FROM HCI TO AAI
Some Bits of History?

Version April-19, 2018
Draft Version
eJournal: uffmm.org, ISSN 2567-6458
19.April 2018
Email: info@uffmm.org

Gerd Doeben-Henisch
Frankfurt University of Applied Sciences (FRA-UAS)
Email: doeben@fb2.fra-uas.de
Emauil: gerd@doeben-henisch.de

✦

CONTENTS

Abstract

The Actor-Actor Interaction (AAI) paradigm as published on the website https://www.uffmm.org/2017/07/27/uffmm-restart-as-scientific-workplace/ has many roots in the history. In this text some of these roots will be presented. The actual text covers the time span from about 1945 until about 2000, and it is not exhaustive. The main focus are not the details of the history but rather the 'big ideas' which have laid the ground for the actual AAI paradigm.

1 USER CENTERED, ENGINEERING, LEARNING

The main characteristics of the actual Actor-Actor Interaction (AAI) paradigm are its 'user-centeredness', the framework of an embracing 'engineering process', as well as the integration of the topic of 'learning systems'. The last topic is today mostly known in the (restricted) version of 'artificial intelligent systems', which show some phenomena of 'intelligence' and – not always – of 'learning'. These main points are further 'enhance' by additional 'formal models' which allow new kinds of precision, of testing, and of simulations.

Looking back in history often can help to 'remember' the original, founding ideas which caused so many processes in research and industry, whose richness can take you 'away' from the underlying basic principles.

Seeing this great variety of concepts and methods which manifest themselves in the hundreds of research papers than one can doubt whether it could be be possible to 'integrate' all these aspects in one coherent view.

2 GUIDES INTO HCI-HISTORY

Asking for guidance in the history of the Actor-Actor Interaction paradigm leads back to the history of the Human-Machine Interaction (HM) paradigm which in turn follows the tradition of the Human-Computer Interaction (HCI) paradigm. For this you can find guidance in the two human-computer interaction handbooks from 2003 and 2008, and here especially in the first chapters dealing explicitly with the history of HCI (cf. Richard W.Pew (2003) [Pew03], which is citing several papers and books with additional historical investigations (cf. p.2), and Jonathan Grudin (2008) [Gru08]. Another source is the 'HCI Bibliography : Human-Computer Interaction Resources' (see: http://www.hcibib.org/), which has a rich historical section too (see: http://www.hcibib.org/hci-sites/history).

3 1945-VISION

I decided to start with the impressive vision of Vannegard Bush, the director of the Office of Scientific Research and Development, who has coordinated the activities of some six thousand leading American scientists in the application of science to warfare. He published an article July 1945 in the newspaper 'THE ATLANTIC MONTHLY', where he opened a vision for a possible better future for mankind by envisioning a multitude of technological improvements. (see: [BUS45]).

In some sense one could begin before 1945, because during the World War II many activities were going on in many countries. Especially in England and in the United states a new kind of 'engineering psychology' has emerged, dealing with the special challenges of humans under new and difficult conditions. In this time the 'Human Factors Society (HFS)' of America has been founded by the American aviation psychologists. Later this title has been changed to 'Human Factors and Ergonomics Society (HFES)'.[1]

But for our purpose it is enough to start with the vision of Bush which he 'donated' to the post-world war society.

Bush started his considerations with the "summation of human experience", which "is being expanded at a prodigious rate, and the means we use for threading through the consequent maze to the momentarily important item is the same as was used in the days of square-rigged ships."(p.3) This explosion of experience can be disastrous, but he interprets the different new technologies as signs of hope for managing the problem.(cf. p.3)

He continuous in his vision with a long list of possible technological improvements. Later he switches to more foundational issues. One is his example of a typical mathematician: "A mathematician is not a man who can readily manipulate figures; often he cannot. He is not even a man who can readily perform the transformation of equations by the use of calculus. He is primarily an individual who is skilled in the use of symbolic logic on a high plane, and especially he is a man of intuitive judgment in the choice of the manipulative processes he employs."(p.11) And he infers from this that a mathematician will only becoming more productive with the new technologies, when the mathematician can use these new technologies somehow directly to interface with his thoughts. And he is sure that "... there will come more machines to handle advanced mathematics for the scientist. Some of them will be sufficiently bizarre to suit the most fastidious connoisseur of the present artifacts of civilization."(p.11)

What one can already be seen here is a clear consciousness about a 'gap' between the actual technology and the way humans (here illustrated with the mathematician) are thinking, this paired with the before mentioned insight in the necessity to improve the tools dedicated to support the human mind confronted with the expanding social reality.

Bush points to the actual simple indexing techniques used in his time and states, that "The human mind does not work that way. It operates by association. With one item in its grasp, it snaps instantly to the next that is suggested by the association of thoughts, in accordance with some intricate web of trails carried by the cells of the brain. It has other characteristics, of course; trails that are not frequently followed are prone to fade, items are not fully permanent, memory is transitory. Yet the speed of action, the intricacy of trails, the detail of mental pictures, is awe-inspiring beyond all else in nature. Man cannot hope fully to duplicate this mental process artificially, but he certainly ought to be able to learn from it."(p.14) And he concludes from this that "selection by association, rather than by indexing, may yet be mechanized."(p.14)

---

[1] See https://www.hfes.org/ContentCMS/ContentPages/?Id=ClyOCRldYfA=

And from these considerations he infers the idea of a future device for individual use, which he calls ’memex’. He describes this ’memex-device’ as ”a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory. It consists of a desk, and while it can presumably be operated from a distance, it is primarily the piece of furniture at which he works. On the top are slanting translucent screens, on which material can be projected for convenient reading. There is a keyboard, and sets of buttons and levers. Otherwise it looks like an ordinary desk.”(p.15)

Today, in the year 2018, many aspects of this vision of 1945 are indeed realized, are accessible, can be used. But there is an interesting difference: while popular cloud-based computer programs can support an individual user with a huge amount of general knowledge (by keeping a lot of private data trapped in a public system), there is no device available which is specifically dedicated to a certain user, bing his personal assistant in a full sense, and keeping full privacy. There is no ’personal enhancement’ in the sense of a close man-machine symbiosis. Rather the discrepancy between the individual and the ’big system’ is growing.

And Bush goes even further in his visions, thinking the future man-machine symbiosis even more radically, by asking, ”whether it could not happen in the future that we can intercept these neural currents [which are interchanged between the body-limbs and the brain] ... either in the original form in which information is conveyed to the brain, or in the marvelously metamorphosed form in which they then proceed to the hand?”(p.18)

4 INTERFACING THE MACHINE

4.1 Grace Murray Hopper

As Bush pointed out in his visionary paper from 1945 there is still a big gap between the computing machines of that time and the needs of human persons who wanted to use these new machines. In a certain sense the spirit of Bush was living in the minds of many experts of these days. A good insight in this long process of ’taming’ the new machines by stepwise improving the interactions with these machines can be gained from the keynote-speech given by Grace Murray Hopper 1978, which worked in the development of coding and programming computers from the earliest beginnings.[2](For the speech see: Hopper (1978) [Hop81])

Hopper sees the main challenge for the development of computing in the task ”to recognizing that we have a large variety of people out there who want to solve problems, some of whom are symbol-oriented, some of whom are word-oriented, and that they are going to need different kinds of languages rather than trying to force them all into the pattern of the mathematical logician. A lot of them are not.”(p.11)

But Hopper made often the experience that in the reality of developing new software – at that time the developers of programs have been rather called ’coders’ than ’programmers’ – the developers often were depending from their actual knowledge, from their patterns of thinking, which produced an obstacle to find something new. And from this experience Hopper draw the conclusion ”that we would do well to train our computer people and specialists in many disciplines rather than in a single discipline because you never know what may prove useful. ”(p.11) A special type of influence on the thinking process was according to Hopper, that the constraints of the machines in usage have been reflected into the thinking of the coders too: ” ... the fact was that we were living in an environment which consisted of a 12 alpha-decimal word, and it became perfectly obvious to us that the entire world operated in 12 alpha-decimal characters of which the first three defined an operation, the next three one input, the next three another input, and the last three the result. And it was clear that the world operated this way, and that the world lived in 12 alpha-decimal characters, and a three-address machine code. ”(p.12f)

