informed public debate about this technology, and its possible effects in the future.

Social forecasters trying to take account of changing public attitudes and political priorities might wish to monitor the reaction to information-exercises such as this. For instance, a claim stressed by the "opponent" of AI in the relevant booklet was the unemotionality of computers, including their inability to take sympathetic account of emotions in others. This view inevitably comes up in any discussion of the relation between man and machine, and will very likely receive general assent from the public. Moreover, at least in the short term, "intelligent" programs in public use will not show anything like emotions (though they might be able to respond differentially to some emotional cues in the verbal input of the user). The result may be an increased emphasis on, or valuation of, the emotional life in human beings. Coupled with the changes in work-patterns and leisure-time brought about by computerization in general, this could radically affect the accepted "masculine" sex-role. Whereas at present Western men are discouraged from expressing emotion in all but a very limited family circle (and even there do not have the emotional freedom that women do), in the future they may be more liberated -- partly because of the computers. This change could affect family-life and lead to significant shifts in other social groupings and patterns of communication.

(I said "in the short term" in the previous paragraph, since there are persuasive reasons for believing that really intelligent computer systems could -- indeed would -- show emotions of one sort or another, even if they were not burdened by the same degree of potential conflict among basic goals as afflicts mankind [Sloman & Croucher, 1981]. A highly complex, wide-ranging, and subtle problem-solver, for example, could be expected to show analogues of emotions such as confidence, diffidence, anxiety, hope, resignation and the like -- at least insofar as these are functionally related to the self-monitoring of cognitive tasks.)

IV: Conclusion:

It is easier to sketch scenarios depicting the possible effects of AI on society than to be sure of what its effects will actually be. In general, social forecasting is a practice in which most of the hostages given to fortune end up losing their lives. A leading group of science-policy forecasters have admitted that we are very bad at predicting the money and time needed for the development of a technology, the extent of its acceptance by users, and its long-term social-psychological and ecological effects [Freeman & Jahoda (1978), p. 207]. Doubtless all social forecasters occasionally draw comfort from historical horror-stories: only a year before the Russian Revolution Lenin said that he did not expect to see it in his lifetime, and Rutherford could conceive of no practical application of the splitting of the atom.

Evidently, then, social anticipation is no science, and is more art than craft. Our picture of the future is only vaguely outlined. AI will surely tint the paint and colour the canvas, but what those colours will be is uncertain: they still lie unmixed on the palette.

REFERENCES

BODEN, M. A. (1977). Artificial Intelligence and Natural Man, New York: Basic Books.

BODEN, M. A. (1981). "The Meeting of Man and Machine/[1] in The Design o£ Information Systems for Human Beings (Informatics 6), eds. K. P. Jones & H. Taylor. London: ASLIB, pp. 4-15.

BODEN, M. A. (1982). [M]The Educational Implications of Artificial Intelligence/[1] in Thinking: An Interdisciplinary Report, ed. W. Maxwell. Philadelphia: Franklin Institute Press, (in press)

BUCHANAN, B. "EMYCIN," in O. Selfridge, M. Arbib, & E. Rissler (eds.), Adaptive Control in Ill-Defined Systems (in press).

CHILANSKY je£ aJU (1976). "Experiments with Computer Induction," Proc. Sixth Annual Int. Symposium jon Multi-Valued Logic (Utah).

DeJONG, G. F. (1979). Skimming Stories in Real Time: An Experiment jn Integrated Understanding. Yale University: Dept. Computer Science, Res. rep. 158.

FEIGENBAUM, E. A. (1979). "Themes and Case-Studies of Knowledge Engineering," in Michie (1979), pp. 3-25.

FIRSCHEIN, O. HJLL. (1973). "Forecasting and Assessing the Impact of Artificial Intelligence on Society," Int. Joint Conf. **Art**. Intell., 3, pp.105-120.

FORESTER, T., ed. (1980). The Microelectronics Revolution. Cambridge. Mass.: MIT Press.

Oxford: Martin Robertson.

GREGORY, R. (1967). "Will Seeing Machines Have Illusions?", in N.L.Collins & D. Michie (eds.). _Machine Intelligence 1_. Edinburgh: Edinburgh Univ. Press. Pp. 169-180.

KAHN, H., W. BROWN & L. MARTEL (1976). _The Next 200 Years_. New York: Morrow.

KAHN, H., & A. J. WIENER (1967). _The Year 2000_. London: Macmillan.

KUZNETS, S. (1972). "The Gap: Concept Measurement Trends," in G. Ranis, ed.. _The Gap Between Rich and Poor Nations_. London: Macmillan.

LEHMAN-WILZIG, S. N. (1981) "Frankenstein Unbound: Towards a Legal Definition of Artificial Intelligence," _Futures_, December 1981 (IPC Business Press), 442-457.

MAY, R., ed. (1961). _Existential Psychology_. New York: Random House.

MICHIE, D., ed. (1979). _Expert Systems in the Micro-Electronic Age_. Edinburgh: Edinburgh Univ. Press.

PAPERT, S. (1980). _Mindstorms: Children, Computers, Powerful Ideas_. New York: Basic Books.

PELTON, J. N. (1981). _Global Talk: The Marriage of the Computer, World Communications, and Man_. Brighton: Harvester Press.

SHORTLIFFE, E. H. (1976). _Computer-Based Medical Consultations: MYCIN_. New York: Elsevier.

SLOMAN, A. (1982). "Why Beginners Need Powerful Systems," _Proc. Inteer-_

press).

SLOMAN, A., & M. CROUCHER (1981). "Why Robots Will Have Emotions," _Int. Joint Conf. Art. Intell._, 6, 197-203.

WEIZENBAUM, J. (1972). "On the Impact of the Computer on Society: How Does One Insult a Machine?", _Science_, 176, pp. 609-614.

WEIZENBAUM, J. (1976). _Computer Power and Human Reason: From Judgment to Calculation_. San Fransicso: Freeman.

WILLICK, M. S. (1983). "Artificial Intelligence: Some Legal Approaches and Implications," _The AI Magazine_, 4 (No.2), 5-16.