What is a Turing Machine?

- Named after (not by) Alan M. Turing.
- Perhaps the most important computational model, from a theoretical viewpoint.
- Simplicitic, yet universal.

Relation to Previous Machines

- Like a FSA but
  - can write as well as read its tape
  - can move in both directions
- Like a PDA with 2 stacks
- More powerful than a locomotive, able to ...

Turing Machine Diagram

Turing Machine (tm) Components
(a la John Martin’s book)

- \((Q, \Sigma, \Gamma, q_0, \delta)\)
- Q: finite set of control states
- \(\Sigma\): input alphabet
- \(\Gamma\): tape alphabet (\(\Sigma \subseteq \Gamma\))
- \(q_0\): initial control state
- \(\delta\): transition function (next page)

Turing Machine (TM) Components
(a la John Martin’s book)

- Let \(\square\) represent "blank", which is not in \(\Gamma\).
- Let \(\varnothing = \square \cup \{\square\}\).
- Let \(h_a, h_r\) be states not in \(Q\) (halting-accepting and halting-rejecting respectively)
- Let \(Q' = Q \cup \{h_a, h_r\}\)
- \(\delta\): transition function
  - \(\delta: Q \times \varnothing \rightarrow Q' \times \{R, L, S\}\)
- R, L, and S represent head motion (right, left, and stationary, respectively)
TM Transition Function

- For each q ∈ Q and X ∈ Γ,
  \[ d(q, X) = (q', Y, M) \]
  where q' ∈ Q, Y ∈ Γ, and M ∈ \{R, L, S\}

- means
  - the machine is in state q and reading X
  - the machine goes to state q', writes Y, and moves
  - Left (L), Right (R), or Not at all (S)

Abbreviations

- No transitions are specified from the two halting states: h_#, h_r.
- To reduce the size of a specification, if a transition is unspecified for a given (q, X), assume it goes to h_r, rewrites X, and does not move.

TM Rule Notation

- Instead of
  \[ d(q, X) = (q', Y, M) \]
we may choose to write
  \[ q, X \rightarrow q', Y, M \]
- This is called "5-tuple" notation.

TM Configuration (i.e. complete state)

- We show the control state q ∈ Q and the tape \[ \square \square \square^* \].
- All but a finite number of symbols are non-blank \[ \square \].
- We underline the symbol over which the tape is positioned, e.g.
  \[ (q_0, 0110\underline{110}) \]
- Outside this string, all tape symbols are \[ \square \].
- Hence the tape may be represented by two finite strings, both elements of \[ \square^* \):
  - The symbols to the left of the head.
  - The symbols to the right of, and including the head.

General Moves

- A machine makes one of the following transitions
  - \[ (q, wXyz) \rightarrow (q', wX'yz) \text{ if } d(q, X) = (q', X', R) \]
  - \[ (q, wXyz) \rightarrow (q', wX'yz) \text{ if } d(q, Y) = (q', Y', L) \]
  - \[ (q, wXyz) \rightarrow (q', wX'yz) \text{ if } d(q, X) = (q', X', S) \]

Standard Starting Configuration

- The machine starts in \[ q_0 \].
- The head is positioned at the over a blank just to the left of the input, which consists of symbols in \[ \square \] (input alphabet) only.
  \[ (q_0, \square X) \text{ where } X \in \square^* \]
Halting Configuration

- The machine halts in one of h_0 or h_r, indicating acceptance or rejection of the original input respectively. No transitions are defined from these states.
- It is possible that the machine never halts, in which case we say it "diverges".

Legend for the Previous TM

- q_0: Just Starting
- q_1: Looking for a, reject anything else.
- q_2: Looking for b, skipping over a’s.
- q_3: Looking for c, skipping over b’s, reject anything else.
- q_4: Moving left after matching one a, b, and c, looking for blank.
- q_5: Moving left after deciding rejection.
- q_6: Moving left after deciding rejection.
- h_0: Halt and accept.
- h_1: Halt and reject.
Conclusion:

- There is a Turing machine that accepts a non-context free language.

