How to Compute Halting

Eric C.R. Hehner

Department of Computer Science, University of Toronto
hehner@cs.utoronto.ca

Abstract: A consistently specified halting function may be computed.

Halting Problem

Here is the function header and specification of a Pascal function \texttt{halts} to compute the termination status of Pascal procedures; the function body is absent. Following that, we have a Pascal procedure \texttt{diag} in its entirety.

\begin{verbatim}
function halts (p, i: string): string;
{ return 'yes' if \( p \) represents a Pascal procedure with one string input parameter }
{ whose execution terminates when given input \( i \); }
{ return 'no' if \( p \) represents a Pascal procedure with one string input parameter }
{ whose execution does not terminate when given input \( i \); }
{ return 'not applicable' if \( p \) does not represent a Pascal procedure }
{ with one string input parameter }

procedure diag (s: string);
begin
  if halts (s, s) = 'yes' then diag (s)
end
\end{verbatim}

We assume there is a dictionary of function and procedure definitions that is accessible to \texttt{halts}, so that the call \texttt{halts ('diag', 'diag')} allows \texttt{halts} to look up 'diag', and subsequently 'halts', in the dictionary, and retrieve their texts for analysis. What should the result of \texttt{halts ('diag', 'diag')} be? This is a question about the specification of \texttt{halts}. Let's look at each possibility in turn.

Should the result of \texttt{halts ('diag', 'diag')} be 'not applicable'? Syntactically, \texttt{diag} is a procedure; to determine that \texttt{halts} is being used correctly within \texttt{diag}, we need only the header for \texttt{halts}, not the body, and we have the header. Semantically, it is a procedure; to determine the meaning of the call to \texttt{halts} within \texttt{diag}, we need only the specification of \texttt{halts}, not its implementation, and we have the specification. (That important programming principle enables a programmer to call procedures written by other people, knowing only the specification, not the implementation. It also enables a programmer to change the implementation of a procedure, but still satisfying the specification, without knowing where and why the procedure is being called.) So there is nothing wrong with the definition of \texttt{diag}, and the result should not be 'not applicable'.

Should the result of \texttt{halts ('diag', 'diag')} be 'yes'? If so, the semantics of \texttt{diag ('diag')} is nontermination, so it should be 'no'.

Should the result of \texttt{halts ('diag', 'diag')} be 'no'? If so, the semantics of \texttt{diag ('diag')} is termination, so it should be 'yes'.

We have ruled out all possibilities. This is inconsistent. Therefore \texttt{halts} cannot be programmed according to its specification.
How to Compute Limited Halting

It is inconsistent to ask for a Pascal function to compute the halting status of all Pascal procedures. But we can ask for a Pascal function to compute the halting status of some Pascal procedures. For example, a function to compute the halting status of just the two procedures

\begin{verbatim}
procedure loop (s: string); begin loop (s) end
procedure stop (s: string); begin end
\end{verbatim}

is easy. Perhaps we can ask for a Pascal function to compute the halting status of all Pascal procedures that do not refer to this halting function, neither directly nor indirectly. Here is its header, specification, and a start on its implementation.

\begin{verbatim}
function halts1 (p, i: string): string;
{ return 'yes' if p represents a Pascal procedure with one string input parameter }
{ that does not refer to halts1 (neither directly nor indirectly) }
{ and whose execution terminates when given input i; }
{ return 'no' if p represents a Pascal procedure with one string input parameter }
{ that does not refer to halts1 (neither directly nor indirectly) }
{ and whose execution does not terminate when given input i; }
{ return 'maybe' if p represents a Pascal procedure with one string input parameter }
{ that refers to halts1 (either directly or indirectly); }
{ return 'not applicable' if p does not represent a Pascal procedure }
{ with one string input parameter }
begin
if (p does not represent a Pascal procedure with one string input parameter)
then halts1 := 'not applicable'
else if (p refers to halts1 directly or indirectly)
then halts1 := 'maybe'
else (return halting status of p, either 'yes' or 'no')
end
\end{verbatim}

The first case checks whether \( p \) represents a (valid) procedure exactly as a Pascal compiler does. The middle case looks like a transitive closure algorithm, but it is problematic because, theoretically, there can be an infinite chain of calls. Thus we may be able to compute halting for this limited set of procedures, but not determine whether a procedure is in this limited set. The last case may not be easy, but at least it is free of the reason it has been called incomputable: that it cannot cope with

\begin{verbatim}
procedure diag1 (s: string); { execution terminates if and only if halts1 (s, s) ≠ 'yes' }
begin
if halts1 (s, s) = 'yes' then diag1 (s)
end
\end{verbatim}

Procedure \( \text{diag1} \) refers to \( \text{halts1} \) by calling it, so \( \text{halts1} \) is not required to determine the halting status of \( \text{diag1} \). Therefore \( \text{halts1 ('diag1', 'diag1')} = \text{'maybe'} \), and execution of \( \text{diag1 ('diag1')} \) is terminating.