Knowing this background it was not obvious that and how thinking about coding could change. One major pressure was always the pressure of time and the demand for results to become faster and better. Thus the tedious ways of programming was a strong stimulus to improve coding. In 1953 ”... People were beginning to have some interest in the possibility of providing some assistance for writing programs....” This happened in ”... an atmosphere which was totally unwilling, or seemingly so, to accept a new development, because it was obvious that you always wrote your best programs in octal. ” (p.14) And: ”I find today that many people are concentrating on the languages, and not doing very much about the tools with which to implement those languages. The art of building compilers, as far as I’m concerned, has been pretty badly stultified.”(p.15)

One nice example for a new programming interface using nearly everyday words is the following one (p.17):

```
INPUT INVENTORY FILE A;
PRICE FILE B;
OUTPUT PRICED INVENTORY FILE C.
COMPARE PRODUCT #A WITH PRODUCT #B.
IF GREATER, GO TO OPERATION 10;
IF EQUAL, GO TO OPERATION 5;
OTHER- WISE GO TO OPERATION 2.
TRANSFER A TO D; WRITE ITEM D;
JUMP TO OPER- ATION 8.
It ended up: REWIND B;
CLOSE OUT FILE C AND D;
and STOP.
```

This is an example of the FLOW-MATIC language. The FLOW-MATIC generator was again used as a tool to develop the first COBOL compiler.(p.19) Another set of programs was developed to validate a COBOL compiler: ”They would

---

[2] To the person of Grace Murray Hopper see: https://en.wikipedia.org/wiki/Grace_Hopper

compare the execution of the compiler against the standard, and monitor the behavior of the actions of the compiler. It was the first set of programs that was built to try to use software to check softwar."(p.19)

With regard to the methods of engineering Hopper mentioned that it was a standard for engineers when designing something new that they will build "either a pilot model or a bread board" to prove feasibility.(p.16) She mentioned as well the importance of documentation with giving an example: "... In the first place, there's a write-up of the compiler and how it's going to work:... There is then a block chart of what it's going to do at a high level; a flow chart at a detailed level. The coding is here. Along with the coding, there are comments on it. The coding and the flow charts are keyed together with little circular things in the flow charts. And following that, there is a line-by-line description of exactly how the compiler works, in English, keyed both to the flow charts and to the coding. All this was done because I knew damn well nobody was ever going to believe it was going to work! And probably maybe that's one of the best reasons for documentation that ever existed– when you have to convince somebody that the darn thing will work. "(p.12)

Besides the problem, that the developers can hinder themselves by having the wrong patterns of thinking Hopper did also point to the managers of the projects which too can be an obstacle to innovation. Inventions of new methods pose always some frictions on the environment and is usually not very welcomed. She mentions the example with the new compiler using 'English vocabulary': "It was a long, tortuous, and difficult job to get that concept accepted, because it was of course obvious [for the managers] that computers could not understand plain English, ..."(p.18) But she firmly confessed her attitude to be oriented for innovation by saying ".. that I always have to push into the future" and at the same time "to remind me that any given moment in time there's always a line out here, and that line represents what your management will believe at that moment in time. And just you step over the line and you don't get the budget."(p.16)

Comparing the vision of Bush how to bridge the gap between the human way of thinking and the new machines with the step-by-step work of interface development during 1945-1955 described by Hopper gives some flavor what this means. At one hand you need some key-technology to be able to built concrete machines, keyboards, cathode-ray-tubes, punched-card readers, memory devices etc., but you need too human persons which are able to 'think in new ways'. The technology as such does not created any single new idea. And these new ideas require lots of new 'languages' for representations and communication. These new languages have to be invented, experimentally being explored, and then applied in new ways. Thus who wants more 'new' future must invest in creative environments enabling real new thinking and new engineering.

4.2 More Interfacing

Clearly, the time between 1945 and 1965 includes many more events, people, and inventions. Here only a small selection of important data.

Appearing in 1956, the TX-0 computer from MIT was an early fully transistorized computer and contained a then-huge 64K of 18-bit words of magnetic core memory.[3] The TX-0 and its direct descendant, the original PDP-1, were platforms for pioneering computer research. Soon, in 1958, followed the TX-2 computer, also transistor-based, with 36-bit words (compared to the TX-0).[4]

Important languages in these time have been FORTRAN (1957, John Backus and IBM), COBOL (1959 by CODASYL, and was partly based on previous programming language design work by Grace Hopper), JOSS (1964, RAND corporation), BASIC (1964, John G. Kemeny, Thomas E. Kurtz and Mary Kenneth Keller), TELCOMP (1964, Bolt, Beranek and Newman (BBN)), BCPL (1967, M.Richards), C (1972, Dennis Ritchie).

According to Grudin [Gru08]:p.5 one of the most influential publications in the history of computer science appeared 1963 with the PhD thesis from Ivan Sutherland [Sut]. Having access to a TX-2 computer he could write his PhD thesis about a new graphical software called 'sketchpad' which revolutionized the work with 3D graphics in the computer. He introduced many new concepts, especially object-orientation. Based on his software the first computer game ever has been created played on the a the at that time popular PDP-1 computer. (see: [Gru08]:p.5)

An innovative invent was the the introduction of the new spreadsheet paradigm exemplified by the VisiCalc program on Apple II (1979).(cf. [Pew03]:p.9)

4.3 Licklider

These new hardware and new programming tools paved the way to a more human-friendly interaction with computers and reinforced more general visions about a possible man-computer symbiosis. One of these inspiring persons was J.C.R.Licklider. In his vision of the upcoming man-computer symbiosis he did not limit the considerations to the individual improvement between one user and the machine but he saw a "close coupling between the human and the electronic members of the partnership."(see: Licklider (1960) [Lic60]:p.4) His extending vision was "to enable men and computers to cooperate in making decisions and controlling complex situations without inflexible dependence on predetermined programs. In the anticipated symbiotic partnership, men will set the goals, formulate the hypotheses, determine the criteria, and perform the evaluations. Computing machines will do the routinizable work that must be done to prepare the way for insights and decisions in technical and scientific

---

[3] Fo more details see McKenzie et.al (1974) [McK]

[4] For more details see Clark et.al (1957) [CFP$^+$]

thinking. Preliminary analyses indicate that the symbiotic partnership will perform intellectual operations much more effectively than man alone can perform them. Prerequisites for the achievement of the effective, cooperative association include developments in computer time sharing, in memory components, in memory organization, in programming languages, and in input and output equipment." [Lic60]:p.4

With such a global extension to networks of computer and people Licklider created the vision of 'Libraries of the future'. (cf. [Pew03]:p.4). Today, 2018, we can confirm Licklider's vision as we have expanded our libraries from the past tremendously.

4.4 Personal Workstations

Alan Kay propagated with his PhD-thesis (1969) the idea of a personal computer thereby introducing new concepts of parallel processing for multiple windows, message passing and object-oriented software. The vision was a 'computer as a tool for logical simulation and modeling of conceptual worlds'. Inspired by Moore's law Kay envisioned a time where hardware was cheap and powerful enough to support this vision. He moved to the Xerox Palo Alto research Center (PARC), which has been founded in 1970. The outcome was the smalltalk programing system, which could be run on the PARC developed computer system ALTO, which has been embedded in a PARC-developed Ethernet.(cf. [Pew03]:p.7)

In this line of thinking emerged from Xerox PARC the common era of 'Graphical User Interfaces (GUIs)' started with the Apple Lisa computer (1983) and then the Macintosh (1984). The 'What You See is What You Get (WYSIWYG)' principle was born, further supported by the new laser printer. The desktop metapher with its icons and the mouse made the user interface to an entity separated from the application code. The technological capabilities were much wider than the HCI community could handle. ([Pew03]:p.9)

In parallel there were activities to improve the programming of GUIs with libraries and tools. One of the first was the 'Steamer' software from Stevens et.al (1983) [SRS83], a GUI for computer aided instruction, which has been improved by Henderson (1986) [Hen86] with the 'Trillium' software. Apple incorporated widges directly in the computeer combined with fixed guidelines for all developers.(cf. [Pew03]:p.10)

4.5 Networks

Prepared by several developments before the 1990s the new Internet, the 'World Wide Web', became a reality (for details see the time line here: 'The Hobbes' Internet Timeline' at https://www.zakon.org/robert/internet/timeline/). Part of the new network world have been 'Computer-Supported Collaborative Work (CSCW) too, e.g. 'Lotus Notes Domino' and 'GroupSystems'. Both raised new challenges for HCI.

