Universal Turing machines

Arno Pauly

March 16, 2021

What we want to show today

Theorem

There is a Turing machine U (called universal TM) which can receive as input a description of a TM M and some word w, and then simulates whatever M would do on input w.

What is a description?

- Let the alphabet be $\{a, b, c, \perp\}$.
- We will use *a*,*b* for binary encoding, and *c* as separation marker.
- There is a lot of flexibility in the encoding, this one is just one convenient choice.

What is a description?

- Let $[n]_2$ denote the binary encoding of $n \in \mathbb{N}$ using a and b as digits. We also encode $d \in \{L, R, S\}$ as $[d]_D$ and $x \in \{a, b, c, \perp\}$ as $[x]_\Sigma$ using digits from $\{a, b\}$.
- A state q_n shall be described by the word

  $$w_n := [n]_2 c[m_a]c[x_a]c[d_a]_D c[m_b]c[x_b]c[d_b]_D c[m_c]c[x_c]c[d_c]_D \dots$$

  $$\dots c[m_\perp]c[x_\perp]c[d_\perp]_D$$

  , where the rules are "on reading an $\alpha \in \{a, b, c, \perp\}$, write the symbol x_α, move in direction d_α and go to the m_α-th state next".
- A TM having states $q_1, \dots, q_k$ is described by the word $w_1 cc w_2 cc \dots cc w_k$.

How to built the universal TM

- U receives an input $\langle M \rangle cccw$, where $\langle M \rangle$ is a description of a TM.
- U uses three tapes, and begins with writing $[1]_2$ on the second tape (which will keep track of what state M would be in)
- and copies w to the third tape (which will play the role of M's tape), and move the head back to the start position.
- then U searches for the rules pertaining to the current state on the firs tape, acts on the third tape as prescribed, and copies the next state description to the second tape.
- U checks whether $[2]_2$ or $[3]_2$ is written on the second tape (assuming that $q_2 = q_{\text{yes}}$ and $q_3 = q_{\text{no}}$ and if so halts, otherwise it repeats from the previous step.

Something else

- Rather than having a TM just answering “yes”-“no” questions, we can have it compute functions.
- Instead of q_{yes} and q_{no} states we just have q_{halt}.
- On reaching q_{halt}, the output is determined as the longest $\perp$-free word on the tape including position 0.
- Practise task: Come up with a TM that performs addition in unary (easy) or in binary (harder). Assume that the inputs are separated by a separation symbol, and use three tapes.

Outlook

- Rather than having a TM deciding a language, we can have it recognize a language (by just requiring halting and saying “yes” for words in the language, and not caring about words outside of the language).
- The languages recognizable like this are the computably enumerable languages – the same generated by unrestricted grammars.
- There is no TM that decides, on input $\langle M \rangle$, whether M ever halts on input ε.