THE WEB—EARLY VISIONS, PRESENT REALITY, GRANDER FUTURE

John McCarthy, Stanford University
mccarthy@stanford.edu
http:/wWww-formal.stanford.edu/jmc/
March 12, 2005

- Turing—1936 Universal computer, Wilkes—1948 Edsac, von Neumann,(Eckert and Mauchly)—1946 stored program,

- Vannevar Bush—1945-Memex, doesn't mention universality

- Turing 1947 AI $\rightarrow$ (1970s applications) $\rightarrow$ 20xx human level

• McCarthy 1958—Proposal for formalizing common sense, far from accomplished today

• 1950s SAGE system, special purpose time-sharing

• McCarthy—1959-1962-1970—universal time-shared computer utility, motivated by Advice Taker proposal

• Licklider—1960—Man-Computer Symbiosis

• Roberts—1970—ARPAnet $\rightarrow$ Internet

• Engelbart—1962-1968—Mouse, linked documents

• Kay—1970—Dynabook

- Berners-Lee—late 1980s and early 90s—World Wide Web

- Brin and Page—1996—Google—first adequate search engine

- other prophets—Nelson, etc. whom I neglect undeservedly from ignorance.

SOME EARLY PREDICTIONS WORKED OUT—OTHERS NOT

• Time-shared public utilities. Modest success. Lack of machine power, needed too much handholding. Worked fine in labs

• Stanford AI Lab news service, 1972-1989. Prototype web newspaper.

• Access to all the world's books. Still hasn't happened. Making steady progress for scientific articles. No economic model for literature except what's out of copyright. John Ockerbloom at the University of Pennsylvania links to more than 20,000 free-to-read books.

• On-line buying and selling. I don't think anyone predicted Ebay auctions in the 1970s.

TIME-SHARING vs. PERSONAL COMPUTERS

• General purpose time-sharing was proposed in 1959 and realized in 1962. Gave each user his share of the computer at his fingertips. What ever were punched cards? The Digital Equipment Corporation PDP-6, PDP-10 series did it best but wasn't pushed in competition with IBM.

• TX-2 about 1960 was a $500,000 personal computer. The first SUN was a $20,000 personal computer.

• The D.E.C. PDP-7 computers on which Unix was developed and first IBM PCs were too weak, but the operating systems kept adaptations to the weaknesses in vastly stronger computers.

TIME-SHARING vs. PCs, part 2

• The PC hell is system administration. Need AI to do it properly on a mass scale.

• Software bloat

• Operating systems as products require the user to do sysadmin for every new version. A time-sharing subscriber would not. There are probably 100 times as many system administrators as should be needed.

• Presently promised "set top boxes" seem to aim at monopoly. Sysadmin is centralized but is probably too little and too late.

FUTURE

• The present web is pretty good. The users will do ok even without new ideas. Those whose business is new ideas will suffer unless they have good new ideas. The dot com crash was substantially due to a large number of bad or trivial new ideas.

• Everyone has trouble using something new. Systems must understand user states of confusion. Trivial example: a user confuses IP address, email address, and URL.

• It is more important for a system to understand a user's confusion than to offer sympathy. (Some advocates of "emotional computing" are hoping to get by with sympathy. My bet is that mere sympathy will only produce annoyance.)

1970 MODEL WORLD OF THE FUTURE

This 1970 conference article "The home computer terminal" was published in *Man and Computer,(Karger, Basel 1972)*. It's available as
www.formal.stanford.edu/jmc/hoter2.html.

Here are some fragments of the 1970 article, with notes in blue.

"At present, a newspaper or a book is a package produced by a large organization.

"In our new system, the physical production disappears allowing a much smaller organization to put out the same packages of text and pictures. Moreover, the user does not face a one shot

decision to buy Life or Look. He will be able to read the 'cover' or table of contents of each and read such items as strike his fancy, and the system will bill him for what he reads from each source. In fact, since the cost of keeping a file of information in the computer and making it publicly available will be small, even a high school student could compete with the New Yorker if he could write well enough and if word of mouth and mention by reviewers brought him to public attention. What, then, is a publication in the new information system?"

