

Halting Problem Proof from Finite Strings to Final States

When we understand that every potential halt decider must derive a formal mathematical proof from its inputs to its final states previously undiscovered semantic details emerge.

When-so-ever the potential halt decider cannot derive a formal proof from its input strings to its final states of Halts or Loops, undecidability has been decided.

The formal proof involves tracing the sequence of state transitions of the input TMD as syntactic logical consequence inference steps in the formal language of Turing Machine Descriptions.

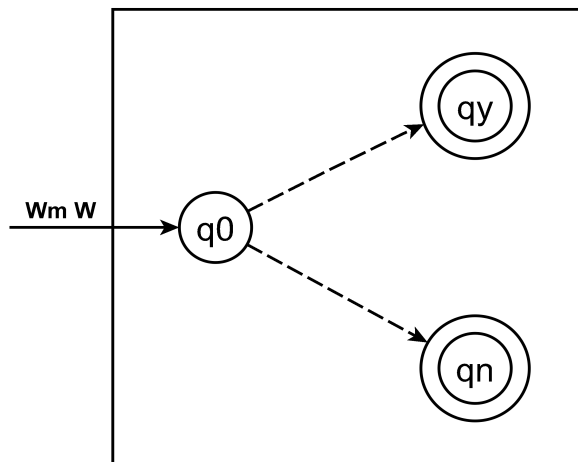
The proof would essentially be a hypothetical execution trace** of the state transition sequences of the input TMD. It cannot be an actual execution trace or the halt decider could have non-halting behavior. **Like step-by-step mode in a debugger.

If the first element of the input finite string pair is a correct TMD then this proof must necessarily proceed from the specified TMD start state through all of the state transitions of this TMD to the halting or looping behavior of this TMD.

The following has been adapted from material from the this book:
An Introduction to Formal Languages and Automata by Peter Linz 1990 pages 318-320

We begin our analysis by constructing a hypothetical halt decider: H.

Figure 12.1 Turing Machine H



The dashed lines proceeding from state (q0) are represented in the text definition as the asterisk \vdash^* wildcard character. These conventions are used to encode unspecified state transition sequences.

Definition of Turing Machine H (state transition sequence)

$H.q0 Wm W \vdash^* H.qy$ // Wm is a TMD that would halt on its input W

$H.q0 Wm W \vdash^* H.qn$ // else

The diagram and the state transition sequence indicate that H begins at its own start state H.q0 and is applied to finite string pair (Wm, W).

Then it proceeds through an unspecified set of state transitions to one of its final states.

H.qy is the final state of H indicating that Wm is a TMD would halt on input W.

H.qn is the final state of H indicating that Wm is Not a TMD that would halt on input W.

In an attempt to provide a valid counter-example proving by contradiction that a Halt Decider cannot possibly exist for every possible TM / Input pair we create Turing Machine \hat{H} by making the following changes to H:

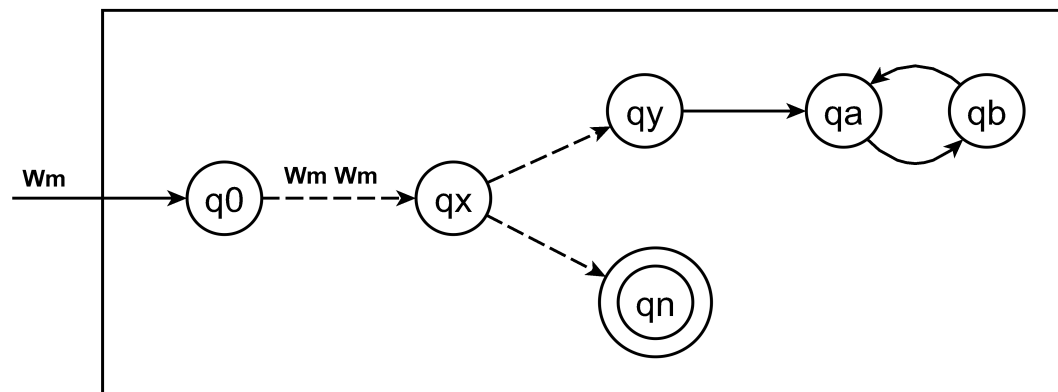
(1) \hat{H} copies its input Wm a its $\hat{H}.q_0$ state and transitions to its $\hat{H}.q_x$ state.