Fact:

- For every context-sensitive language L, there is a Turing Machine that accepts L.

- Proof: A TM can be programmed to generate all derivable strings from a type 1 grammar, up to the length of the input. Those strings are strictly increasing in length.

  At some determinable point, no new strings of length less-than-or-equal-to the length of the input will be generated. The TM can determine acceptance or rejection by then.

Another Possible Rendering

/cs/cs81/tm provides a TM simulator

There is documentation (tm.doc) and examples:

add1.tm  Adds 1 to a binary numeral
add.tm   Adds two binary numerals

add1.tm:

<table>
<thead>
<tr>
<th>start</th>
<th>_</th>
<th>left</th>
<th>add1</th>
</tr>
</thead>
<tbody>
<tr>
<td>add1</td>
<td>0</td>
<td>1</td>
<td>right</td>
</tr>
<tr>
<td>add1</td>
<td>_</td>
<td>1</td>
<td>right</td>
</tr>
<tr>
<td>add1</td>
<td>1</td>
<td>0</td>
<td>left</td>
</tr>
<tr>
<td>end</td>
<td>0</td>
<td>0</td>
<td>right</td>
</tr>
<tr>
<td>end</td>
<td>1</td>
<td>1</td>
<td>right</td>
</tr>
</tbody>
</table>

How to use or get JFlap.

- Go to Susan Rodger’s (the author’s) page:
  
  www.cs.duke.edu/~rodger/tools/jflap/

  - Either:
    - Download the .jar file and run it on your machine, or
    - Use the web applet provided.

  - JFlap simulates: FA, NFA, PDA, TM, etc.

Possible TM behaviors on a given input

- Halt in the accepting state.
- Halt in the rejecting state.
- Fail to halt (**diverge**).
Acceptance vs. Recognition

- This is technical (but very important) distinction, and the terminology is not universal.
- We will use that of John Martin:
  - A TM accepts a language L if for every string in the language, the machine halts in the accepting state.
  - If L is not in the language, it may halt in the rejecting state, or may diverge.
- A TM recognizes the language if it accepts the language and always halts (never diverges).

Acceptance vs. Recognition

- One reason the distinction is important is that there are languages that are:
  - accepted by some TM
  - not recognized by any TM

Terminology

- A language that is recognized by some TM is called "recursive".
- A language that is accepted by some TM is called "recursively enumerable" (R.E.).
- So recursive R recursively enumerable, by definition (but not conversely).

Note on Terminology

- "recursive" doesn't mean that the language has some kind of self-referential structure.
- Rather this term comes from the "recursive functions" model of Godel and Kleene. These functions can be shown to be equivalent to Turing machines.

More on the Recursive vs. R.E. Distinction

- If a language is recursive, then its complement is also.
- This can be seen by simply swapping the accepting and rejecting halting states in a TM that recognizes the language.

More on the Recursive vs. R.E. Distinction

- If a language is R.E. and its complement is R.E., then the language is recursive.
- How to show this?
TMs Computing Functions

- Rather than accepting a language, we can use a TM to compute a function:
  - The machine starts with some string $x$ on its tape initially.
  - The machine halts with some string $y$ on its tape finally.
  - Then the corresponding function $f$ would have $f(x) = y$

Characteristic Functions

- A characteristic function is just a way to represent a subset of a given universe.
- If $W$ is the universe, and $S \subseteq W$, then
  $$c_S(x) = \begin{cases} 
1 & \text{if } x \in S \\
0 & \text{otherwise}
\end{cases}$$
- $c_S: W \rightarrow \{0, 1\}$ is the characteristic function of $S$.
- So sets are equivalent to these special kind of functions.

TMs Computing Partial Functions

- For some initial strings $x$ the TM might diverge.
- More generally, we say the machine computes a partial function, and $f(x)$ is undefined for such $x$.
- An ordinary function is a special case of a partial function, one which is nowhere undefined. We call such a partial function a total function to emphasize the distinction.