To mention are the growing possible extensions of the Internet in everyday life by new devices, making computing somehow 'ubiquitous' (see Weiser (1993) [Wei93] 'Some scientific issues in Ubiquitous Computing'). Touching the body or even becoming part of the body raises the vision of new 'cyborgs' (see: Norman (2001) [Nor01]).

The simple usage of web-pages, as it appeared at a first glance, was accompanied by a bunch of new problems which all called for new support by HCI experts. The need for usability became stronger than ever. More and more companies installed usability departments to serve this demand.(cf. [Pew03]:p.13)

4.6 New HCI

All the before mentioned developments of new hardware, new graphics, networks, new kinds of interactions, etc. have been new challenges for the field of HCI.

In this time HCI became a professional discipline, more and more textbooks appeared, with 1982 first conferences dedicated only to HCI started, and in 1982 first Journals dedicated only to HCI began publishing.(cf. [Pew03]:p.11)

One of these new groups was based in the Xerox PARC started by Newell and Simon, who revolutionized thinking by their theory of 'human problem solving' (1972), and together with Simon he was proposing a new project for applied information-processing psychology. The new project started in 1974 with the team Stuart K.Card, Thomas P.Moran, and Allen Newell. They did ground-breaking HCI research. They worked out 'predictive models', the 'model human processor', the GOMS (Goals, objects, methods, selection rules) paradigm, and this work culminated in the highly influential book 'The Psychology of Human-Computer Interaction' (1983) [CMN83].

In parallel there was some activity in Great Britain in the realm of psychology and ergonomics (Brian Shackel, D.E.Broadbend, Alan Baddley) as well as computer based learning (T.R.G.Green).(cf. [Pew03]:p.8)

One of the first style-guides was published by Pew and Rollins (1976) 'GENERIC MAN-COMPUTER DIALOGUE SPECIFICATION: AN ALTERNATIVE TO DIALOGUE SPECIALISTS' (http://journals.sagepub.com/doi/pdf/10.1177/154193127602001302);
in contrast a huge amount of specific guidelines by Sidney L. Smith and Jane N. Mosier (1986) http://www.dfki.de/~jameson/hcida/papers/smith-mosier.pdf

Many new textbooks appeared. A very influential one was Nielsen (1993) [Nie93]. He tried to find 'discount' versions of the official psychological procedures, more tailored for the practical applications (e.g. 'prototypes on paper', 'simplified thinking aloud', 'heuristic evaluations of expert thinking').

The other paradigm was the introduction of iterative development models, especially the incorporation of the human factors aspect into the software life cycle process (see e.g. Mantei et.al. (1989) "Incorporating Behavioral

Techniques into the Systems Development Life Cycle" ([MT89]).The paper shows nicely that the restriction to psychology for the development of a system interfaces without integrating this into a more general systems engineering framework is not sufficient! This idea has been repeated by Zhang et.al (2004) [ZCT^{+}04]. Anotheer variant is given by Baxter et.al. (2011) [BS11].

4.7 Resume

From the point of view of the AAI paradigm it is interesting that the idea of integrating the HCI paradigm within the more general systems engineering framework is there, but very rare. Until now I could find only three papers in the time-span 1989 - 2011.

But there is an interesting 'side case' in that the discussion about the necessity and possibility of a trans-disciplinary discipline of 'cognitive science' generates also arguments that psychology alone is not sufficient to handle the case of Human-Computer Interaction.[5]

5 BRIDGING THE PHYSICAL-MENTAL GAP

In the rest of this small introduction about the history before the Actor-Actor Interaction paradigm I will mention a few selected positions, which are focused around the kind of interactions between a human actor and a machine as actor.

5.1 Norman - 1986

A good starting point seems to be the position of David Norman which is described as the *User Centered* perspective. (cf. Norman and Draper (1986) [ND86]) Within the user-centered perspective it is the 'gulf' of 'execution' and of 'evaluation' ([ND86]:p.38f) which is revealing itself in a more detailed analysis. For Norman a human user starts with some 'goals', deriving from these some 'intentions', this leads to the specification of an 'actions sequence', followed by an 'action' in the real world, associated with a system state, which can be 'perceived' and which has to be 'interpreted'. What does follows from this interpretation for the 'intentions' and the driving 'goal'?(cf. [ND86]:p.41)

For Norman follows from this situation that an 'improvement' of this interaction can only be reached if either the mental system of the human actor comes closer to the reality of the system or vice versa the reality of the system can be adapted to a higher degree to the mental states.(cf. [ND86]:p.43)

It is in this context where Norman assumes that humans organize as part of their interactions with the environment 'mental models' of the environment and their interactions. The success of the interactions depends very strongly from the degree of matching the reality with the mental model.(cf. [ND86]:p.46)

From the point of the designer of a system Norman claims further that the designer has to organize the 'system image' (interface, documentation, instructions...) in a way that a user will be supported to get an optimal 'mental model' from the working system. Thus their is the pair 'design model/ system image' versus 'mental model/ user model/ conceptual model'. Besides this Norman mentions at least the possibility, that the system is an 'intelligent system' that constructs a model of the interacting (human) user. This would be the systems model of the user. (cf. [ND86]:p.46f)

Norman mentions then many examples of design models with a good 'system image' which for him are all systems which function as 'powerful tools for the user'.(cf. [ND86]:p.48-51)

The main requirement which Norman gets from his analysis assumes that the user needs should have priority before the system design. First one has to design the interface which meets the user needs optimally, and then one has to implement the system which is supporting the designed interface.(cf. [ND86]:p.59-61)

5.2 Engelbart - 1962/3

Although Douglas Engelbart precedes Norman about 20 years I discuss his ideas here. In this text I am focusing on his report from 1962 [Eng62](but, see also Engelbart (1963) [Eng63]). With a good funding Engelbart created a computer-supported environment for to accomplish intellectual work at the Stanford Research Institute, which culminated in an impressive demonstration at the Fall Joint Computer conference in San Francisco 1968.(cf. [Pew03]:p.4f) The rationale behind this technology is described in this paper.(cf. [Eng62]:p.1)

For Engelbart the situation of a human person is characterized by a growing complexity, where the problems grow faster than the capabilities of man are developing. Therefore he thinks that "augmenting mans intellect" is an important task. In the paper he describes the report which "covers the first phase of a program aimed at developing means to augment the human intellect. These means can include many things... and we consider the whole system of a human and his augmentation means as a proper field of search for practical possibilities. It is a very important system to our society, and like most systems its performance can best be improved by considering the whole as set of interacting components rather than by considering the components in isolation."(p.1f)

He mentions that this kind of a system approach, which is needed to analyze and thereby supporting to improve 'human intellectual effectiveness', can not be found as a ready-made conceptual framework such as it is the case for established disciplines. (cf. [Eng62]:p.2)

He thinks "that there is no particular reason not to expect gains in personal intellectual effectiveness from [a] concerted system-oriented approach that compares to those

---

[5] See for this the introduction of Johnson-Laird (1980) [JL80]

made in personal geographic mobility since horseback and sailboat days."(cf. [Eng62]:p.2f)

What has to be find are (i) "the factors that limit the effectiveness of the individuals basic information handling capabilities" ... " and (ii) to develop new techniques, procedures, and systems that will better match these basic capabilities to the needs, problems, and progress of society."(cf. [Eng62]:p.6)

For this analysis to do Engelbart assumes a basic system layout as follows: " The entire effect of an individual on the world stems essentially from what he can transmit to the world through his limited motor channels. This in turn is based on information received from the outside world through limited sensory channels; on information, drives, and needs generated within him and on his processing of that information. His processing is of two kinds: that which he is generally *conscious* (of recognizing patterns remembering visualizing abstracting deducing inducing etc.), and that involving the *unconscious* processing and mediating of received and self-generated information, and the *unconscious* mediating of *conscious* processing itself."(cf. [Eng62]:p.8)

Thus, he assumes an input-output system filled up with processes, where Engelbart distinguishes between some 'conscious' processes and all the other 'unconscious' processes. Interesting is his assumption that the 'conscious' processes are presupposing 'unconscious' processes for their conscious functioning.