(2) \hat{H} would begin to evaluate Wm Wm at its $\hat{H}.q_x$ state in exactly the same way that H would begin to evaluate its Wm W input at its H.q0 state.

(3) States (qa) and (qb) are appended to existing final state ((qy)) such that any transition to state (qy) will cause \hat{H} to loop.

Since Turing Machine \hat{H} is created by adapting H, it would have exactly the same behavior at its $\hat{H}.q_x$ state as H would have at its H.q0 state.

Figure 12.3 Turing Machine \hat{H}



Definition of Turing Machine \hat{H} (state transition sequence)

$\hat{H}.q_0 W_m \vdash^* \hat{H}.q_x W_m W_m \vdash^* \hat{H}.q_y \infty$

$\hat{H}.q_0 W_m \vdash^* \hat{H}.q_x W_m W_m \vdash^* \hat{H}.q_n$

If Turing Machine H is applied to Turing Machine descriptions [\hat{H}] [\hat{H}] would H transition to H.y or H.n ?

H [\hat{H}] [\hat{H}^2] // We append a “2” to the second \hat{H} for clarity

To answer this question we perform an execution trace on H

(1) H begins at its start state H.q0.

(2) H begins to evaluate what [\hat{H}] would do on its input [\hat{H}^2].

Because H is performing a mathematical proof on finite strings $[H]$ $[H2]$ the required syntactic logical inference chain is simply the state transitions that $[H]$ would make on its input $[H2]$. In simplest terms H is performing a step-by-step debug trace of $[H]$.

Step-by-step debug trace of what $[H]$ would do on its input $[H2]$

- (3) $[H]$ would begin at its start state $[H].q_0$
 - (4) $[H]$ would make a copy of its input $[H2]$, we will call this $[H3]$.
 - (5) $[H]$ would transition to its state $[H].q_x$.
 - (6) $[H]$ would begin to evaluate what $[H2]$ would do on its input $[H3]$.
- Can you see the infinite recursion?

H $[H]$ $[H]$ specifies an infinitely recursive evaluation sequence. Every HP proof by contradiction depends this same infinite recursion. What no one ever noticed before is that the debug trace by the halt decider must abort its evaluation long before it ever reaches the appended infinite loop. Because the formal proof is aborted before ever reaching the final states of H , H is not a decider for H $[H]$ $[H]$.

Because of this every TMD in the infinitely recursive sequence is defined in terms of H each would reject the whole sequence as semantically incorrect before even beginning any halting evaluation, and transition to $H.q_n$.

As further evidence that infinitely recursive evaluation sequence has been overlooked we only need to know that every TMD always requires its own TM/Input pair. For any finite sequence of input H could always decide halting.

If H did not copy its input and instead simply took two inputs then the Halting Decision would be easy. H would transition to its $H.q_y$ state because $H2$ would transition to its $H2.q_n$ state on null input. This would cause H to transition to its $H.q_n$ state deciding halting for $[H]$ $[H2]$.

Copyright 2004, 2006, (2012 through 2018) Pete Olcott

Glossary of Terms

A halt determiner H determines halting by deriving a formal mathematical proof from its input strings (b,c) to its own final states. It does this in the language of Turing Machine Descriptions (shown below).

A halt decider is free to decide halting by always saying yes.

A halt determiner H that determines halting for a single pair of input strings (b,c) is a halt determiner for these input strings.

A halt decider that decides halting for a single pair of inputs does not have a conventional term of the art.