Strings as Natural Numbers

- We can view any string over a fixed finite alphabet as an encoding of a natural number.
- The encoding is a 1-1 correspondence.
- Example: Consider the alphabet $\{a, b\}$.
  - $0 \rightarrow \epsilon$
  - $1 \rightarrow a$
  - $2 \rightarrow b$
  - $3 \rightarrow aa$
  - $4 \rightarrow ab$
  - $5 \rightarrow ba$
  - $6 \rightarrow bb$
  - $7 \rightarrow aaa$
  - $\ldots$ (called the “2-adic” encoding)

Consequence

- Turing machines can be viewed as computing with (natural) numbers just as well as they can with strings.

Turing Machines as Enumerators

- Suppose $M$ is a Turing Machine where we view the input as a natural number and the output as a string over alphabet $\{\}$.
- Then $M$ enumerates a subset of $\{\}^*$, in the sense that it computes a function
  $$f(i) = \text{some element of } \{\}^*, \text{ for each } i \in \mathbb{N}$$
- The set enumerated is $\{f(0), f(1), f(2), \ldots\}$, which of course is a language.
We will eventually see:

- A language is recursively enumerable iff there is a TM that enumerates it.

(Recall that the definition of R.E. is that there is a TM that accepts it.)

---

**Encoding TMs as Strings**

- We know that a TM can be described fully by a list of transition rules, of the form:
  - \( q_0 \rightarrow \rightarrow q_0 \), \( R \)
  - \( q_0 \rightarrow \rightarrow q_1 \), \( L \)
  - \( q_0 \rightarrow \rightarrow q_0 \), \( R \)
  - \( q_0 \rightarrow \rightarrow q_1 \), \( L \)
  - \( \ldots \)

- from earlier examples.

- We can devise a way to encode an arbitrary set of such rules into a fixed alphabet.

- Although the number of states and tape symbols can be arbitrary, we can encode these by using strings of symbols, say \( \{a, b\} \), and concatenate the symbols to get rules, and concatenate rules to get machines.

---

**Fact:**

- There are functions that cannot be computed by any TM.

- The simple way to see this is to use the concept of countability.

---

**Definitions:**

- Two sets are **equipotent** if there is a 1-1 correspondence between them. (Equipotence is sort of an equivalence relation.) Write \(|A| = |B|\) to denote that \( A \) and \( B \) are equipotent.

- A set is **infinite** if it is equipotent with a proper subset of itself, otherwise it is **finite**.

- A set is **countably-infinite** if it is equipotent with \( \mathbb{N} \), the set of natural numbers.

- Example: The set of all TMs is countably infinte.

- A set is **countable** if it is finite or countably-infinite; otherwise is **uncountable**.

---

**Further Definitions:**

- Write \(|A| < |B|\) to denote that there is a 1-1 mapping \( f \) from \( A \) into \( B \). This means that \( B \) is “at least as big” as \( A \).

- It can be shown (Schröder-Bernstein theorem) that \(|A| < |B|\) and \(|B| < |A|\) together imply \(|A| = |B|\).

- Write \(|A| < |B|\) to mean \(|A| < |B|\) but not \(|B| < |A|\).

- Example: \( A \) is **countable** and \( B \) is **uncountable**, then \(|A| < |B|\).
Theorem of Cantor

- The set of functions of the form \( w \mapsto \{0, 1\} \) is uncountable.
- **Proof by contradiction:**
  Suppose that the set of such functions is countable.
  This means that there is an enumeration of all such functions:
  \( f_0, f_1, f_2, \ldots \)
- Define the function \( f^*(i) = 1 \) if \( f_i(i) = 0 \), and 0 otherwise.

