

The Future of AI and its Impact on Society

Christian F.-Freksa

UC Berkeley

Two panel discussions on AI were held in the San Francisco Bay Area in May:

1. "The Future of AI", Stanford University, 11 May 1978, moderator: Peter Hart (SRI); panel: Daniel Bobrow (Xerox PARC), Nils Nilsson (SRI), Raj Reddy (CMU), Ed Feigenbaum, John McCarthy (both Stanford).
2. "AI and its Impact on Society", UC Berkeley, 23 May 1978, moderator: Lotfi Zadeh (UC Berkeley); panel: Daniel Bobrow, Nils Nilsson, Ira Pohl (UCSC), Jane Robinson (SRI), Milton Waxman (Hughes Aircraft), Bob Moore (MIT-Stanford), Hubert Dreyfus (UC Berkeley).

The first discussion held what its title promised: it was an occasion for theoreticians and implementers to make their well-known controversial points about the purpose of AI, how AI research should be conducted, and which results should be expected.

Peter Hart poses four questions for the panelists:

1. What are the most significant technological problems of AI?
2. What progress is likely by 1990?
3. What will be the effect of AI on society?
4. What could be done institutionally to speed up progress?

Feigenbaum starts off by noting that a focus on knowledge has been emerging in AI and feels a need for programs to read scientific text books to get knowledge into computers. Nilsson indicates that how to structure and process knowledge is an unsolved problem. Bobrow's greatest concern is the improvement of human-machine interaction to increment knowledge bases and Reddy analyzes the issue by distinguishing four types of knowledge, 1) algorithmic knowledge, 2) formal knowledge, 3) informal (commonsense) knowledge, and 4) new (presently not existent) knowledge. All panelists agree that AI is still an infant field and McCarthy points out that it took 90 years from the formulation of Mendel's laws to Watson/Crick's discovery of the genetic code - even though quite a few rather smart people had been working in the field. Reddy suggests that the short-term success of AI will be dominated not by the better understanding of intelligence but by the increasing availability of computing power. Nilsson believes "knowledge engineering" will emerge from AI in the next 10 years and asks McCarthy whether progress would not be faster if we worked on specific applications rather than on abstract problems. McCarthy admits that applications are easier to work on but points out that thinking is much more efficient.

The panel could not determine the kind of research administration that would best guarantee progress in AI. Opinions were split on the question of whether it would be desirable to have a large number of researchers working in the field.

Feigenbaum called on his experience as a defense project adviser to speculate that by 1990, in addition to commercialized games, AI research would come out with a "defense biggie" (like an information retrieval system) which would loosen up enough money to let the research coast along for a long time. A member of the audience disliked AI support

for the Army. Feigenbaum reacted sensitively and distinguished between Army and Defense Department but did not explain the difference in this context.

To a question from the audience as to whether we want a "defense biggie" which might interfere with people's privacy Reddy responded that sophistication in encoding of private information would proceed more quickly than the cracking of codes.

McCarthy believes that we have a genuine defense problem which we should help to solve. No one in the panel or audience contradicted him.

The debate was very disciplined: one argument was followed by a counterargument or a different point of view; the ball was played back and forth. The panel almost did not require an audience.

The title of the second panel "AI and its impact on society" was promising. Traditionally there was an annual debate between Dreyfus (author of "What computers can't do - a critique of artificial reason", 1972) and "AI" workers on the question of whether artificial intelligence is possible, at Berkeley. Should this debate give way for a critical reflection on our own role in society? Not only Dreyfus made this impossible!

Zadeh first explains the theme by noting that AI has by now come out of its closet and its impacts on society are visible to laymen. Therefore, the question of whether or not a machine can think appears less significant even though it has not yet been (or never may be) resolved.

The first speakers try to satisfy a request from the audience to explain what they mean by "artificial intelligence" and do so mostly by examples of what AI systems might do in the near future. Nilsson points out that robotics is beginning to have some impact in factories and says it is unclear whether the computer revolution will create more jobs than it will eliminate, as it was the case in previous industrial revolutions. He admits that AI systems might be misused: he fears that some day the message sending system could be used by commercial advertisement agencies to send him ads he is not interested in. But of course there will be help: we simply will need a program to filter out the undesired messages.