Furthermore does Engelbart assume that there is no direct interaction between the real outside situations and some complex background knowledge. There are some 'intermediate' processes necessary. He cites the example of "an aborigine who possesses all of our basic sensory-mental-motor capabilities, but does not possess our background of indirect knowledge and procedure, cannot organize the proper direct actions necessary to drive a car through traffic, request a book from the library, call a committee meeting to discuss a tentative plan, call someone on the telephone or compose a letter on the typewriter."(cf. [Eng62]:p.8)

These ways in which human capabilities are thus extended does Engelbart call 'augmentation-means' and he distinguishes "four basic classes of them: *Artifacts*, *Language Methodology* , and *Training*.(cf. [Eng62]:p.9)

Important seems to be his assumption that all these augmentation means, which are realized in processes, consist always of simple steps, which can be combined to larger and larger units. "... human beings nevertheless do solve complex problems. It is the augmentation means that serve to break down large problem in such a way that the human being can walk through it with his little steps, and it is the structure or organization of these little steps or actions that we discuss as *process hierarchies*." (cf. [Eng62]:p.10) For Engelbart is this repertoire of processes organized as a 'hierarchical'.(cf. [Eng62]:p.11)

An interesting point is Engelbarts assumption, that the processing of all these augmentation means is managed by the '*execution* capability.' Such an "executive process (i.e. the exercise of an executive capability) involves such sub-processes as planning, selecting, and supervising, and it is really the executive processes that embody all of the methodology in the H-LAM/T (:= Human using Language, Artifacts, Methodology in which he is Trained) system."(cf. [Eng62]:p.12)

Although Engelbart does not discuss explicitly the definition and measuring of intelligence he seems to assume that there exists observable behavior which can be classified as 'intelligent behavior' and therefor can arise the question "where that intelligence is embodied?". And Engelbart concludes that "we are forced to concede that [the intelligence] is elusively distributed throughout a hierarchy of functional processes... whose foundation extends down into natural processes below the depth of our comprehension." From this follows that 'intelligence' seems to be embedded in this '*organization*'. For this the term '*synergism*' is in use. And Engelbart states that "This term seems directly applicable here ... synergism is our most likely candidate for representing the actual source of intelligence." (cf. [Eng62]:p.18)

Following Ross Ashby who coined the term 'intelligence amplifier' Engelbart thinks that this term is applicable to "the goal of augmenting the human intellect in that the entity to be produced will exhibit more of what can be called intelligence than an unaided human could;..."(cf. [Eng62]:p.18) And Engelbart continues "What we have done in the development of our augmentation means is to construct a superstructure that is a synthetic extension of the natural structure upon which it is built. In a very real sense, as represented by the steady evolution of our augmentation means, the development of artificial intelligence has been going on for centuries."(cf. [Eng62]:p.19)

Using this new concept of an 'intelligence amplifier' Engelbart constructs a "historical progression in the development of our intellectual capabilities" with the following levels of complexity: " (i)*Concept Manipulation* ... capability for developing abstractions and concepts.... (ii) *Symbol Manipulation* Humans... learned to represent particular concepts in their minds with specific symbols. ... (iii) *Manual, External, Symbol manipulation* ... means for externalizing some of the symbol-manipulation activity, particularly in graphical representation." And he is considering "if language is (as it seems to be) part of a self-organizing system then it seems probable that the state of language at a given time strongly affects its own evolution to succeeding state."(cf. [Eng62]:p.23f) And he suggests a "fourth stage to the evolution of our individual human intellectual capability: (iv) *Automated external symbol manipulation*", which could be enabled by computer "with which we could communicate rapidly and easily, coupled to a three-dimensional color display within which it could construct extremely sophisticated images."(cf. [Eng62]:p.25)

In a final metaphor Engelbart compares the repertoire hierarchy of process capabilities with a "executive superstructure by considering it as though it were a network of contractors and subcontractors in which each capability in the repertoire hierarchy is represented by an independent contractor whose mode of operation is to do the planning, make up specifications, subcontract the actual work, and supervise the performance of his sub contractors. This means that each subcontractor does the same thing in his turn. At the bottom of this hierarchy are those independent contractors who do actual 'production work'. ... We can readily recognize that there are many ways to organize and manage such a superstructure resulting in vastly different degrees of efficiency in the application of the worker's talent."(cf. [Eng62]:p.42)[6]

Summing up these ideas one can perhaps say that Engelbart constructed a 'hypothetical machinery' located in a human actor which gives some rationale behind the manifold and perplexing behavior of a human actor. Although Engelbart identified some types of behavior like 'concept manipulation', 'symbol manipulation' etc. it stays open how exactly these different behaviors are 'matching' with this machinery. His machinery is like a 'working metaphor' to inspire some more analysis and more experiments.

With regard to Norman – who published later – one has to say that this machinery of Engelbart shows no direct relationship to the different kinds of 'models' mentioned by Norman. The 'conceptual model (mental model)' of the human actor about the system to interact with is not necessarily identical with the 'machinery'of the system; the same holds for the 'user model' of a potential intelligent machine interacting with a human actor as user. Only the 'design model' (including the 'system image') of the designer could eventually include the machinery of the system, but methodological the design model should not deal with the details of the system.

5.3 Johnson-Laird - 1980/2010

For the following discussion I am using the papers from 1980 (1980) [JL80] as well as from 2010 [JL10].

The first extensive paper from 1980 did Johnson-Laird write while he was a Professor for Experimental Psychology at the University of Sussex (England).

After analyzing many interesting topics he is coming to the conclusion, that an appropriate version of 'Cognitive science' does not yet exist! It is necessary to invent it. He sees some deficiencies on the side of the psychological experimenter who "exerts a dangerous pull in the direction of empirical pedantry, where the only things that count are facts, no matter how limited their purview" and on the other side the computer science programmer, who "exerts a dangerous pull in the direction of systematic delusion, where all that counts is internal consistency, no matter how remote it is from reality."(cf. [JL80]:p.110) What Johnson-Laird is calling for is "to develop general and comprehensive theories of the mind, couched in the theoretical vernacular of the discipline; to make explicit models of at least parts of them in the form of computer programs; and to combine this process with a regime of experimental investigation. This route may lead us to a discipline that is a general science of the mind."(cf. [JL80]:p.110) And one has to generalize this claim with Johnson-Laird in saying "If we are ever to understand cognition, then we need a new science dedicated to that aim and based only in part on its contributing disciplines."(cf. [JL80]:p.71)

Behind this claim stands the experience of different "cases where a particular problem or concept has been a focus for work in a number of different disciplines." All these cases "show an increasing overlap in the research carried out in different academic departments." And to integrate these different points of view in a coherent way it deserves a methodology which serves these needs.(cf. [JL80]:p.72) Johnson-Laird points to the different 'truth-theories' of Psychology and an Artificial Intelligence Discipline: "*Psychologists* want their theories to *correspond* to the facts; *artificial intelligencers* want their theories to be *coherent*; both groups have adopted the methods best suited to their aims. *Cognitive science*, however, needs theories that both cohere and correspond to the facts. Hence a rapprochement is required."(cf. [JL80]:p.73)

Johnson-Laird continues then with three topics which are intended to illustrate his diagnosis and his thesis. All three topics "implicate the notion of a mental model" which is understood as an "internal model of the world inside an organism." The main questions here are "the way in which [the mental models] are mentally represented and the use to which they are put in cognition." (cf. [JL80]:p.73)

After discussing logical inference in the classic syllogistic Aristotelian logic as well as inference in modern logic Johnson-Lairds draws some conclusions about properties of mental models, e.g. "... mental models can obviously be generated so as to represent all sorts of quantified assertions. They accommodate multiply-quantified assertions ... which cannot be represented by Euler circles. They can even represent sentences that are claimed to demand 'branching' quantifiers that go beyond the resources of the ordinary predicate calculus,... They can accommodate such quantifiers as most, many, several and few. They enable distinctions to be drawn between each and every, and any and all,... Models also allow a clear distinction to be drawn between class-inclusion and class-membership. "(cf. [JL80]:p.82) And he continues: "The psychological theory posits a process of inference that involves, not the mobilization of quasi-syntactic rules of inference, but the direct manipulation of a model of the assertions in the premises."(cf. [JL80]:p.82)

These general remarks do not explain the details of a psychological model of 'mental models'. The statement of Johnson-Laird, that "A computer program that I have devised

---

[6] Reading this text from Engelbart I am somehow reminded to the vision of Minsky (1985) [Min86] in his thought experiment 'Society of Minds'. It seems that Minsky did not know this text from Engelbart.)

works according to the theory and uses no rules of inference. Its power resides in the procedures for constructing and manipulating models – a power which in turn demands at the very least the recursive power of list-processing operations"(cf. [JL80]:p.84) does not answer the question. Thus we have observations of certain behavior – here logical inferences – and then there are some hypotheses about a 'machinery' called 'mental model' which shall in some way 'explain' the observable behavior. This explanation seems to be missing.

