Turing Machine Universality

Data/Program duality for Turing machines: Turing machine control tables can be written on a Turing machine’s tape as input.

The enumerability of Turing machines: We can impose an ordering $M_1, M_2, M_3, \ldots$ on Turing machines.

Universality of Turing machines: There is a universal Turing machine $U$ that takes any program $M_i$ and any input $x$ as input and simulates what $M_i$ would do given input $x$. That is,

$$U(M_i, x) \begin{cases} \text{outputs } M_i(x) & \text{if } M_i \text{ halts on } x \\ \text{runs forever} & \text{if } M_i \text{ runs forever on } x \end{cases}$$

Simplifying Assumption: Though our results hold in the general case, for ease of argument we will assume that we are only looking at Turing machines whose input is a finite sequence of A’s. We will also use the simplifying notation that running $M_i$ on input $j$ A’s can be written as $M_i(j)$.

Definition: The halting function $h(M_i, j)$ takes as input a Turing machine $M_i$ and an input string of $j$ A’s and has value 1 if $M_i$ halts when running on $j$ A’s as input, and has value 0 if $M_i$ runs forever when running on $j$ A’s an input. That is,

$$h(M_i, j) = \begin{cases} 1 & \text{if } M_i(j) \text{ halts} \\ 0 & \text{if } M_i(j) \text{ runs forever} \end{cases}$$

Uncomputability Theorem: The halting function is not computable. That is, there is no Turing machine that can compute $h$.

We will prove this by assuming such a Turing machine exists and reaching a contradiction.

Proof Outline:

Definition: The Universal Termination Detector, $D$, is a Turing machine that computes the halting function$^1$:

$$D(M_i, j) \text{ outputs } \begin{cases} 1 & \text{if } h(M_i, j) = 1 \\ 0 & \text{if } h(M_i, j) = 0 \end{cases}$$

If $D$ exists, we can build an “evil” Turing machine $X$ using $D$ which takes $j$ A’s as input, simulates $D$ on $M_j$ and $j$ as input, and has the following output behavior:

$$X(j) \begin{cases} \text{outputs } 1 & \text{if } D(M_j, j) = 0 \\ \text{runs forever} & \text{if } D(M_j, j) = 1 \end{cases}$$

This evil Turing machine $X$ will allow us to reach a contradiction.

It is $D$ that allows us to build $X$, so there cannot be a Universal Termination Detector$^2$. That is, the halting function is uncomputable.

---

$^1$Remember, $D$ is a Turing machine; it computes the halting function, but is not the halting function itself.

$^2$Note that this is not a proof that the halting function does not exist. The halting function is a well-defined mathematical object. It is a proof that the halting function cannot be computed by any Turing machine.