Proof of Theorem of Cantor (cont’d)

- \( f^*(i) = 1 \) if \( f_i(i) = 0 \), and 0 otherwise.
- Clearly \( f^*: \mathbb{N} \mapsto \{0, 1\} \), so it must be in the list somewhere.
  That is, there is a \( j \) such that \( f^* = f_j \).
- Then we have
  \[
  f^*(j) = 1 \text{ if } f_j(j) = 0, \text{ and 0 otherwise}
  \]
  which is to say
  \[
  f_j(j) = 1 \text{ if } f_j(j) = 0, \text{ and 0 otherwise}
  \]
  which is absurd.

Diagonalization Proofs

- The preceding type of proof is known as proof by “diagonalization”, because the function \( f^* \) can be seen as being derived from the “diagonal” in a listing of all the characteristic functions (as if infinitely-long bit vectors).

Corollary

- There are functions of the form \( w \mapsto \{0, 1\} \) that are not computed by any Turing machine.
- Proof: The set of such functions was shown uncountable. But the set of Turing machines were shown countable.
  So there functions computable by Turing machines (interpreting acceptance as 1 and rejection as 0) form a proper subset of the functions of the indicated form.

Unsolvability of the Halting Problem

- We showed that there are non-Turing computable functions, but this was non-constructive; we didn’t exhibit such a function.
- Now we will:
  The partial function \( h \) given by
  \[
  h(x_i) = \begin{cases} 
  1 & \text{if } T_j \text{ diverges on } x_i \\
  \text{undefined (diverges)} & \text{if } T_j \text{ defined on } x_i 
  \end{cases}
  \]
  is not computable by any Turing machine.

Proof of Unsolvability of the Halting Problem

- Suppose that \( h \) is computable by some TM, say \( T_x \).
  Then from the definition of \( h \):
  \[
  T_x(x_i) = \begin{cases} 
  1 & \text{if } T_j \text{ diverges on } x_i \\
  \text{undefined (diverges)} & \text{if } T_j \text{ defined on } x_i 
  \end{cases}
  \]
  But if we supply the corresponding argument \( x_i \) :
  \[
  T_x(x_i) = \begin{cases} 
  1 & \text{if } T_j \text{ diverges on } x_i \\
  \text{undefined (diverges)} & \text{if } T_j \text{ defined on } x_i 
  \end{cases}
  \]
  which is absurd, because no machine can both diverge and give answer 1.
Corollary

- These functions are not computable:
  \[ g(j) = \begin{cases} 
  1 & \text{if } T_j \text{ diverges on } x_j \\
  0 & \text{if } T_j \text{ defined on } x_j 
  \end{cases} \]
  \[ f(j, k) = \begin{cases} 
  1 & \text{if } T_j \text{ diverges on } x_k \\
  \text{undefined} & \text{if } T_j \text{ defined on } x_k 
  \end{cases} \]

- Proof: If these were computable, it can be argued that \( h \) would be computable as well, by a simple modification of the TM that computes \( g \) or \( f \).

Variants on Turing Machines

- The next few slides present variations that, at first glance, might be either simpler or more powerful than our TM model.
  
  - Upon further reflection, they can be seen to be equivalent to our model.

TM without S (stationary) Move

- This is simpler, but \( S \) can be simulated by a combination of \( L \) and \( R \) moves.

TM that can Write or Move but not both

- This is simpler, but the more complex original machine can be simulated by a finite sequence of steps of the restricted machine.

TM with Added Registers

- Some finite number of **registers** is added, each with a finite set of possible values.
  - The transitions can query the registers for values and change their contents.
  - Everything can be simulated by the enlarging control states.
TM with Non-recursive Subroutines

- As long as the subroutines are not recursive, they can be simulated with added control states to remember the return state.
- These subroutines can also take arguments provided the values are drawn from a finite set.

TM with Multi-track Tape

- This machine has the tape divided into separate tracks.
- Each transition specifies which track being read and written.
- Tracks can be simulated by taking the tape alphabet to be the Cartesian product of the track alphabets.
- Another approach is to interleave the cells of the tracks.