Johnson-Laird continues with another example taking from the realm of 'Meaning', the semantics of language expressions. He circumscribes the problem as "a major burden for the meaning of words is to account for the relation between such assertions as 'Polly is a parrot' and 'Polly is a bird' – if the first assertion is true, then plainly so is the second." What "is the nature of the semantic machinery needed to explain such relations. "(cf. [JL80]:p.85)

Johnson-Laird cites as one position "that the meaning of a word such as 'parrot' is represented in the *mental lexicon* as a set of semantic elements" including the decomposition of a sentence "into semantic primitives".(cf. [JL80]:p.85) In another cited position there are "*meaning postulates*"... which "stipulate the semantic relations between words, e.g. for any x. if x is a parrot then x is a bird." This hypothesis is associated with another one telling that "sentences in a natural language are translated into '*propositional representations*' in a corresponding *mental language*, and that meaning postulates couched in the mental vocabulary are used to make inferences from these propositional representations."(cf. [JL80]:p.86)

The hypothesis with the 'meaning postulates' is ruled out according to Johnson-Laird because "human beings do not have an unlimited capacity for storing information, or the ability to learn an infinite number of rules."(cf. [JL80]:p.88) And after mentioning further arguments against 'meaning-postulates' Johnson-Laird argues for the assumption, that "the only way to account for the proper relations between words, and for inferences based upon them, is by giving a specification of their meanings that includes their relations to the world." And he continuous: "the[se] relations are so basic that there is no way to define them in ordinary English. It is for this reason that a complete theory of meaning must rely upon some sort of decomposition into more primitive notions."(cf. [JL80]:p.89)

From this it seems that the theoretical concept of a 'mental model' is somehow 'touching' basic {elements/ events/ processes} of the brain machinery which serve as the building blocks of a possible 'meaning'.

Finally Johnson-Laird is looking to the case of 'images'. He assumes that "No one seriously doubts the existence of the psychological phenomena of imagery. What is problematical, however, is the explanation of the phenomena and the ultimate nature of images as mental representations. It seems unlikely that they are simple pictures in the head, because this conjecture leads to a number of undesirable consequences including the need for an homunculus to perceive the pictures, and thus to the danger of an infinite regress."(cf. [JL80]:p.91f)

Johnson-Laird distinguishes "two schools of thought. On the one hand, there are those who argue that an image is distinct from a mere representation of propositions." ... "On the other hand, there are theorists who argue that the subjective experience of an image is epiphenomenal and that its underlying representation is propositional in form." (cf. [JL80]:p.92)

After many pages of considerations about the relationship of 'mental models' and 'propositions', Johnson-Lairds draws the following conclusion: "Mental models and propositional representations can be distinguished on a number of criteria. They differ preeminently in their function: a *propositional representation* is a description. A description is true or false, ultimately with respect to the world. But human beings do *not* apprehend the world directly: they possess only internal representations of it. Hence, a propositional representation is *true or false with respect to a mental model of the world*. In principle, this *functional difference* between models and propositions could be the only distinction between them: there need be nothing to distinguish them in form or content.... A model represents a state of affairs and accordingly its structure is not arbitrary like that of a propositional representation, but plays a direct representational or analogical role. Its structure mirrors the relevant aspects of the corresponding state of affairs in the world. "(cf. [JL80]:p.98)

And this difference between a 'concrete form' and of a 'general form' is a property shared between 'images' and 'models': "Images, like models, have the property of arbitrariness, which has often drawn comment from philosophers. You cannot form an image of a *triangle in general*, but only of a specific triangle. Hence, if you reason on the basis of a model or image, you must take pains to ensure that your conclusion goes beyond the specific instance you considered."(cf. [JL80]:p.98) Therefore it "follows that images correspond to those components of models that are directly perceptible in the equivalent real-world objects. Conversely, models may underlie thought processes without necessarily emerging into consciousness in the form of images. Models are also likely to underlie the perception of objects by providing prototypical information about them in a form that can be directly used in the interpretation of the output of lower level visual processes"(cf. [JL80]:p.100)

At this point the question arises, how far an intended theory of cognitive science has to go 'inside' a hypothesized system? Which 'level of detail' has to be assumed for the 'semantics' of the used theoretical terms? (A question which is virulent for the ideas of Engelbart too). Johnson-Laird does exclude that cognitive science has to deal with the neuronal activity in case of the brain or with the machine code in the case of the computer.(cf. [JL80]:p.100f) But this exclusions does not explain how to proceed. Johnson-Laird cites ideas from the formal semantics of computer languages where a compiler translates a high-level language into some machine code. For a programmer it is not important which kind of machine code will be generated; the programmer thinks in the 'meaning categories' of the high-level language.(cf. [JL80]:p.100) But the question is then, to what corresponds these 'meaning categories' in the case of the cognitive scientist? Do these meaning categories correspond to the 'mental model', and if so, what dies this mean?

Johnson-Laird continues to discuss this question with an example of programming a computer with spatial objects: "A programmer needs to know no more: one can write procedures for manipulating arrays simply by thinking of them as n-dimensional spaces where each location is specified by an n-tuple of integers. A student of the "psychology" of computers, however, may be curious about the invisible machinery that makes such an array possible. Its representation in the computer does not involve an actual physical array of locations in core store. That is quite unnecessary. Indeed, the physical embodiment of an array is irrelevant. What matters is that it should function as an array, that is, it has a set of addresses that are functionally equivalent to an array, its elements can be accessed as in an array, and its contents displayed or printed out in the form of an array. A psychological description should accordingly be a functional one."(cf. [JL80]:p.101) Thus "The program functions as though it uses an array, and one seen from a particular viewpoint, too."(cf. [JL80]:p.102)

This example again allows the distinction between some 'presupposed lower level' and some language expressions having a meaning on a more 'abstract level'. But what can we say about this more abstract level? In the following paragraphs Johnson-Laird points to a 'solution' which works 'somehow', but does not answer all questions.

He writes: "In general, a model is only a model at a certain level of description: that level at which it functions as one. A listing of the original spatial inference program in machine code is a level of description that obscures the program's use of models. ... There is, of course, nothing inconsistent about calling such a representation [= the program] a propositional theory. Indeed, the controversy can be resolved in a still more direct way to support the view that any plausible theory of any psychological phenomenon is propositional. If you accept Church's thesis that any "effective procedure" can be computed by a Turing machine, then it follows that the psychological theory, granted the reasonable criterion that it is intended to characterize an effective procedure, can also be computed by a Turing machine. This device, however, can be completely described by a set of propositions- linear strings of symbols from a defined alphabet that characterize the rules governing its change of state and behaviour as a function of its current state and input."(cf. [JL80]:p.102)

But this is not the answer we have asked for. The point is not, that every kind of a theory is 'propositional', but what are these 'meaning categories' which constitute the possible 'truth' for the expressions?.

In a final summary Johnson-Laird states: "the real context of an utterance consists of the mental models of the current conversation that the speaker and the listener maintain. These models represent the relevant individuals, events, and relations. They also represent what is known about the other participants' state of mind. Hence, a speaker chooses his words partly on the basis of his model of the listener's discourse model; and a listener interprets these remarks partly on the basis of his model of the speaker's discourse model. A number of referential phenomena depend critically on the characteristics of mental models ... For example, what really controls the use of a definite description is, not uniqueness in the world, but uniqueness in a model. Likewise, the most important characteristic underlying the coherence of texts is continuity of reference – a feature that was explicitly manipulated in the experiments on spatial inference. There are of course other aspects of coherence, but none is likely to be so preeminent as referential continuity: if a text never refers to the same entity more than once, it rapidly acquires the characteristics of a telephone directory rather than a passage of prose. Mental models evidently play a part in a variety of phenomena other than those that I have considered in detail in this paper. They appear to have a unifying role to play in Cognitive Science."(cf. [JL80]:p.106f)

After these very interesting ideas of Johnson-Laird supported by many experiments, which have not being reported here, the hypothesis of 'mental models' is on the table; additional some requirements of 'propositional models' which get their truth values from these models. But it is further unclear how one can or should 'define' these models and their interplay in a full cognitive science theory.

