
Artefactual Intelligence: The Extended Mind Hypothesis & Distributed Cognition

Nils Dahlbäck

Department of Computer and Information Science

Linköping University

www.ida.liu.se/~nilda

Cognition as computation:

The theoretical core of early CogSci

A common theoretical perspective shared by a number of cognitive sciences

- AI: cognition as symbolic processing; the Physical Symbol System Hypothesis
- Philosophy:
 1. (Machine) Functionalism
 2. Language of thought
- Psychology: Human Information Processing

Some common characteristics

- Cognition is (or requires) an independent level of description
- Cognition as a separate system or module possible to study in isolation from other parts of the agent (body, sensory organs, I/O gates, etc.)
- Cognition as computation

Some common characteristics (cont.)

- Functional descriptions – does not require knowledge of the material base (neurology/hardware)
- Focus on the (isolated) individual – all processes are intracranial

New directions in Cognitive Science

- Two (three) new theoretical perspectives
- Both keep some assumptions of the standard model and revise some – but in different ways

Developmental trail 1

“the biology trail”

- Neurocognition
 - Keeps the focus on the individual
 - Emphasizes the connection cognition-neurology
 - Laboratory studies
- Embodied cognition
 - Keeps the focus on the individual
 - Emphasizes the connection between cognition and perceptual and motor systems
- Evolutionary cognitive science

Developmental trail 2

“the cultural trail”

- Situated cognition
 - Focus on the use of artifacts
 - Field studies
 - Also lab studies – e.g. Tetris players
- Distributed cognition
 - Cognition is (also) a process external to the individual brain/mind
 - Cognitive processes are distributed between many persons and artifacts
 - Field studies (“Cognition in the Wild”)

Some representative titles of early papers and books

- Adapting the Environment instead of Oneself (Kirsh, 1966)
- The Intelligent Use of Space (Kirsh, 1995)
- On Distinguishing Epistemic from Pragmatic Actions (Kirsh and Maglio, 1995)
- Cognition in the wild (Hutchins, 1996)
- How a cockpit remembers its speeds (Hutchins 1996)
- Being there: Putting brain, body, and world together again (Clark, 1997)

Today's seminar papers

- Situated cognition:
Clark and Chalmers (1998) *The Extended Mind*
- Distributed Cognition
Hutchins (1995) *Cultural Cognition*
- Both present radical theoretical perspectives on cognition – how similar and different are they?

The extended mind

Clark & Chalmers

The Extended Mind

“Where does the mind stop and the rest of the world begin?”

“We advocate an *active externalism*, based on the active role of the environment in driving cognitive processes”

The parity principle

”If, as we confront some task, a part of the world functions as a process which, *were it done in the head*, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world *is* (so we claim) part of the cognitive process”

Illustrative case 1: Tetris

- Tetris players decide faster what to do if they manipulate the zoids (tiles) on the screen than if they perform mental transformation a la Shepard and Metzler

On distinguishing epistemic from pragmatic action (Kirsh & Maglio, 1995)

- C&C also add a third version where a person has a neurological implant which works as fast as the computer

Illustrative case 2: Inga and Otto

- Is claimed to show that also beliefs can be extended, i.e. is an argument for an extended mind

Inga

Inga hears from a friend that there is an exhibition at the Museum of Modern Art, and decides to go see it. She thinks for a moment and recalls that the museum is on 53rd Street, so she walks to 53rd Street and goes into the museum. It seems clear that Inga believes that the museum is on 53rd Street, and that she believed this even before she consulted her memory. It was not previously an *occurrent* belief, but then neither are most of our beliefs. The belief was sitting somewhere in memory, waiting to be accessed.

Otto

Otto suffers from Alzheimer's disease, and like many Alzheimer's patients, he relies on information in the environment to help structure his life. Otto carries a notebook around with him everywhere he goes. When he learns new information, he writes it down. When he needs some old information, he looks it up. For Otto, his notebook plays the role usually played by a biological memory. Today, Otto hears about the exhibition at the Museum of Modern Art, and decides to go see it. He consults the notebook, which says that the museum is on 53rd Street, so he walks to 53rd Street and goes into the museum.

Conclusion

- “Clearly, Otto walked to 53rd Street because he wanted to go to the museum and he believed the museum was on 53rd Street. And just as Inga had her belief even before she consulted her memory, it seems reasonable to say that Otto believed the museum was on 53rd Street even before consulting his notebook. For in relevant respects the cases are entirely analogous: the notebook plays for Otto the same role that memory plays for Inga. The information in the notebook functions just like the information constituting an ordinary non-occurrent belief; it just happens that this information lies beyond the skin.”

