Mind: Architectures, Models, Formalisms

2. Is the Universal Turing Machine (UTM) a mind?

(Or: If the UTM is not a mind, then what is?)

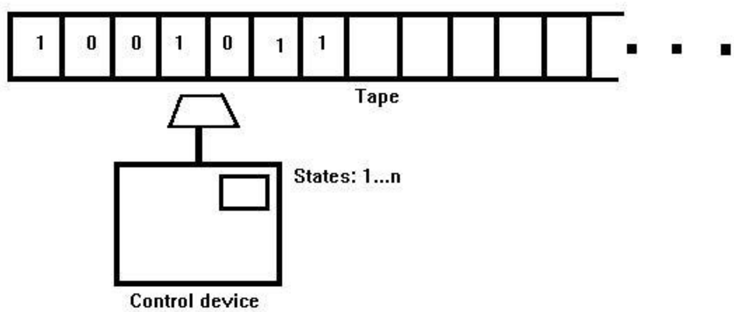
Luís M. Augusto

Institute of Philosophy, University of Porto / FCT

A 1900 question and an answer in1936

- The (implicit) question (Hilbert, 1900): What is an effective calculation procedure, or what is an effectively calculable function?
- The answer (*Church-Turing Thesis*, 1936): An effectively calculable function is a function computable by a Turing machine.

The Turing Machine



(ucla.edu)

17-05-2013

A (simplified) Turing Machine (TM) is a 5-tuple $\mathcal{M} = (Q, \Sigma, \Gamma, q_0, \delta)$,where

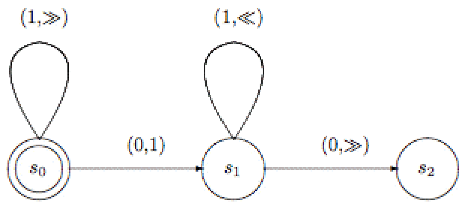
- Q is a finite set of states; $q_0 \in Q$
- Σ is an input alphabet (typically, $\Sigma = \{0, 1\}$)
- Γ is a tape alphabet (e.g., $\Gamma = \{0, 1, \mathrm{X}\}$)
- δ is a transition function $\delta : (q_i, s) \longrightarrow (q_j, A)$, where
  - q_i and q_j are the current and next states, respectively
  - A is a two-step action which defines the operation to be carried out in terms of tape symbols ($s*$) and the direction instruction (d) for the tape:
    - (i) $s* =$ keep s, or erase s and leave a blank square (_), or erase s_k and write s_l
    - (ii) $d =$ move to the right (R) / left (L); of course the machine can just stop (S)

The program for a TM is the list of all its transition functions. E.g.:

(q_i, s)	$(q_j, s*, d)$
$q_0, 1$	$q_0, 1, R$
$q_0, 0$	$q_1, 1, R$
$q_1, 1$	$q_1, 1, R$
$q_1, 0$	$q_2, 0, L$
$q_2, 1$	$q_2, 0, R$

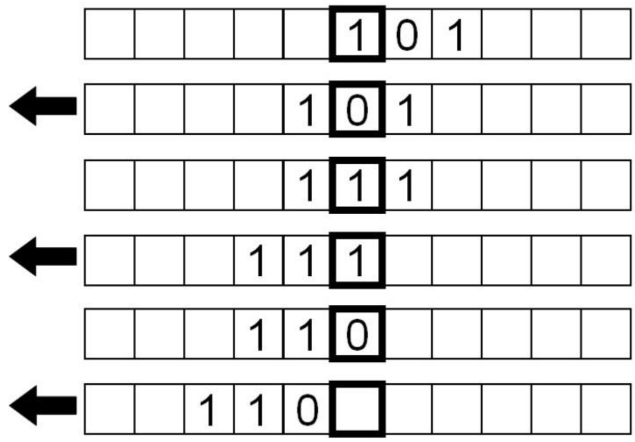
17-05-2013

(To verify what the computed function is, apply the program to the input (e.g.) 11011. If a square is blank, then it has a 0.)



(stanford.plato.edu)

(cam.ac.uk)



17-05-2013

1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think ?'

The new form of the problem can be described in terms of a game which we call the 'imitation game'. It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either 'X is A and Y is B' or 'X is B and Y is A'. The interrogator is allowed to put questions to A and B thus:

C: Will X please tell me the length of his or her hair ?

Now suppose X is actually A, then A must answer. It is A's object in the game to try and cause C to make the wrong identification. His answer might therefore be

'My hair is shingled, and the longest strands are about nine inches long.'

In order that tones of voice may not help the interrogator the answers should be written, or better still, typewritten. The ideal arrangement is to have a teleprinter communicating between the two rooms. Alternatively the question and answers can be repeated by an intermediary. The object of the game for the third player (B) is to help the interrogator. The best strategy for her is probably to give truthful answers. She can add such things as "I am the woman, don't listen to him !' to her answers, but it will avail nothing as the man can make similar remarks.

We now ask the question, 'What will happen when a machine takes the part of A in this game ?' Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman ? These questions replace our original, 'Can machines think ?'

Turing (1950), 434-4

17-05-2013

6. *The universal computing machine.*

It is possible to invent a single machine which can be used to compute any computable sequence. If this machine $\mathfrak{U}$ is supplied with a tape on the beginning of which is written the S.D of some computing machine $\mathcal{M}$, then $\mathfrak{U}$ will compute the same sequence as $\mathcal{M}$.

Turing (1936-7), 241-2

Given the table corresponding to a discrete state machine it is possible to predict what it will do. There is no reason why this calculation should not be carried out by means of a digital computer. Provided it could be carried out sufficiently quickly the digital computer could mimic the behaviour of any discrete state machine. The imitation game could then be played with the machine in question (as B) and the mimicking digital computer (as A) and the interrogator would be unable to distinguish them. Of course the digital computer must have an adequate storage capacity as well as working sufficiently fast. Moreover, it must be programmed afresh for each new machine which it is desired to mimic.

This special property of digital computers, that they can mimic any discrete state machine, is described by saying that they are *universal* machines. The existence of machines with this property has the important consequence that, considerations of speed apart, it is unnecessary to design various new machines to do various computing processes. Turing (1950), 441

Objections against [human(-like)] thought in machines (i.e., digital computers) (Turing, 1950)

- Theological: Thinking is a function of man's God-given immortal soul. Humans are the only recipients of this gift.
- "Heads in the sand": Thinking machines would entail terrible consequences.
- Mathematical: There are computation limitations to discrete state machines (e.g., incomputable functions).
- Argument from consciousness: Machines cannot feel. In particular, they cannot have feelings regarding self-accomplishment and limitations; namely, they cannot feel themselves thinking.
- Arguments from various disabilities: Machines cannot do X (e.g., fall in love).
- Etc.

What makes a mind / what does a mind make?

- Problem solving, decision making, routine action
- Memory, learning, skill
- Perception, motor behavior
- Language
- Motivation, emotion
- Imagining, dreaming, daydreaming,...

[the author felt daunted and could not bring himself to complete the list]

(Areas to be covered by a unified theory of cognition; Newell, 1990, p. 15)

References & Internet sources of the figures

Hilbert, D. (1902). Mathematical problems. *Bulletin of the American Mathematical Society*, *8*, 437-479. (Originally published in German in 1900.)

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Turing, A. M. (1936-7). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society (2)*, *41*, 230-265.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *59*, 433-460.

- http://www.cl.cam.ac.uk/projects/raspberrypi/tutorials/turing-machine/one.html
- http://plato.stanford.edu/entries/turing-machine/
- http://www.read.cs.ucla.edu/111/notes/lec1