If we are looking to the paper from 2010 – 30 years later! – it seems that the theory of Johnson-Laird did not evolve to new insights. The paper repeats more or less the same topics with more or less the same wording. Look to his conclusion: "Human reasoning is not simple, neat, and impeccable. It is not akin to a proof in logic. Instead, it draws no clear distinction between deduction, induction, and abduction, because it tends to exploit what we know. Reasoning is more a simulation of the world fleshed out with all our relevant knowledge than a formal manipulation of the logical skeletons of sentences. We build mental models, which represent distinct possibilities, or that unfold in time in a kinematic sequence, and we base our conclusions on them. When we make decisions, we use heuristics and some psychologists have argued that we can make better decisions when we rely more on intuition than on deliberation. In reasoning, our intuitions make no use of working memory and yield a single model. They too can be rapid – many of the inferences discussed in this article take no more than a second or two. However, intuition is not always enough for rationality: a single mental model may be the wrong one."([JL10]:p.18249)

5.4 Kieras et.al - 1984

The paper of Kieras and Bovair (1984) [KB84] about ’Mental Models’ can be distinguished from Johnson-Laird’s paper because they try to reduce the concept ’Mental Model’ to that internal structure, which corresponds to that system, with which the human actor as user is directly interacting. Therefore they propose to call this internal model ’device model’. Whether this distinction compared to Johnson-Laird (1980) helps to gain more clarity is an open question. They do not really introduce more information than Johnson-Laird with his experiments.

In their three reported experiments they are concerned ”with learning how to operate a simple control panel device, and how this learning is effected by understanding a device model that describes the internal mechanism of the device.”(cf. [KB84]:p.255) In their first experiment they compare two groups, one of which is learning a set of operating procedures for the device by rote, and the other group is learning a device model before they receive the identical procedure training.

Already at this point one can ask whether this description is not somehow ’deceptive’ in that sense that what here is called a ’device model’ is not the ’inner’ device model which is perhaps some ’internal structure’ in the test person, but a ’real’ device model (’operate an unfamiliar piece of equipment’) which serves as a reference object to generate a possible ’internal’ device model.

When Kieras et.al then report from the first experiment, that the group trained with the real device model learned the procedure faster, retained them more accurately, executed them faster, and simplified inefficient procedures more often than the rote group, then this correlates a different behavioral input with a different behavioral output and an ’unknown internal factor’, which perhaps has become ’activated’ by the different behavioral context.

From the second experiment do Kieras et.al report, that the group trained with the real device model is able to infer the procedures much more easily than the rote group, which would lead to more better learning and better recall performance. (cf. [KB84]:p.255) It is here not clear what exactly the difference is to the first experiment.

From the third experiment is reported, that the important content of the real device model was the specific configuration of components and controls, and not the motivational aspects, component descriptions, or general principles. The used concepts are not very clear. When Kieras and Bovair state, that ”this specific information is what is logically required to infer the procedures.”, then one has to clarify what means ’logically required’.(cf. [KB84]:p.255)

Besides the ’first confusion’ in the usage of the term ’device model’ ((i) ’the real thing in front of the human actor,or (ii) some ’internal structure in the human actor’) there seems to be a ’second confusion’. Kieras and Bovair introduce the example of ’the modern telephone system’ which is ”extremely complex, but the typical telephone book contains only ’how-to-do-it’ instructions; very few people know how the system works beyond the crudest principles. However, almost everybody can successfully operate a telephone. Detailed knowledge of how the system works seems to be irrelevant.”(cf. [KB84]:p.256)

Here one can really ask whether ’device model’ has to be identified with a model of the ’internal system’ of the real system in front of the user, which in most cases never is known. This is a reason why Norman speaks of the ’system image’.(cf. [ND86]:p.46) From this we gain at least three different ’readings’ of the term ’device model’: (i) The inner structure of some real system; (ii) the system image of the real system, or (iii) the inner structure in the user dealing with the real system.

As the series of the three experiments showed (according to the authors), it was only the ’general version’ of the device model, which was ’psychological significant’. We look therefore directly to this kind of a ’device model’.

In their third experiment Kieras and Bovair state, that ”the critical how-it-works information is the specific descriptions of the controls and their path relations to the internal components. Therefore, neither the fantasy context, nor details about the nature of the components, nor general principles about how the system works, should be of value in enabling subjects to infer the procedures. This set of assertions was tested in Experiment 3...” (cf. [KB84]:p.266f)

Inferring from the context it seems, that the Kieras and Bovair mean with ’device model’ an ’internal model’ of the real system. But as the three experiments and the last lines reveal, the term ’(inner) device model’ as such is also not very clearly defined. Otherwise it would be difficult to understand why the experimenters used different kinds of descriptions under the label of ’device model’. It is really not clear why ’the nature of the components’ should not belong to an ’(inner) device model’, and what ’specific descriptions of the controls’ should mean.

The device model materials have been presented in the third experiment without any discussion of how the system components worked or why they were present. In the paper no image of this third device model is shown. The notice, that ”The diagram is shown in Figure 3” is wrong.

The experimenters conclude, that ”These results [in experiment 3] show that the effectiveness of the device model instructions ... was ... due to ... the critical how-it-works information [of] the specific items of system topology that relate the controls to the components and to the possible paths of power flow.”(cf. [KB84]:p.271) And the experimenters continue with the conviction, that ”If this definition of what constitutes a useful device model is adopted, several practical suggestions can be offered for when, and what kind of, device model information should be taught to users of a device:

1. The device model information must *support inferences*

about the exact and specific *control actions*...
2. The relevant *how-it-works knowledge* can be very superficial and incomplete, because *the user does not need to have a full understanding of the system in order to be able to infer the procedures for operating it*.
3. Teaching a device model will not always be of value; *it depends on whether the user in the actual task situation both needs to infer the procedures, and also needs the supplied information* in order to be able to infer the procedures. ...
4. Learning and using a device model may have its own pitfalls. That is, *knowledge of the model may be subject to misunderstandings and distortions*, like any other knowledge. ..."(cf. [KB84]:p.272))

This final statements can cause some confusions:

Ad 1) From the text of the paper it is not clearly defined what is meant with 'inferences' and what are 'control actions'.
Ad 2) The term 'how-it-works knowledge' is confusing too. This term 'how-it-works knowledge' reminds inner states of the user which in the preceding passages have been excluded. Furthermore in the preceding paragraphs the experimenters have talked about a '(real) design model' which usually is a 'real object' and not a 'description'. And, how can a user 'infer the right procedures' without having a certain kind of knowledge which enables such inferences? If one assumes – like Norman – a 'system image' which is given without a '(real) device model' of the inner structure of the system then this could be plausible, but this has been excluded by Kieras and Bovair.
Ad 4) If a 'device model' can be understood as 'wrong' how should this be tested in the experiments to exclude some 'noise' from the measurement?

Summing up, the text of Kieras and Bovair shows some interesting points in the discussion of the topic 'mental model'. Nevertheless some questions do arise: (i) Even the talk of a 'design model' calls for distinctions ('real inner states', 'real image states', 'inner states'); (ii) the talk about 'inferences' is fuzzy. What exactly shall this be?, (iii) How is the relation between a 'real design model as a real object' and some kind of a 'description' of this model? Is the description only a 'text' with some undefined meaning or a 'text' with a 'certain meaning'? What then is this 'meaning'? (iv) Under which circumstances can one say that a user 'needs' information, and which kind of 'information'? How can one 'measure' this need?

5.5 Rasmussen - 1987/90

I am following the text of Rasmussen which has been published as part of a book in 1990 [Ras90]. This text is a 1-t-1 copy of the text of a report published 1987 [Ras].[7] In the citations I am using the page numbering of the book.