Calling is one kind of referring, but not the only kind. In the specification of \( \text{halts1} \), the name \( \text{halts1} \) appears, and also in the body. These are self-references, whether or not \( \text{halts1} \) calls itself. We exempt \( \text{halts1} \) from having to determine the halting status of procedures containing any form of reference to \( \text{halts1} \); the result is \text{'maybe'}. We might try to circumvent the limitation by writing another function \( \text{halts2} \) that is identical to \( \text{halts1} \) but renamed (including in the specification, the return statements, and any recursive calls).
function \( \text{halts2} \) (\( p, i \): string): string;
{ return 'yes' if \( p \) represents a Pascal procedure with one string input parameter }
{ that does not refer to \( \text{halts2} \) (neither directly nor indirectly) }
{ and whose execution terminates when given input \( i \); }
{ return 'no' if \( p \) represents a Pascal procedure with one string input parameter }
{ that does not refer to \( \text{halts2} \) (neither directly nor indirectly) }
{ and whose execution does not terminate when given input \( i \); }
{ return 'maybe' if \( p \) represents a Pascal procedure with one string input parameter }
{ that refers to \( \text{halts2} \) (either directly or indirectly); }
{ return 'not applicable' if \( p \) does not represent a Pascal procedure }
{ with one string input parameter }
begin
if (\( p \) does not represent a Pascal procedure with one string input parameter)
then \( \text{halts2} := 'not applicable' \)
else if (\( p \) refers to \( \text{halts2} \) directly or indirectly)
then \( \text{halts2} := 'maybe' \)
else (return halting status of \( p \), either 'yes' or 'no')
end;

procedure \( \text{diag2} \) (\( s \): string);
begin
if \( \text{halts2} \) (\( s, s \)) = 'yes' then \( \text{diag2} \) (\( s \))
end

Now \( \text{halts2} \) ('\text{diag2}', 'diag2') = 'maybe' because \( \text{diag2} \) refers to \( \text{halts2} \) by calling it. And \( \text{halts1} \) ('\text{diag2}', 'diag2') = 'yes' because \( \text{diag2} \) does not refer to \( \text{halts1} \) and execution of \( \text{diag2} \) ('\text{diag2}') terminates. Even though \( \text{halts2} \) is obtained from \( \text{halts1} \) just by renaming it, they produce different results when given the same input.

**How to Compute Unlimited Halting**

In Pascal, as originally defined, identifiers cannot contain underscores. I now define a new programming language, Pascal\_\_, which is identical to Pascal except that identifiers can contain underscores. Pascal\_\_\_ is a larger language than Pascal, but no more powerful: they are both Turing-Machine-equivalent. In this new language, perhaps we can write a function named \( \text{halts\_\_} \) that determines the halting status of all Pascal procedures. Pascal procedures are syntactically prevented from referring to \( \text{halts\_\_} \), so the problem of determining whether a Pascal procedure refers to \( \text{halts\_\_} \) disappears, along with the 'maybe' option.

function \( \text{halts\_\_} \) (\( p, i \): string): string;
{ return 'yes' if \( p \) represents a Pascal procedure with one string input parameter }
{ whose execution terminates when given input \( i \); }
{ return 'no' if \( p \) represents a Pascal procedure with one string input parameter }
{ whose execution does not terminate when given input \( i \); }
{ return 'not applicable' if \( p \) does not represent a Pascal procedure }
{ with one string input parameter }
begin
if (\( p \) does not represent a Pascal procedure with one string input parameter)
then \( \text{halts\_\_} := 'not applicable' \)
else (return halting status of \( p \), either 'yes' or 'no')
end

If it is possible to write a Pascal function to compute the halting status of all Pascal procedures that do not refer to this halting function, then by writing in another language, we can compute
the halting status of all Pascal procedures.

