

The Halting Problem is Undecidable

We define the language HALT to be the set of all strings of the form $\langle M \rangle w$ such that M halts with input w . HALT is the language which is equivalent to the halting problem.

Theorem 1 HALT is not decidable.

Proof: By contradiction. Suppose HALT is decidable. Let D be a machine which implements the following program:

```
read a machine description  $\langle M \rangle$ .
if  $M$  halts with input  $\langle M \rangle$ 
  run forever.
else
  halt.
```

We now run D with input $\langle D \rangle$. One of the following two cases must hold.

Case 1. D halts with input $\langle D \rangle$. That means that, when D reads $\langle D \rangle$, it runs forever, hence D does not halt with input $\langle D \rangle$, contradiction.

Case 2. D does not halt with input $\langle D \rangle$. That means that, when D reads $\langle D \rangle$, it halts, hence D halts with input $\langle D \rangle$, contradiction.

In either case, we obtain a contradiction, hence HALT is undecidable. █

Theorem 2 HALT is recognizable.

Proof: The following program P recognizes HALT.

```
read  $\langle M \rangle w$ 
run  $M$  with input  $w$ .
if  $M$  halts with input  $w$ 
  accept  $\langle M \rangle w$ .
```

Thus, P accepts every member of HALT, but no other string. █

Note that the program will run forever if $\langle M \rangle w \notin \text{HALT}$. HALT is recursively enumerable, since every language recognized by a machine is recursively enumerable.

Theorem 3 Every recursively enumerable language is decidable.

Proof: Let L be a language enumerated by a machine M . Let w_1, w_2, \dots be the enumeration of L given by M . For any $n \geq 0$, let $L[n]$ be the finite set of all members of L of length no greater than n . Each member of $L[n]$ must be w_i for some i . Define $T(n) = \max\{i : w_i \in L[n]\}$. The following program decides L .

```
read  $w$ , let  $n = |w|$ 
If( $M$  writes  $w$  within time  $T(n + 1)$ ), ACCEPT  $w$ .
Else REJECT  $w$ . █
```

Thus, HALT is decidable. But it's undecidable! What's wrong with this proof?