If we accept the conclusion – where do we stop?

- It is absurd to claim that my mind includes all information I can find on the Internet
- C&C list a number of criteria for when to count something external as part of the mind
 - It is a constant in the life of the individual
 - It is directly available without difficulty
 - The information retrieved is automatically endorsed
 - (The information in the notebook has been consciously endorsed in the past)

Summary

- Primarily a thesis about individual cognition
- Cognition extends into the world through the use of artifacts
- Also beliefs can be external → the mind extends into the world
- Empirical base includes also invented cases
... has lead to an intense debate – but unfortunately mostly by philosophers

Distributed cognition

Edwin Hutchins, UCSD

Some delimitations

- Many overlapping perspectives
 - Situated cognition
 - Distributed cognition
 - Embodied cognition
 - ...
- Focus here primarily on Hutchins' work

A revolution in cognitive science?

”The central claim of distributed cognition is so radical that it may well reorganize the whole of cognitive science”

Latour (1996)

Revolution? ... but not new thoughts

”In order to explain the highly complex forms of human consciousness one must go beyond the human organism. One must seek the origins of conscious activity and 'categorical' behavior not in the recesses of the human brain or in the depths of the spirit, but in the external conditions of life. Above all, this means that one must seek these origins in the external processes of social life, in the social and historical forms of human existence.”

(Luria, *Language & Cognition*).

Cognition is distributed between many persons and artifacts

“Cognition has nothing to do with minds nor with individuals, but with the propagation of representations through various media, which are coordinated by a very lightly equipped human subject working in a group, inside a culture, with many artefacts and who might have internalized some parts of the process”

Latour (1996)

Cognition is distributed between many persons and artifacts

“Cognition has nothing to do with minds nor with individuals, but with the propagation of representations through various media, which are coordinated by a very lightly equipped human subject working in a group, inside a culture, with many artifacts and who might have internalized some parts of the process”

Latour (1996)

Distributed cognition

- Often studied complex work environments
 - Navigation teams on large ships
 - Pilots cooperation in the cockpit
- Often mis-interpreted as “applied cognitive science”

The relationship between specific cases and the general theory

”there are powerful regularities to be described at a level of analysis that transcends the details of the specific domain.

”It is not possible to discover these regularities of the domain without understanding the details of the domain, but the regularities of the domain are not about the domain specific details, they are about the nature of human cognition in human activity”

Hutchins (1992)

What is “culture”?

“Whatever one needs to know in order to behave appropriately in any of the roles assumed by any member of a society”

Goodenough (1957)

- Emphasizes knowledge, not things/artifacts

Hutchins' two main messages

- Culture is not a collection of things – culture is a process
- Cognition is a cultural process

"I am proposing an integrated view of human cognition in which a major component of culture is a cognitive process (...) and cognition is a cultural process"

"Culture is not any collection of things (...). Rather it is a process. It is a human cognitive process that takes place both inside and outside the minds of people"

Two more quotes on culture

- "Culture is an adaptive process that accumulates partial solutions to frequently encountered problems"
- "The understanding of the individual that has developed without consideration of cultural process is fundamentally flawed"

Hutchins' conclusion

“Some of what have been done in cognitive science must now be undone so that these things can be brought into the cognitive picture”

Gardner's Five Key Features of (early) CogSci

- Representations
- Computers
- De-emphasis on affect, context, culture, and history
- Belief in interdisciplinary studies
- Rootedness in classical philosophical problems

How did we get here?

Hutchins' answer: because of a fundamental misunderstanding of what a computer is

We need to re-interpret the history of cognitive science

- The idea that cognition is a representation of mental symbols is based on a misunderstanding of the similarities and differences between people and computers
- Doing this requires a re-interpretation of the "standard model" of cognition as symbol manipulation (the PSS-hypothesis)
- This in turn requires an understanding of how the "error" was made in the first place

The symbol system hypothesis

Physical Symbol System Hypotesis: The necessary and sufficient condition for a physical system to exhibit general intelligent action is that it be a physical symbol system

- Both a computational theory of mind” and a ”computational theory of cognition”

Hutchins agrees that cognition is a computational process

- Formal computation is central to our culture
- Formal computation is central to cognition

” I consider the mastery of formal systems to be the key to modern civilization. This is a very, very powerful idea”.

CiTW, p 300

Computers again ...

- Computers are machines that manipulate formal symbols
- First developed as an abstract theoretical machine: The Turing machine

Turings own description (according to Dennet)

- Reflected on how he did when he solved mathematical problems
- Broke down the process into its smallest constituents
- Process description:
 - Decide on which rule to use
 - Apply it and write down the answer
 - View the result and decide on which rule to be used

What is this?