Note 2000: I underestimated the resistance to being displaced these organizations would be able to mount. Resisters include the publication organizations of non-profit scientific societies. I also underestimated the fraction of the cost of producing a newspaper that would persist even if the newspaper were entirely on-line.

Note 2004: Four years later, the biologists have taken the lead in creating on-line journals that compete with print journals. Their financial basis is page charges, which works for science, because page charges are a small fraction of the cost of doing the research, but that model won't work for people who make a living by writing.

Blogs come closest to my predictions, but they can't afford famous professional writers.

2004 November note: Some of the politcal bloggers made a lot of money from advertisements during the election campaign. The resulting competition should professionalize blogging.

"A publication is an organization that puts out a list of material it has edited and recommends to its readers. It helps its authors

produce material that it thinks will suit the readers, and it has a financial arrangement with them about splitting the proceeds.

"There can be a wide variety of publications of different standards of writing and editing and different budgets for carrying out these activities.

"However, they will all be equally accessible to all readers, and the only justification for an expensive editorial organization will be that it can produce a more popular package. The price of reading a package can be set by the publishers."

Note 2004: This ignores the copying problem. There still isn't a general purpose pay-by-the read mechanism. Moreover, the difference between professional (full time) writing and semi-professiona writing is likely to persist.

The star phenomenon will persist and become even more dominant.

"A reader may feel that he needs help in finding his way through the totality of literature available to him. Various people will be eager to make a living by providing it. A bookstore or library is a program that when called shows the 'covers' of publications. Reviewers will produce lists for him and make money when he reads their lists or by kickbacks from the publishers. 'Reading advisers' under some catchier name will offer to generate lists just for him according to a profile of his interests."

Note 2004: This hasn't happened enough to make writers independent of publications. The 1970 article didn't take into account the importance of publicity and advertising.

"Advertising in the sense of something that can force itself on the attention of a reader will disappear because it will be too easy to read via a program that screens out undesirable material."

Note 2004: This didn't happen, perhaps fortunately. Also I didn't predict spam. I am temperamentally an extreme optimist, but the pessimists didn't predict spam either.

"Another effect is the possibility of frequent revisions of articles and books. An author can take into account new facts or other people's criticisms, and the revision will take effect immediately." Readers of an old version will be unlikely to read a work again even if it contains important changes. Better put the changes in a new article.

Note 2004: I put dated footnotes on my old articles, but I doubt their existence lures anyone to read the article again.

"Public controversy can be carried out more expeditiously than at present. If I read something that seems controversial, I can ask the system if anyone has filed a reply. This, together with an author's ability to revise his original statement, will lead people to converge on considered positions more quickly than at present even if they do not come to actual agreement."

Note 2000: There are various proposals, but this hasn't happened yet. One can imagine Bush and Gore "truth squads" putting on their candidates' web sites arguments against the positions of the other guy. Personal attacks too.

Note 2000 June 1: Today's New York Times has an article entitled "E-Mail Messages to the Press Have Made the Gore-Bush Race a Cyberwar" recounting how the Gore and Bush campaigns send dozens of messages per day to reporters. I suppose this is a partial realization of my 1970 prediction.

Note 2004: The campaigns have their web sites, but they still aren't the main places undecided people go to see arguments refuting those of the other side. Alas, TV advertising is still the main way of influencing the voters.

"Famous authors will not need publishers because their loyal readers will have the system find their stuff automatically."

Note 2004: A try at this failed because of copying

"To summarize: the new information system will promote intellectual competition by reducing the price of entry, will permit readers to be selective, and will allow authors to revise material until they are satisfied that it withstands criticism as well as it ever will. This should make intellectual life more interesting."

Note 2004: This doesn't seem to happen much. Instead of perfecting their earlier analyses, bloggers just bombard their opponents with new stuff.

