Turings „Computing Machinery and Intelligence“ (1950) vs. Searles „The Myth of the Computer“ (1982)

Turings argumentation:

Turing in 1950 has strong beliefs that computers in the future will be able to compete with men in all intellectual fields. He tries to show that with enough speed and storage capacities computers will be able to carry out operations that cannot be distinguished from human thinking. His argumentation goes as follows:

- The question „Can machines think?“ cannot be discussed because the words *machine* and *thinking* would have to be defined in the way they are commonly used, which seems absurd to him.
- Instead he replaces it by the *closely related* (Turing p.524) and *more accurate* (Turing p.527) question „Can a computer successfully play the role of A in the Imitation Game?“
- He then introduces the Imitation Game: Person A, a male and person B, a female are interrogated by person C, who is in a separate room. The communication is done by teleprinter, so other sensual inputs are excluded. C has to find out, who is the male and who the female. Person A has the task to lead C to the wrong conclusion, person B has the task to lead C to the right conclusion.
- He then changes the task of the Imitation Game to distinguishing a computer from a human. Person C now has to find out who is which.
- He restricts the meaning of *machine* to discrete state machines, digital computers „which move by sudden jumps from one quite definite state to another.“ (Turing p.526).
- He then refutes eight contrary views to the original question „Can machines think“.
- At last he introduces two possible ways to construct machines that could become successful in playing the imitation game.

Critique on Turing:

Focussing his string of argumentation I notice several problems. His starting point that definitions of *thinking* and *machine* would have to follow a poll on how these words are commonly used is unclear. He could as well define them for his problem, as he later does with *machine* anyway. He

thus believes the original question „Can machines think?“ to be too meaningless to deserve discussion (Turing p.527), which wouldn‘t be the case otherwise.

The stated *close relatedness* of his question „Can machines be successful in the interpretation game“ is not argued. Turing also rules out the difference of imitation and thinking for his question without argument: „May not machines carry out something which ought to be described as thinking but which ist very different from what a man does? This objection is a very strong one, but at least we can say that if, nevertheless, a machine can be constructed to play the imitation game satisfactorily, we need not be troubled by this objection.“ (Turing p.526). This is easily criticized, since imitating and thinking are two distinctable actions, although not necessarily from a third persons view.

Turing turns back to the original question and offers his own point of view as conjectures and beliefs without arguments for their plausibility: When speed and storage capacities can be provided, computers will be able to compete with humans in all intellectual fields (Turing p.535). The meaning of *thinking* will have changed, so playing the imitation game successfully will be accepted as *thinking* (Turing p.527).

Although he himself does not discuss the original question, he refutes contrary views to it, partially with straw man arguments: e.g. in the *Head in the Sand Objection* he creates a stance that is dequalified by its lack of argument. In the solipsism reproach of his refutation of the c*onsciousness objection* he builds on theory of mind ideas, which construct minds as completely separated entities that have to bridge a gap to other minds by reconstructing them within themselves.

Turing admits that he criticises opposers more than giving positive evidence for his own view (Turing p.534). When he finally promises to do so, he merely repeats his own beliefs as hopes for the future without further argumentation.

Searles argumentation:

Searle in 1980 reviewed the book *The Mind's I: Fantasies and Reflections on Self and Soul*

by Hofstadter and Dennett, in which he criticises the Turing-Test as the criterion of the mental. His argumentation goes as follows:

- He introduces three theses of *Strong Artificial Intelligence*, which he opposes:
  a) *Mind as a program*: any system with the right input and output has mental states.
  b) *Irrelevance of the neurophysiology of the brain*: minds are abstract sort of things. There is no essential connection between the mind and the organic brain.
  c) *Turing-Test as the criterion of the mental*: passing the Turing-Test is the conclusive proof of the presence of a mental state.
- Searle first refutes the i*rrelevance of the neurophysiology of the brain* by showing a computers incapacity to produce thirst by simulating the formal properties of chemical and electrical phenomena involved. He then equates such strongly bodily related phenomena with mental states such as understanding, feeling, worrying.
- He refutes *mind as a program* by analysing computer actions down to elements that themselves do not symbolize any content. „[…] the computer has a syntax not the semantics.“ (Searle p.539), the interpretation of the symbols being only up to their programmers and users (Searle p.538).
- Challenging the Turing-Test he introduces the *Chinese-Room* thought experiment, in which he himself performs predefined actions as computers do: he does not know any Chinese, but is locked in a room with Chinese symbols and rules how to match these symbols together. He is passed Chinese symbols from the outside, which are called questions, and has to pass back the symbols, which are called answers, according to the defined rules.
- Searle concludes that sufficiently acting according to a formal program does not necessarily imply understanding.