A halt determiner H that cannot derive a formal proof from its input strings (b,c) to its own final states because these input strings derive the pathological self-reference error is called undecidable by H for this input pair (b,c).

A halt decider that cannot decide halting for a single pair of inputs because these inputs derive the pathological self-reference error has no conventional term of the art.

A hypothetical halt determiner H that can decide halting for an infinite set of input strings (b,c) is called a universal halt determiner.

A hypothetical halt decider that can decide halting for an infinite set of input strings is called a halt decider.

A TM H that determines Halting / Looping or the error of pathological self-reference for inputs (b,c) is called a halting determinability determiner.

A TM H that decides Halting / Looping or the error of pathological self-reference for inputs (b,c) is called a halting decidability decider.

Turing Machine Description Language

(Generalized from: TM, The Turing Machine Interpreter by David S. Woodruff)

The following is space delimited 'quintuples', each one of which is a five-symbol Turing Machine instruction. It is written in AWK regular expression syntax.

- 1) Initial State
- 2) Current Tape Symbol
- 3) Next State
- 4) Write Tape Symbol
- 5) Move Tape Head Left or Right

```
--1--    --2--    --3--    --4--    --5--  
[0-9]+ [\x20-\x7e] [0-9]+ [\x20-\x7e] [LR]
```

If we make a fixed number of (hexadecimal) states then The UTM could simply hold a sorted list of `std::strings`. A sorted list would allow state transitions to occur by binary search. This would allow a very simple DFA based syntax checker.

The tape could be constructed using a `std::vector`. Binary Zero end markers would be in the tape alphabet and not the input alphabet. These trigger memory reallocation to enlarge the tape.

Postscript November 2018 (Halting Problem Solution)

The following analysis applies very recent notational simplifications and clarity to the same insights already discussed in the above paper.

Linz 1990 page 233 Definition 9.1

A Turing Machine M is defined by

$$M = (Q, \Sigma, \Gamma, \delta, q_0, \square, F)$$

where

Q is a finite set of internal states,

Σ is a finite set of symbols called the input alphabet,

Γ is a finite set of symbols called the tape alphabet,

δ is the transition function,

$\square \in \Gamma$ is a special symbol called blank,

$q_0 \in Q$ is the internal state,

$F \subseteq Q$ is the set of final states.

Provability() is defined in the language of Turing_Machine_Descriptions: \vdash_{TMD}

$\forall tm \in \text{TMD} \forall i \in \text{Finite_Strings}(\Sigma)$

$\text{Halts}(tm, i) \leftrightarrow (\vdash_{\text{TMD}} \text{unify_with_occurs_check}(tm, i) \wedge \vdash_{\text{TMD}} \text{halts}(tm, i))$

The higher order logic predicate: $\text{Halts}(tm, i)$ is materially equivalent to the existence of a formal proof in the language of Turing Machine Descriptions to the final states of:

(a) $\text{halts}(tm, i)$ // (tm, i) halts

(b) $\text{unify_with_occurs_check}(tm, i)$ // valid input (see details below)

Copyright 2018 Pete Olcott

254 Chapter 10 The Relation of Prolog to Logic

Programming in Prolog Using the ISO Standard Fifth Edition by Clocksin and Mellish

...

Most Prolog systems will allow you to satisfy goals like:

`equal(X, X).`

`?- equal(foo(Y), Y).`

that is, they will allow you to match a term against an uninstantiated subterm of itself.

In this example, `foo(Y)` is matched against `Y`, which appears within it. As a result, `Y` will stand for `foo(Y)`, which is `foo(foo(Y))` (because of what `Y` stands for), which is `foo(foo(foo(Y)))`, and so on. So `Y` ends up standing for some kind of infinite structure.

...

the built-in predicate `unify_with_occurs_check/2` ... fails if an occurs check detects an illegal attempt to instantiate a variable.