"The new information system will have a profound effect on buying and selling. Sellers of movies, groceries, automobiles, plumbing services and cures for baldness will find it advantageous to list their wares in the information system together with current

prices and availability. The user can place an order through the system as he can by telephone, but he can do much more:"

Note 2004 : This happened but isn't revolutionory.

"(1) He can call on someone's program to scan the sellers of sports cars and propose what it considers the best deal. This program might even negotiate with programs representing the sellers.There's some of it now.

"(2) He can tell the system whether last year's cure for baldness worked and get a summary of the opinions of those who bothered to record their opinions of the cure he contemplates trying now.

"(3) He can make an airplane or hotel reservation by interacting with a program the airline or hotel reservation company has

written to tell him what is available. He need not suffer the delays you now get when you call an airline or travel agent at peak hours." All this has happened.

"(4) Individual design and construction services can be offered through the system although this requires the development of computer-controlled manufacturing techniques for various types of articles. The idea is that automated design programs can produce designs for articles meeting individual specifications.

Note 2004: This hasn't happened yet. Maybe it will.

"There are many more useful services that can be offered through the new information system and again the system is conducive to competition. Writing and storing a program and announcing its

availability can be a very low capital operation, and the system can collect whatever price has been set for its use."

Note 2004: This has happenened, but In the world of pcs, this is far less convenient than in a world of time-sharing—or than it should be.

Note 2004: The above greatly underestimates the role refereeing and publicity of all kinds plays in creating reputation and getting attention to ideas.

2004 note: My 1970 article did not see AI as as essential as I now see it to help people use computers.

THE FUTURE OF PERSONAL COMPUTING INVOLVES AI, BETTER THAN THE AI OF TODAY.

Example: swindle protector

- A low level protector knows about specific swindles.
- Higher level can identify variants of the Nigerian scam.
- High level—knows facts about swindling in general.

Example: Understanding a user's confusion.

• Suppose the user confuses IP addresses and URLs. Suppose a program asks for an IP address, and the user gives a URL. Most present programs will simply put up an OK box that says "wrong format". The user may just worry about the format of the URL. A system designer who anticipated the confusion would have the program say "You gave me a URL when I asked for an IP address."

• More generally, system administration requires knowledge and reasoning. Evidence: The people who spend several hours fixing my problems obviously think a lot. They understand enough to fix my problems, but they don't understand enough about how they do it to automate their work.

LOGICAL AI IS PROBABLY NEEDED TO UNDERSTAND USERS' PROBLEMS

Logical AI involves expressing what is known about the world, especially common sense knowledge, in languages of mathematical logic. The logical AI program infers from the sentences about the world and sentences about a particular situation that a certain course of action is appropriate. The main scientific basis for logical AI is mathematical logic as developed since 1879. Newer methods of *non-monotonic reasoning* developed since the 1970s are also needed.

Logical AI is based on study of the world and the actions that achieve goals. Its main rival is based on stuying human and animal neurophysiology. Both approaches have been pursued for

50 years, and neither has reached human-level AI yet. Excuse: It took the geneticists about 100 years from Mendel to the genetic code, and genetics isn't done yet.

BIG ADVANCES REQUIRE LOGICAL FORMULATION OF COMMON SENSE KNOWLEDGE AND REASONING

• Humans mainly communicate in facts, not just rules or programs.

• Humans reason to get new facts from old. Logicians formalized these rules. Gödel proved them complete.

• Reasoning programs require full first order reasoning.

• Advanced help requires understanding the problem domain and usually understanding the user's state of mind.

THE FUTURE OF THE WEB—PART I

• Everyone will be able to read anything. The problem of paying authors will be solved.

• While world population won't even double, the public will incread by a factor of five.

• Very specialized interest will have adequate publics.

• There will be more rich stars—for better or worse.

THE FUTURE OF THE WEB?—PART II

Programs that understand

- substantial parts of natural language documents,
- facts about the world,
- facts about people's states of mind, including confused states of mind,
- can give good advice,
- and can put together programs from this information.