Critique on Searle:

Searle refutes the *irrelevance of the neurophysiology of the brain* by showing a computers incapacity to produce bodily phenomena by simulating the formal properties of chemical and electrical phenomena involved. He thereby defines the mental as a bodily phenomenon, which, taken as an argument, is an induction that doesn‘t prove that the mental couldn‘t also appear on other grounds.

His appeal to common sense: „No one in his right mind thinks ...“(Searle p.537) is polemically dequalifying his opposers more than arguing his point, that simulation and the simulated action are different. Searles amazement about people believing that computers may produce mental phenomena is also not an argument that proves that this cannot be the case. His belief that everything we have learned from biology suggests that mental phenomena are part of our biological natural history (Searle p. 538) does not really prove his point either.

With his *Chinese-Room* thought experiment Searle cleverly undermines the *consciousness objection* reproach by putting himself in the role of the computer.

As does Turing, Searle also uses straw man argumentations, when e.g. towards the end he without being asked consoles his imagined opposers by suggesting that they don‘t have to be afraid: computers will never be a threat by becoming similar to us (Searle p.540).

Discussion:

Both Turing and Searle seem to be building on reduced ideas of thinking. Turing equates it to deceiving a person, Searle concedes that everything instantiates some program or other, „So, in a trivial sense brains, like everything else, are digital computers.“ (Searle p.540).
To accept such ideas of thinking as basic, produces the image that it is a predisposition for other more complex ways of thinking. But isn‘t it more so just one very abstract form of thinking, stripped of its context in life, cut off from feeling, wanting something, be bothered by something, deciding, dealing with the lived world?
Searle believes that brain processes *cause* mental states, which a program cannot. But that certain bodily processes correlate with mental states does not prove such a causal chain. He cannot and does not explain what causes mental states, meaning and understanding.

People are different from machines, organic interaction differs from anorganic. Why should we lump them together? Turings comparison of babies to computers shows very poor observation: Babies are not basically programmed computers but actively interacting beings from the very beginning of their existence. Also the role of award and punishment in education is very small compared to the complex learning in interaction and living together in a shared world.

Searle with his *Chinese-Room* thought experiment gives a convincing example that sufficient acting according to a formal program does not necessarily imply understanding. But does that prove that it could never develop in that direction? Humans acquire language by associating experiences with meaningless syllables, concededly taught and learned by complex interaction with their environment and other beings. Must understanding already be there beforehand, does it develop alonghand, or even afterwards, only made possible by the acquisition of language?

Theory of Mind is not the only approach to understanding human minds. Phenomenologist attempts show an inseparable connection of minds (Merleau-Ponty, Heidegger, Ratcliffe), and interaction theories a direct access to other minds in common action (Slaby).

We know circumstances in which it happens, but we do not know how consciousness and meaning come to being. Consciousness develops in interaction in a shared world. *Dasein* (being-there) is *Mitsein* (being-with), as Heidegger says.
To say that thinking is caused by electric or chemical signals in the brain remains a mere statement. These may also be just the bodily manifestations of thoughts, feelings and other mental acts.
We do not know how consciousness and thinking comes to the world. To propose that computers will develop this phenomenon, once we provide sufficient speed and storage capacities may well be a myth inducing human development across its borders. Denying that possibility may come from the very same difficulty in thinking expressed in the science fiction story by Terry Bisson, where members of a non-organic species discover human life from their spaceship:

> "They're made out of meat."
> "No brain?"
> "Oh, there's a brain all right. It's just that the brain is made out of meat!
> "So ... what does the thinking?"

They finally decide to erase the records and forget their discovery, not to have to go through the trouble of explaining that.

Literature:

Bisson, T.: „They‘re made out of meat“. Omni 1990. https://web.archive.org/web/20190501130711/http://www.terrybisson.com/theyre-made-out-of-meat-2/ 2020-02-01 11:00)

Heidegger, M.: *Sein und Zeit*. GA. I. Abt.: Veröffentl. Schriften 1910–1976. Bd. 2, hg. von Friedrich Wilhelm von Herrmann. Vittorio Klostermann. Frankfurt am Main, 1977.

Merleau-Ponty, M.: *Phenomenology of Perception*. Routledge and Kegan Paul. London and Henley, 1962.

Ratcliffe, M.: *Rethinking Commonsense Psychology – A Critique of Folk Psychology, Theory of Mind and Simulation*. Palgrave Macmillan. Houndsmills, 2007.

Searle, J. R.: „The Myth of the Computer“ in Sinnott-Armstrong, W., Fogelin, R. *Understanding Arguments*. Wadsworth, 2010, 536-541.

Slaby, J.: “Empathy’s blind spot” in Med Health Care and Philos 17, 249-258. Springer. Dordrecht, 2014.

Turing, A.M.: „Computing Machinery and Intelligence“ in Sinnott-Armstrong, W., Fogelin, R. *Understanding Arguments*. Wadsworth, 2010, 524-536.