TM with Semi-infinite Tape

- The tape is extendible on one side only.
- This is simpler, but a two-infinite tape can be simulated by folding the latter at some arbitrary point on the tape into two "tracks", each simulating one direction of tape expansion.

TM with Insertible Cells

- Rather than allowing extension only by means of writing on formerly blank cells at the tape extremities, this machine has a special move that inserts a new cell just to the right (say) of where the head is. The symbol written would go in that new cell and the head would stay put.
- Other variations are possible.

TM with Multiple Tapes

- This machine has multiple tapes, rather than a single tape. Each tape has its own head.
- An n-tape machine can be simulated by a 2n-track machine, where half of the tracks are devoted to keeping the head position for the corresponding tape.

TM with Recursive Subroutines

- As long as the return is to specific control states, recursion can be carried out by allocating a separate tape to implement a stack of return states.
TM with 2-Dimensional Tape

- This machine has a 2-D tape and moves can be Up or Down as well as L, R, S.
- To simulate a 2-D tape, divide the 2-D plane into 4 quadrants and use a separate track for each.
- Each track will contain strips representing the non-blank portions of the corresponding quadrant.
- The simulating machine would need to move cells to add to a strip.
- It would add cells to add a new strip.

Non-Deterministic Turing Machines

- A TM that specifies more than one transition for a given control state and read-symbol is non-deterministic (NTM).
- An NTM accepts an input x if some sequence of moves leads to the halting accepting state.
- An NTM recognizes an input x if it accepts x and every sequence of moves leads to the halting accepting state.

Non-Deterministic Turing Machines

- The language \( L(M) \) accepted by an NTM M is the set of all x that are accepted by M.
- If M halts on every input, and recognizes exactly the strings in L, then M is said to recognize L.

Type 0 Languages are R.E.

- Recall the definition of type 0 grammar. Their productions have an unrestricted LHS and RHS.
- Theorem: If L is a Type 0 language, then there is a TM accepting L.
- Proof: Let G be a grammar for L. Construct an NTM that works as follows:
  - With input x, allocate the remainder of the tape to simulate a derivation in G. This part of the tape starts with just S.
  - Non-deterministically choose a rule in G and rewrite the working area according to that rule.
  - After applying a rule compare with x. If x has been generated, then accept.
R.E. Languages are Type 0

- **Theorem**: If $L$ is accepted by a TM, then $L$ is a Type 0 language.
- **Proof**: Let $M$ be a TM. We want to construct a grammar $G$ that generates exactly the terminal strings that $M$ accepts:
  - $G$ will generate two adjacent copies of every string in $L^*$ non-deterministically.
  - $G$ will have productions corresponding to each transition of $M$.
  - Once $G$ generates a pair of strings in $L^*$, $G$ will simulate the behavior of $M$ on one of the copies.
  - If $M$ accepts the string, then $M$ will erase all symbols but the other copy, leaving it as a terminal string.

Linear-Bounded Automata

- A Linear-Bounded Automaton (LBA) is a non-deterministic Turing machine with its tape a fixed length. There are end-marker symbols that it can use to detect the boundary.
- Analogous to the previous result can be proven:
  - **Theorem**: A language $L$ is accepted by an LBA iff $L$ is context-sensitive (type 1).

Universal Turing Machines

- There are other uses of strings that describe Turing machines besides unsolvability arguments.
- Such a string is effectively a **program** for a Turing machine.
- A Turing machine that interprets such a program to carry out the actions specified is called a **Universal Turing Machine** (UTM).

Universal Turing Machines

- UTMs can be shown to exist by constructing them.
- Think about what would be required:
  - The tape has to hold the tape of the machine being simulated.
  - The tape has to hold the program of the machine being simulated.
  - The program must be laid out in such a way that the necessary markers can be inserted to keep track of the current state, etc.
  - All this is possible, if somewhat laborious to construct.
- Whether a machine is universal will depend on the particular encoding used.