Waxman points out that there is a lot of interest in computer vision in industry at present which might have "consequences for management - labor relations" and suggests AI technology should develop an affordable "friendly household robot" which would not compete for jobs (since there is nobody who wants to clean houses anyway). He indicates that speech understanding might have a strong social impact if it were to be used by "Big Brother" to monitor conversations.

Dreyfus apparently had not realized that the object of the discussion was not to debate whether artificial intelligence can be created, but rather what impact whatever AI workers are creating has. He discusses whether or not people are programmed, thus preventing other discussants from talking about social impact of computer technology. After 20 minutes a computer enthusiast from the audience stops this aberration by asking how and when computer systems could be made more widely available for home use.

Moore returns to the theme of the panel and indicates that there might be a dramatic impact of advanced technology due to the tremendous availability of personal computers. He suggests that college dropouts rather than

academic researchers might raise the level of competence in AI with the help of homo computers. If a substantial amount of the work people do in our society will be done on a computer, they could do it at home thus decentralizing the work place and revolutionizing transportation patterns.

Robinson mentions that computers have become more intelligent in the past 20 years and might develop into clever cooperative systems that are more pleasant to deal with than inflexible people. She agrees with a statement from the audience that computers may build walls between people but says they also might take away walls by letting them communicate more easily.

Robrow views AI as a study of intelligent behavior independent of the medium used and expects the first impact of AI visible to the general public to be intelligent consulting programs which are starting to emerge now.

Pohl compares the possible impact of interactive computing systems on society to that of television but says it might be stronger, since TV leaves the user passive whereas computers demand user participation. He sees a need to socialize computers to prevent the possibility that children are able only to talk to computers but cannot play baseball anymore with other children. A positive aspect of interactive computing is that "technological dropouts" might find it more pleasant to study on their own pace by using a computer rather than studying in a classroom situation.

The discussion is lively, the panel is integrated into the audience, the audience participates and tries to redirect the discussion to its nominal topic - with very limited success. One participant is concerned about the possible effect if goals and purposes of people are changed by computers. He is countered (guess by whom) that there will not be any effect on society since artificial intelligence is not going to exist. Another participant detects a mismatch between all the "groovy things" the panel talks about and the fact that a lot of weapons research and industrial research is not directed towards "groovy things". He compares AI research to atomic research and asks the panel how they feel about what they are doing. Nilsson's response sounds frightening in the absence of any background to his conclusion: he feels his work is important for society and imagines that there could be circumstances under which he would work on "some defense item".

Waxman explains the prevalence of military-type research in image understanding by noting that they (the military) are the ones most interested in it and they pay for it.

A participant generalizes his experience during the discussion by saying that mankind apparently has not learned anything from the past. Another participant asks why there are no experts on society on the panel but only computer experts. Another participant says he is scared at the irresponsible unconcern of the panel and wonders why they justify their field so poorly. Bobrow concludes by admitting potential dangers of AI and says that we (the AI workers) have to decide how to use intelligence appropriately.

The two panels discussed mainly direct effects of AI research and of those effects mostly the intended ones. It seemed the panelists had not been thinking very much about possible side effects. They emphasized the potential power of information and of information processing systems and made clear that it is non-trivial to use them in order to obtain the desired results. But they did not assume responsibility for intentional or unintentional consequences of their work in our

social environment. Sometimes there seemed to be an implied consent that "society" will decide what it wants and that it will have the expertise to make good decisions.

Many issues remained untouched. In particular, the panels did not discuss

- how they viewed their responsibility to society
- who should be allowed to use their research results or how misuse can be prevented
- whether the comparison to the situation of the atomic researchers is valid, and if so, how can we deal with it?
- the promising work of AI for medical diagnosis and for the handicapped
- that machines can do work which would be dull or dangerous for people
- whether a machine should replace people for economic reasons
- that computers may alter the self-image of people
- that computers have become bureaucratic scapegoats
- whether there might be developments for which the society is not ready yet - in the sense that it might not be ready for TV if people become dominated by it, for example
- how they imagine "society" should decide how to use advanced technology. Society is made up of mostly non-experts and thus usually able to judge impacts of new developments only after the fact.

In conclusion, it seems desirable that discussions on the impact of AI continue within the AI community and the suggestion seems good that the discussion should include sociologists, psychologists and other experts in fields indirectly related to AI.