Cognition as symbol manipulation
What is ”computation”?
The PSS hypothesis

Nils Dahlbäck
Department of Computer and Information Science
Linköping University
www.ida.liu.se/~nilda

Overview of contents

- What is a computer?
- What is ’computation’?
- Cognition as symbol manipulation/Computation
- The PhysicalSymbol System hypothesis
- Comments on today’s seminar texts

What is computation?

- Computation is one of mankinds oldest inventions
- Formal systems are used to calculate conclusions that are formally valid
- The specific contents are irrelevant to the validity or truth of the conclusion

Two simple examples – both equally valid and correct

- All humans are mortal
- All athenians are human

➔

- All athenians are mortal

- All teachers are educated in how to teach
- All educated to teach are brilliant lecturers

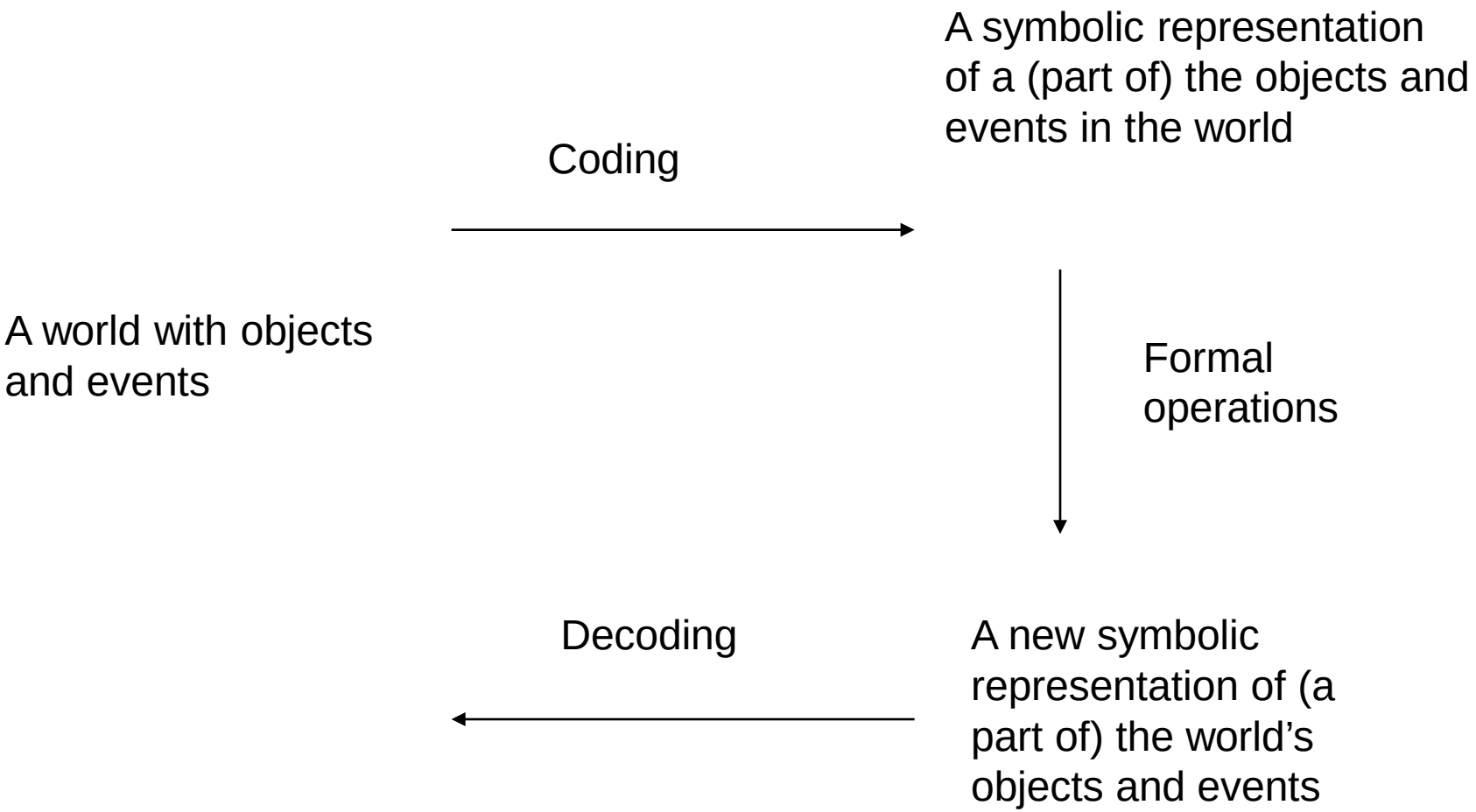
➔

- All teachers are brilliant lecturers

Three features of formal calculations

- Based on symbols (representations)
- The symbols are manipulated/transformed through formal procedures (rules)
- Both representations and rules are materially independent – but not immaterial!

The basic model



Some tricky questions

- How is this possible?
- How can syntactic operations preserve semantic relations?
- Why is it possible to mathematically model the world?
- Answer: Noone knows, but it works!
- Haugeland *”Take care of the syntax, and the semantics will take care of itself”*

What is a computer?

- A computer is not a (traditional) machine!
- A traditional machine transforms energy
- Computers transform formal symbols (representations)
- Computers are physically realized automatic symbol representation transformers (i.e. they are Computing machines ☺)

A conclusion (or hypothesis): Cognition is Symbol Transformations

- Formal transformation systems are independent of their particular physical instantiations
- Cognitive processes are (some kind of) symbol transformations
- Conclusion: Cognitive processes *can* and *must* be studied independently of the specific details of the physical implementation (i.e. the properties of the brain)

The Pyshical Symbol Systems Hypothesis (1)

- Developed by Newell & Simon
- Claimed to be an empirical discovery from the early days of computer science
- *Symbol systems* are the same as *universal machines* (Newell, 1981, p 56)
- Computers are (almost) universal machines
- Universal machines can compute anything that can be computed

The Pyshical Symbol Systems Hypothesis (2)

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

The Pyshical Symbol Systems Hypothesis (3)

Necessary means that any physical system that exhibits general intelligence will be an instance of a physical symbol system

Sufficient means that any physical symbol system can be organized further to exhibit general intelligent acdtion

General intelligent actions means that the same scope of intelligence seen in human action: that in real situations behavior appropriate to the ends of the system and adaptive to the demands of the environment can occur, within some physical limits

(Newell, 1981 p. 72)

A comment on the Chinese Room (Searle, 1980)

- An often misunderstood argument
- Searle does not deny that machines can think
  - ”We are precisely such machines”
- Searle’s point is that computers can’t have a mind, since since they only transform formal symbols
- Formal symbols does not have a semantics (cannot have intentional states)

Today's seminar papers and topics

- Newell's paper summarizes his view on cognition as a physical symbol system
- Newell also introduces an additional theoretical level – the knowledge level (similar to Dennet's *Intentional Stance*)
- The discussion and critique of Newell's views by Brian C. Smith, computer scientist and philosopher, and by Daniel Dennet, philosopher