Specific Universal Turing Machines

- The first was constructed by Turing himself.
- Shannon showed how a UTM could be converted to a 2-symbol machine and to a 2-state machine (not at the same time), by augmenting states or symbols, respectively.
- Minsky (1960) gave a 7-state 6-symbol machine.
- Watanabe (1961) gave an 8-state 5-symbol machine.
- Minsky (1962) gave a 7-state 4-symbol machine.
- Rogozhin (1996) gave a 4-state 6-symbol machine, among others.

A Specific Universal Turing Machine

- 8-state, 5 symbol UTM published by Watanabe in 1961
- To understand this, you would need to understand the encoding used.

<table>
<thead>
<tr>
<th>$s$</th>
<th>$t$</th>
<th>$a$</th>
<th>$c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
</tr>
<tr>
<td>i</td>
<td>j</td>
<td>k</td>
<td>l</td>
</tr>
<tr>
<td>m</td>
<td>n</td>
<td>o</td>
<td>p</td>
</tr>
</tbody>
</table>

**Table 30**
A Specific Universal Turing Machine

- 2-state, 5 symbol UTM published by Wolfram in 2002

Wolfram expresses belief that there is a 2-state, 3-symbol UTM:
No 2-state, 2-symbol UTM exists.

Turing machines that can print their own description
- At first this might sound impossible.
- Machines that can print their own description have value in establishing certain theoretical results.
- Such machines (and programs in various languages) have been called "Quines" after Willard Van Orman Quine, a famous logician who wrote on self-referential paradoxes.

Willard Van Orman Quine (1908 - 2000)

There is a TM that can print its own description
- **Proof:**
  - Let C be the contents of an arbitrary region of tape (say one delimited by blanks).
  - **First** construct M, a machine that can read the tape contents C and print the description of a machine M_C that will print C, followed by C itself. Let the machine M_C have initial state q_0.
  - [Note that M does not attempt to "understand" the meaning of C. It simply creates a machine that prints a sequence.]

Behavior of M

<table>
<thead>
<tr>
<th>C (arbitrary)</th>
<th>\text{d(M)} = \text{&quot;description of&quot;}</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>C</td>
</tr>
<tr>
<td>\text{d(M_C)}</td>
<td>C</td>
</tr>
<tr>
<td>(blank)</td>
<td></td>
</tr>
<tr>
<td>where M_C</td>
<td>C</td>
</tr>
</tbody>
</table>

A TM that can print its own description, continued
- Next construct M' a machine that can print the description of the fixed machine M on an initially blank tape, then transfer control to a state named q_0 which is not otherwise in M' (but as we know, is the name of the initial state of M_C).
Behavior of M’

\[ M' \rightarrow d(M) \]

where M is as on the previous diagram.

There is a TM that can print its own description

- Consider the combined program \( M';M \).
- This sequence will:
  - Print a description of M.
  - Print the description of \( M' \) (a machine that prints the description of M) along with the description of M.
- So the description written will be that of \( M';M \).
- Hence this machine prints its own description.

Summary of \( M';M \)

\[ \text{blank} \]

\[ M' \rightarrow d(M) \]

\[ M \rightarrow d(M_{\text{print}}) \rightarrow d(M) \]

But \( d(M_{\text{print}}) \) is \( d(M') \), so the final tape is \( d(M';M) \).

Difficulties Constructing Self-printing Programs

- One difficulty is quoting. To get a literal symbol to print, the normal way is to quote it. But that means that the quote marks have to be printed too. And of course, they will have to be quoted, which means that escape characters may be needed. And the escape characters themselves will have to be printed.
- An escape for this is to use numeric codes that print as characters in lieu of the actual literal symbols. The numeric codes don’t normally have to be quoted.