In this paper Rasmussen wants to discuss the concept 'mental model' "as seen from the point of view of analysis and design of interfaces between humans and their work based on advanced information technology. What is the nature of humans' conception of their work content, and how can computer-based information systems be made transparent and support the proper mental models?"(cf. [Ras90]:p.41)

Rasmussen uses the concept of 'mental model' different to Kieras and Bovair not in the sense of a '(real) design model' but rather in the sense of Johnson-Laird as something 'internal' to the human actor. He writes, it "is used to characterize features of the resident *knowledge base*, *representing* properties of the task environment which can serve the *planning* of activities and the *control* of acts when instantiated and *activated by observation of the actual state of affairs.*(cf. [Ras90]:p.42)

It seems that Rasmussen makes a distinction between the 'actual state of affairs' which point to some 'real situation'; these states of affairs can by observations 'activate' some 'properties' of an internal 'knowledge base'. The set of these properties of the internal knowledge base is associated with the term 'mental model'. If there is 'outside', in the 'real situation' some observable structure which is associated with a system to interact with then the 'internal model' represents properties of this real system, which can be used for 'planning'.

From a discussion of Craik [Cra43] Rasmussen infers that "for the representation of human knowledge in a complex working context more than one 'mental model' should be considered."(cf. [Ras90]:p.42)

In the public discussion Rasmussen identifies "two different points of view" for the term 'mental model': (i) "Mental models are the bridge between the *work environment to be controlled* and the *mental processes underlying this control*. Consequently, a study can be approached by a study of human mental processes as well as by a study of work requirements, and these approaches result in different concepts." (ii) "The approach from the *psychological point of view* quite naturally focuses on the explanation of human performance, which often will be influenced by the AI related cognitive science. The focus of this research will be on the nature and form of the mental model together with its role in human reasoning and its relations to the 'mind'. Consequently, the criterion of success will often be whether a theory can be phrased explicitly in procedural form for simulation on computer."(cf. [Ras90]:p.43)

From a methodological point of view it is difficult to see how these two positions can be distinguished. Psychology is looking to the behavior which is embedded between the work environment and the acting human. And doing its job in the right way psychology will set up models of the inner structures of the acting human which 'explain' the behavior in that specific environment. Part of these models can be 'mental models'. Somehow Rasmussen seems to agree with this not existing difference when he states: "The two approaches are supplementary rather than competing and interaction between them is important for the development

---

[7] The report can be found online.

of modern information technology."(cf. [Ras90]:p.43) But his term 'modern information technology' has no clear meaning in this context.

Rasmussen mentions further the 'Cognitive Science Approach' without introducing its methodological position exactly, but associates 'cognitive science' with the position of Johnson-Laird.[8] Rasmussen cites Johnson-Laird in the context of 'mental representations' and cites the three types of mental representations found with Johnson-Laird: "*propositional* representations which are strings of symbols that correspond to natural language, *mental* models which are structural analogues of the world, and *images* which are perceptual correlates of models from a particular point of view."

Thus it seems that the term 'mental representation' is the more general term and 'mental model' is only one variety of a 'mental representation'. Furthermore Rasmussen seems to assume that Cognitive Science is dedicated only to analyze and model 'mental representations'. But if Cognitive Science is not completely apart from Psychology than this view is too restrictive because Psychology can generate their models only based on the observable behavior and this includes the work environment.

Somehow Rasmussen seems to underline this position, when he under the label of 'Cognitive Engineering' stresses that "the interaction of performance under control of basically different kinds of mental representations has to be considered." And "Several different research models of the various mental representations have to be accepted ..., and validation of the models will basically be a test of their predictive ability for systems design, i.e., validation depends to a large degree on evaluation of system during actual work conditions. In order to have a framework for mapping the properties of different kinds of mental representations, a discussion of the cognitive control of skilled work performance will be useful.(cf. [Ras90]:p.46f)

Therefore, for Rasmussen "it is necessary to study the interaction of a wide variety of mental strategies and models. In particular, study of the interaction and interference between different modes of cognitive control appear to be important for the understanding of erroneous performance.(cf. [Ras90]:p.47)

And Rasmussen distinguishes then three typical levels of performance: skill- , rule-, and knowledge-based performance:

"*Skill-based* behavior represents sensori-motor performance during acts or activities that, after a statement of an intention, take place without conscious control as smooth, automated, and highly integrated patterns of behavior. .. At the next level of *rule-based* behavior, the composition of a sequence of subroutines in a familiar work situation is typically consciously controlled by a stored rule or procedure that may have been derived empirically during previous occasions, communicated from other persons' know-how as an instruction or a cookbook recipe, or it may be prepared on occasion by conscious problem solving and planning. The point is here that performance is goal-oriented, but structured by "feed-forward control" through a stored rule, in other words, the person is aware that alternative actions are possible and has to make a choice. The choice is based on 'signs' in the environment which have been found to be correlated to one of the alternative actions. Very often, the goal is not even explicitly formulated, but is found implicitly in the situation releasing the stored rules. The control is teleological in the sense that the rule or control is selected from previous successful experiences. The control evolves by "survival of the fittest" rule.(p.49) ... "In general, skill-based performance rolls along without conscious attention, and the actor will be unable to describe the information used to act. The higher-level rule-based co-ordination in general is based on *explicit know-how*, and the rules used can be reported by the person, although the cues releasing a rule may not be explicitly known. During unfamiliar situations for which no know-how or rules for control are available from previous encounters, the control must move to a *higher conceptual level*, in which performance is *goal-controlled*, and *knowledge-based* (knowledge is here taken in a rather restricted sense as *possession of a conceptual, structural model* or, in AI terminology, of deep knowledge. The level, therefore, might also be called *model-based*'). In this situation, the goal is explicitly formulated, based on an analysis of the environment and the overall aims of the person. Then a useful plan is developed - by selection. Different plans are considered and their effect tested against the goal, physically by trial and error, or conceptually by means of 'thought experiments'. At this level of *functional reasoning*, the *internal structure of the system* is explicitly *represented by a "mental model"* that may take several different forms. A very important aspect of the cognitive control to be captured by models of human behavior is the *dynamic interaction between the activities at the three levels*. (cf. [Ras90]:p.49)

In the preceding paragraph I have additional 'highlighted' some expressions to point to the implicit structure of the text. While the sections with 'skill-based' and 'rule-based' behavior can simply be separated is the text centering around the 'knowledge-based' performance enriched with some additional terms which are not yet readily introduced and defined. The 'explanation' of 'knowledge-based' by the expression 'possession of a conceptual, structural model' does not really explain something because these terms as such are also not explained, neither is the term 'mental model' of any help here. Further it is not really clear what is meant with a 'higher conceptual level' no what has to be understood with the term 'goal controlled'.

After introducing the basic terms Rasmussen talks about a certain performance context called 'problem solving'. He describes this as follows: "*Problem solving* takes place when the reaction of the environment to possible human actions is *not known from prior experience*, but *must be deduced* by means of a *mental representation of the 'relational structure' of the environment*. This structure must be *represented symbolically* in a mental model. A major

---

[8] See text above.

task in knowledge-based problem solving is to transfer those properties of the environment which are related to the perceived problem to a proper symbolic representation. The information observed in the environment is then *perceived as symbols, with reference to this mental model.*(p.50)

The situation seems to be one in which the actor has reacted to the environment, and this environment will be perceived in a way, that a mental representation in the actor will be activated which represents some of the relational structure from this environment. Parts of this structure can be translated in a way that these receive some meaning from the corresponding mental model. This availability of a meaning enables some of the perceived structures to function as 'symbols' in the sense of Semiotics. Within this context Rasmussen claims that the representation of the environment can be 'different' from what can be 'expected'. In that case other human actions shall be 'deduced'. How can this be done? The whole machinery is not very clear.