- A description of a person doing formal computations using his hands and eyes and paper and pencil
- The functional cognitive system is the person doing computations with external aids
- The properties of this system are completely different from the properties of the person
- Hutchins: "What Turing modeled was the computational properties of a socio-cultural system"

What Turing did with the Turing machine

- Turing described how one could make the formal symbol manipulation automatically
- This was done by placing the rules of the person/mathematician within the system (the computer)
- But:
”The physical symbol system architecture is not a model of individual cognition. It is a model of the operation of a sociocultural system from which the human actor has been removed”

In summary

- Not a critique of "thinking as computation"
- But a re-analysis of what the view of "thinking as computation" entails for cognitive science

What went wrong according to Hutchins(1)

”Having failed to notice that the central metaphor of the physical symbol system hypothesis captured the properties of a sociocultural system rather than those of the individual mind, AI and information processing proposed some radical conceptual surgery for the modeled human. The brain was removed and replaced by a computer.”

What went wrong according to Hutchins(2)

“The surgery was a success. However, there was an apparently unintended side effect: the hands, the eyes, the nose, the mouth, and the emotions all fell away when the brain was replaced by a computer”

Conclusions

- It is not an accident that it is difficult to integrate emotion, culture, body etc. in the "standard model"
- It is a logical consequence of the original flawed analysis
- It is not possible to just "add" emotion, culture and body to the original model – also that need to be revised

Hutchins' conclusion

"The computer was not made in the image of a person. The computer was made in the image of the formal manipulations of abstract symbols. And the last 30 years of cognitive science can be seen as attempts to remake the person in the image of a computer."

A very common mis-understanding of Hutchins work

- Hutchins is often claimed to be a defender of the computer metaphor of mind, e.g. Button:

“Following a cognitivist tradition, ‘distributed cognition’ understands ‘mind’ computationally, running processes on inner representations, and it is by reference to these inner representations that human action is understood”

“Thus, unlike many who use ethnographic techniques for studying human action in context where all that is on hand is the observed data, ‘distributed cognition’ is attempting to relate that data to computational processing in the brain.”

But ...

... Hutchins separates the computation part of cognition from the computer part of the classical view

The Chinese Room Again

- The Chinese Room is also a sociocultural system
- Hutchins: What Searle shows is that the cognitive properties of the entire system is something different from the cognitive properties of the isolated person
- Comment: OK, but is cognitive science then still "The Mind's New Science"?

Some consequences of the standard model's view of cognition(1)

- Too much of the cognitive processes are placed inside the head

"When one commits to the notion that all intelligence is inside the inside/outside boundary, one is forced to cram inside everything that is required to produce the observed behaviors"

"If we fail to bound the system properly, we may attribute the right properties to the wrong system or (worse) invent the wrong properties and attribute them to the wrong system"

Some consequences of the standard model's view of cognition(2)

- Creates a too sharp dividing line between perception and cognition
- Too much of the cognitive processes are located inside the head
- Research is occupied with studying mentally difficult problems (puzzels) ... often in lab settings

Differences between everyday cognition and cognition in the lab

Laboratoriet

- Uncommon
- Isolated
- Poor meta-cognition
- Experimenter defines the problem
- Special purpose skills
- Novel tasks
- One time performance
- *"In captivity"*

Vardag

- Common
- Ecological
- Rich meta-cognition
- Actor defines the problem
- General abilities
- Familiar tasks
- Repeated performances
- *"In the wild"*

Samspel konkreta fall – generell teori

”there are powerful regularities to be described at a level of analysis that transcends the details of the specific domain.

”It is not possible to discover these regularities of the domain without understanding the details of the domain, but the regularities of the domain are not about the domain specific details, they are about the nature of human cognition in human activity”

Hutchins (1992)

Comparing Hutchins to the standard model (1)

- **Similarity:** Cognition is a process of transformations of representations
- **Difference:** View of mental representations
 - Standard: central to mental processes
 - Hutchins: exist, but not central

Comparing Hutchins to the standard model (2)

- View of computational theories (Computational theory of mind – computational theory of cognition)
 - Standard model: Same thing
 - Hutchins:
 - For "Computational theory of cognition"
 - Against "Computational theory of mind"
- Relationship between cognition-mind
 - Standard: closely connected; two sides of the same coin
 - Hutchins: ??????

Some final questions

- How does Hutchins' view fit with the other perspectives presented here
- How much of what we usually call "cognition" is left out of the picture