A C Quine (author unknown)

```
char f[] =
"char f[] =%c%c%s%c;%cmain()

(printf(10,34,f,34,10,10);}%c";
main() {printf((f,10,34,f,34,10,10);}
"
```

Example: Prolog Quine

```
quene:–Q="write(quine),write((:–)),put(81),put(61),put(34),writes(Q),put(34),put(44),writes(Q),put(46),put(13),put(10),write(writes),put(40),put(91),put(72),put(124),put(84),put(93),put(41),write((:–)),write(put),put(40),put(72),put(41),put(44),write(writes),put(40),put(84),put(41),put(46),put(13),put(10),write(writes),put(40),put(91),put(93),put(41),put(46)

writes([H|T]):–put(H),writes(T).
writes([ ]).```

Author: Pekka P. Pirinen (pekka@harlequin.co.uk)
Prolog Quine Execution Log

```
% compiling file /home/keller/quine.pl
% quine.pl compiled in module user, 0.080 sec 3,648 bytes
```

Example: A rex Quine constructed by a Pomona College Student

```
Example:
Arex
Quine
constructed
by
a
Pomona
College
Student

```

For many Quines in many diverse languages, see:

http://www.nyx.net/~gthompso/quine.htm

For other varieties of “signature” programs, see:

http://home.planet.nl/~faase009/Signindex.html

Extended Quines

• A Quine can be extended to have arbitrary other functionality that could be computed in the given framework.

• For example, suppose f: \( \mathbb{N} \rightarrow \mathbb{N} \) is a function in such a framework. We could define \( f'(n) = f(n-1) \) for \( n > 0 \) and have \( f'(0) \) return ("print") the description of \( f' \) itself.

Other Uses of Self-Reference

• Programs, e.g. Turing machine programs, can be assumed to have reference to their own source code (as a built-in constant).

• Equivalently, in an enumeration of the Turing computable functions \( T_0, T_1, T_2, \ldots \) the definition of a function \( T_j \) can involve use of its own index \( j \).

Using Self-Reference to Establish the Unsolvability of the Halting Problem

• Consider the function

\[
h(x_j) = \begin{cases} 
1 & \text{if } T_j \text{ halts on input } x_j \\
0 & \text{otherwise}
\end{cases}
\]

• Suppose \( h \) were computable.

• Let \( f \) be the machine that, with input \( x_j \), runs \( h(x_j) \), where \( x_j \) is the description of \( f \), then if \( h(0) = 1 \) returns diverges intentionally, otherwise returns 1.

• If \( h \) were computable, so would \( f \) be. Now if we give \( x_j \) as the argument to \( f \), then we get a contradiction.
The machine $f$ always gets it wrong. This is because $h$ can’t really exist.

Blank-Tape Halting is Unsolvable

- It turns out that blank-tape halting is just different, not really simpler, as the following argument shows.
- Suppose that there is a Turing machine $B$ that implements function $b$.
- Try to use $B$ to solve the original halting problem.

The Blank-Tape Halting Problem

- This problem is ostensibly simpler to solve than the halting problem:
  
  $b(x_j) = \begin{cases} 
  1 & \text{if } T_j \text{ halts on the all-blank tape} \\
  0 & \text{if } T_j \text{ diverges on the all-blank tape}
  \end{cases}$

  vs.

  $h(x_j) = \begin{cases} 
  1 & \text{if } T_j \text{ halts on input } x_j \\
  0 & \text{if } T_j \text{ diverges on input } x_j
  \end{cases}$

Blank-Tape Halting is Unsolvable

- Let $M$ be a machine such that, with any input $x_j$, $M$ constructs the description of a machine $N$ that
  - writes the description $x_j$ of $T_j$ on its tape, then moves its head to the left end of what it wrote and acts like $T_j$.
- So $N$ on a blank tape is equivalent to $T_j$ on $x_j$.
- So $b(N) = 1$ if $T_j$ halts on $x_j$.
- Therefore $b(M(x_j)) = h(x_j)$, i.e. if a TM for $b$ existed, then so does a TM for $h$ (since a TM for $M$ obviously exists). But the latter can’t be.

The Outer Box Solves the Halting Problem

Key Ideas in the Blank-Tape Problem

- Another problem (the halting problem) was established as unsolvable.
- TM’s can be composed like functions.
- TM’s can manipulate descriptions of other TM’s.