Rasmussen sees a mutual interrelation between 'knowledge about the basic laws governing the behavior of the environment' and the 'formation of a proper representation'. When such a representation is obtained – which means a mental model – then this representation turns into a prescriptive *system of signs* that control the application of stereotyped process rules.... The efficiency of formal, mathematical models and technical graphs and diagrams, as e.g., control engineers Bode plots and pole-zero graphs, depends on the existence of a large *repertoire of stereotyped manipulation rules* used for solutions and predictions...The conclusion of this discussion is that patterns in a symbolic model configuration, as is the case with perceptual patterns of the physical environment, can act as signs. (cf. [Ras90]:p.50)

Although the details of this machinery allow many questions it emerges an interesting idea insofar Rasmussen uses the mental machinery as a framework to motivate that and how in this framework elements of this machinery can function as 'signs' bounded by 'meaning' which is constructed as part of the machinery. As far as all actors have such a representational machinery with embedded sign-structures the exchange of actions including externalized sign-elements can trigger the meaning structures grounded in the mental representations of the other actors. This is seems to be a very strong position. For the semiotic part Rasmussen mentions explicitly Morris (1971) [Mor71] and Eco (1979) [Mor79].

Although Rasmussen continues in this paper with many more exciting ideas we stop here the discussion. The reason is that the details of this machinery are for precise measurements and hard discussions too fuzzy. Nevertheless his approach is impressive and very stimulating. He shows – together with the work of Johnson-Laird – in which direction future work of a more theory based Psychology has to go in close connection with computer science as an important modeling tool, which after theoretical success can probably contribute with many exciting new applications.

What is becoming more clearer now too is the great opportunity which is provided by the new Actor-Actor Intraction paradigm embedded in Systems Engineering and including Artificial Intelligence Methods.

5.6 Cañas et.al. 2001

We will not discuss the paper 'The role of working memory on measuring mental models of physical systems' of Cañas, Antolí, and Quesada (2001) [CnAQ] in detail. The only aspect which is special to these authors is the explicit inclusion of the psychological theoretical constructs of 'Short Term Memory (STM)' as well as 'Long Term Memory (LTM)'. Instead of speaking in a broad – and mostly a bit 'fuzzy' – manner of 'mental representations' or 'mental models' they primarily refer to the theoretical memory concepts of psychology. Assuming such a framework they distinguish two meanings of the term 'Mental Model': "For some researchers, a Mental Model is a representation stored in Working Memory, while for others it is the knowledge stored in Long Term Memory.(cf. [CnAQ]:p.26)

The rest of the paper is problematic. The authors take nowhere a clear position what exactly is their position. Furthermore most concepts are not clearly defined, even the main terms of their paper 'short-term' as well as 'long-term memory' are not discussed with all the results and questions which are known in the literature.

6 CONCLUSION (PRELIMINARY)

Because of the draft character of this text the conclusions here are preliminary.

The main idea was to look back in the history of Human Computer Interaction to clarify in which sense the actual paradigm of Actor-Actor Interaction as part of Systems Engineering including also the Artificial Intelligence paradigm makes sense, or not.

Although the selected texts so far are only a small subset of a much bigger set of historically relevant papers it can be summarized already that the main road paved by the many always impressive HCI experts is in close agreement with the visions of our project.

Indeed, what we are interested in is the development and inclusion of the computational theories and technologies to improve the life of man – and all of biological life ! – on the planet earth as well as in the whole universe as much as possible. As it becomes more and more clear today the new computational technologies are substantially lacking something which biological life, especially we, the homo sapiens, have, and vice versa these new technologies have something, we as humans have not. This points in the direction of a new, deeper and more interesting symbiosis of biological life and machines. Although in 2018 some interesting new insights could be reached with artificial intelligence, the existing algorithms are far below that level which really could be called 'a learning intelligence'. At the same time we have to accept that the knowledge of humans about humans is highly un-developed.

There is some motivation to elaborate this HCI-Review project to improve and to extend in the future. At the same time the AAI-in-SE-with-AI project has to be developed much further.

REFERENCES

[BS11] Gordon Baxter and Ian Sommerville. Socio-technical systems: From design methods to systems engineering. *Interacting with Computers*, 23(1):4–17, 2011.

[BUS45] VANNEVAR BUSH. As we may think. pages 1–19. THE ATLANTIC MONTHLY, Washington, D.C., 1945.

[CFP+] W.A. Clark, J.M. Frankovich, H.P. Peterson, J.W. Forgie, R.L. Best, and K.H. Olsen. The lincoln tx-2 computer, October 1.

[CMN83] Stuart K. Card, Thomas P. Moran, and Allen Newell. *The Psychology of Human-Computer Interaction*. Lawrence Erlbaum Associates, Inc., Mahwah (NJ), 1 edition, 1983.

[CnAQ] José J. Cañas, Adoración Antolí, and José F. Quesada. The role of working memory on measuring mental models of physical systems. *Psicológica*, 2001:25–42.

[Cra43] Kenneth Craik. *The Nature of Explanation*. Cambridge University Press, Cambridge (UK), 1 edition, 1943.

[Eng62] D.C. Engelbart. Augmenting human intellect: A conceptual framework. October 1962. Stanford Research Institute,Technical Report, No.3578.

[Eng63] Duglas C. Engelbart. A conceptual framework for the augmentation of man's intellect. pages 1–29. Spartan Books, Washington, D.C., 1963.

[Gru08] Jonathan Grudin. A Moving Target: The Evolution of HCI. In A. Sears and J.A. Jacko, editors, *The Human-Computer Interaction Handbook. Fundamentals, Evolving Technologies, and emerging Applications*. 2 edition, 2008.

[Hen86] D. A. Henderson, Jr. The trillium user interface design environment. *SIGCHI Bull.*, 17(4):221–227, April 1986.

[Hop81] Gracy Murray Hopper. History of programming languages i. chapter Keynote Address, pages 7–20. ACM, New York, NY, USA, 1981.

[JL80] P.N. Johnson-Laird. Mental models in cognitive science. *Cognitive Science*, 4(1):71–115, 1980.

[JL10] Philip N. Johnson-Laird. Mental models and human reasoning. *Proceedings of the National Academy of Sciences*, 107(43):18243–18250, 2010.

[KB84] David E. Kieras and Susan Bovair. The role of a mental model in learning to operate a device. *Cognitive Science*, 8(3):255–273, 1984.

[Lic60] J. C. R. Licklider. Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics*, HFE-1(1):4–11, March 1960.

[McK] John A. McKenzie. Tx-o computer history, October 1.

[Min86] Marvin Minsky. *The Society of Mind*. Simon and Schuster, New York, 1986.

[Mor71] Charles Morris. *General Theory of Signs*. Mouton, Paris, 1971.

[Mor79] Charles Morris. *A Theory of Semiotics*. Indiana University Press, Bloomington, 1979.

[MT89] Marilyn M. Mantei and Toby J. Teorey. Incorporating behavioral techniques into the systems development life cycle. *MIS Quarterly*, 13(3):257–274, 1989.

[ND86] Donald A. Norman and Stephen W. Draper. *User Centered System Design; New Perspectives on Human-Computer Interaction*. L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 1986.

[Nie93] Jakob Nielsen. *Usability engineering*. Morgan Kaufmann, Mountain View (CA), 1993.

[Nor01] Donald A. Norman. Cyborgs. *Commun. ACM*, 44(3):36–37, March 2001.

[Pew03] Richard W. Pew. Introduction. Evolution of human-computer interaction: From memex to bluetooth and beyond. In J.A. Jacko and A. Sears, editors, *The Human-Computer Interaction Handbook. Fundamentals, Evolving Technologies, and emerging Applications*. 1 edition, 2003.

[Ras] J. Rasmussen. Mental models and the control of actions in complex environments.

[Ras90] J. Rasmussen. Mental models and the control of action in complex environments. In D. Ackermann and M.J. Tauber, editors, *Mental Models and Human-Computer Interaction 1*, volume 1, pages 41–69. Elsevier Science Publishers B.V. (North-Holland), Amsterdam, 1990.

[SRS83] A. Stevens, B. Roberts, and L. Stead. The use of a sophisticated graphics interface in computer- assisted instruction. *IEEE Computer Graphics and Applications*, 3(2):25–31, March 1983.

[Sut] Ivan Edward Sutherland. Sketchpad: A man-machine graphical communication system, September. PhD Thesis.

[Wei93] Mark Weiser. Some computer science issues in ubiquitous computing. *Commun. ACM*, 36(7):75–84, jul 1993.

[ZCT+04] Ping Zhang, Jane Carey, Dov Te'eni, , and Marilyn Tremaine. Integrating human-computer interaction development into sdlc: A methodology. *AMCIS 2004 Proceedings*, (574):4620–4626, 2004. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.231.9408&rep=rep1&type=pdf.