There is an argument that, at first sight, seems to refute the possibility of computing the halting status of all Pascal procedures just by programming in another language. Suppose that in writing `halts_` we do not use any underscores in any other identifiers, and we do not use the identifier `halts`. Then we can easily obtain a Pascal function `halts` just by deleting the underscore from the `halts_` identifier. We thus obtain a Pascal function with the same functionality: `halts (s) = halts_ (s)` for all `s`. But there cannot be a Pascal function that computes the halting status of all Pascal procedures. Therefore, the argument concludes, there cannot be a Pascal function to do so either.

As compelling as the previous paragraph may seem, it is wrong. We have already seen that renaming `halts1` to `halts2` produces a function with different results. The phenomenon can be understood in everyday experience. If I say “My name is Eric Hehner.” I am telling the truth. If Margaret Jackson says exactly the same words, she is lying. When I say it, there is a self-reference; when Margaret Jackson says it, there is no self-reference. The truth of that sentence depends on who says it.

Here is a simple example of the failure of program translation: a Pascal procedure that prints its own name.

```
procedure A_: { this procedure prints its own name }
begin print (A_) end
```

Translating this procedure to Pascal, we face a dilemma. We could translate it as

```
procedure A; { this procedure prints its own name }
begin print (A) end
```

arguing that the two procedures have the same output, but clearly this translation does not preserve the intention. The Pascal procedure `A_` meets its specification; the Pascal translation `A` does not. Or we could translate it as

```
procedure A; { this procedure prints its own name }
begin print (A') end
```

arguing that we have preserved the intention, but clearly the two procedures do not have the same output. Translating from `halts_` to `halts` has the same problem. We cannot preserve the intention because the specification at the head of `halts_`, which is perfectly reasonable for a Pascal function, becomes inconsistent when placed at the head of a Pascal function. If we just use the same Pascal procedure but delete the underscores, we obtain a Pascal procedure that no longer satisfies the specification.

There is another argument that, at first sight, also seems to refute the possibility of computing the halting status of all Pascal procedures just by programming in another language. In Pascal, we can write an interpreter for Pascal programs. So if we could write a halting function `halts_` in Pascal for all of Pascal, we could feed the text of `halts_` to this interpreter, and thus obtain a Pascal function to compute halting for all Pascal procedures. But there cannot be a Pascal function that computes the halting status of all Pascal procedures. Therefore, the argument concludes, there cannot be a Pascal function to do so either.

The reason this argument fails is the same as the reason the previous argument fails. The interpreter interpreting `halts_` is just like the translation of `halts_` into Pascal by deleting underscores. The interpreter interpreting `halts_` can be called by another Pascal program;
halts cannot be called by a Pascal program. That fact materially affects their behavior. Pascal program halts can be applied to a Pascal procedure $d$ that calls the interpreter interpreting halts applied to $d$, and it will produce the right answer. But the interpreter interpreting halts applied to $d$ calls the interpreter interpreting halts applied to $d$, and execution will not terminate.

### the Barber

A town named Russellville consists of some men (only men). Some of the men shave themselves; the others do not shave themselves. A barber is a person who shaves all and only those men in Russellville who do not shave themselves. There is a barber: a man named Bertrand who lives just outside the town, in the Greater Russellville Area called Russellville. Without any difficulty, he satisfies the specification of barber.

One of the men in Russellville, whose name is Bertrand, decided that there is no need to bring in a barber from outside town. Bertrand decided that he could do the job. He would shave those men whom Bertrand shaves, and not shave those men whom Bertrand does not shave. If Bertrand is fulfilling the role of barber, then by doing exactly the same actions as Bertrand, Bertrand reasoned that he would fulfill the role of barber. But Bertrand is wrong; those same actions will not fulfill the role of barber when Bertrand performs them. To be a barber, Bertrand has to shave himself if and only if he does not shave himself. A specification that is perfectly consistent and possible for someone outside town becomes inconsistent and impossible when it has to be performed by someone in town.

And so it is with the halting specification, and for the same reason. For Bertrand, the barber specification has no self-reference; for Bertrand, the barber specification has a self-reference. For halts, the halting specification has no self-reference; for halts, the halting specification has a self-reference (indirectly through $\text{diag}$ and other procedures that call $\text{halts}$).

### Conclusion

By weakening the specification a little, reducing the domain from “all procedures” to “all procedures that do not refer to the halting function”, we obtain a specification that may be both consistent and computable. Equivalently, we may be able to compute the halting status of all procedures in a Turing-Machine-equivalent language by writing a halting function in another Turing-Machine-equivalent language, assuming that the procedures of the first language cannot refer to the halting function written in the second language. In any case, we do not yet have a proof that it is impossible.