Perspective

- Let F be a family of machines.
- Is there an algorithm that determines, for any machine $M \in F$, whether or not $L(M)$?
- For finite-state machines: Yes.
- For pushdown automata: Yes.
- For Turing machines: No.

The Some-Tape Halting Problem

- Is there an algorithm for determining whether any TM halts on some input tape?
- Try to use the Key Ideas stated earlier.
- Note that it is enough to show that some-tape halting can't be determined for some TM's (It definitely can for others.)

The Some-Tape Halting Problem

- $s(x_j) = \begin{cases} 1 & \text{if } T_j \text{ halts on some tape} \\ 0 & \text{if } T_j \text{ diverges on every tape} \end{cases}$

Some-Tape Halting is Unsolvable

- Let $M$ be a machine such that, with any input $x_j$, $M$ constructs the description of a machine $N$ that
  - erases whatever is on its tape*; then acts like $T_j$ (on a blank tape).
- $N$ on any tape is equivalent to $T_j$ on a blank tape.
- So $s(N) = 1$ iff $T_j$ halts on a blank tape.
- Therefore $s(M(x_j)) = b(x_j)$, i.e. if a TM for $s$ existed, then so does a TM for $b$ (since a TM for $M$ obviously exists). But the latter can't be, as we already showed. (* see Incremental Erasure)

The Outer Box Solves the Blank-Tape Problem

- The idea of a machine erasing whatever is on its tape may seem like an impossibility.
- Consider the technique of incremental erasure:
  - Instead of erasing everything at the very beginning, the machine only erases when it would have have encountered a non-blank square for the first time.
Incremental Erasure

- Overwrite the cell under the head with a blank.
- Introduce a new symbol, say $, and overwrite the cells on either side of the head with it.
- Introduce new rules that behave as follows: Whenever a cell $x$ with $\$ is moved onto from the left for the first time, write a blank in its place and a $\$ to its right, then move back and behave as the original machine would at cell $x$ (which now contains a blank). Do the same for when cells containing $\$ are moved to from the right, except place a $\$ to their left.
- This modified machine’s behavior, in the long run, behaves as the original machine would have on a blank tape.

Gordian-Knot Erasure

- It would not be a violation of the spirit of Turing’s argument to endow the machine with a special “erase” action.
- Add E to the repertoire of moves \{L, R, S\}.
- Doing E erases the entire tape.

Problem Reduction Formalized

- In the preceding couple of proofs we implicitly used the idea of “problem reduction”.
- Let P and Q be problems to be solved.
- Write P \leq Q (P is reducible to Q) if there is an algorithm that will transform any instance of P into an instance of Q in such a way that the answer to the instance of P is the answer to Q.

Examples of Reductions

- Halting Problem \leq Blank-Tape Problem
  - If the blank-tape problem were solvable, then the halting problem would be solvable.

- Blank-Tape Problem \leq Some-Tape Problem
  - If the some-tape problem were solvable, then the blank-tape problem would be solvable.

  Note that this may seem somewhat “backward”: we are reducing the problem that can’t be solved to the problem in question.

Notes on Reductions

- We typically use reductions in the “contrapositive direction”: to show that a problem is unsolvable.
  - If the blank-tape problem were solvable, then the halting problem would be solvable.
  - The halting problem is unsolvable, therefore the blank-tape problem is unsolvable.
  - The second statement is the contrapositive of the first.

- Reduction is transitive:
  - If P \leq Q and Q \leq R, then P \leq R.
Reduction Formalized

- Suppose P and Q are language recognition problems.
- P is the problem: Give an algorithm (TM) that will determine, for any input x ∈ S*, whether or not x ∈ L_P.
- Q is the problem: Give an algorithm (TM) that will determine, for any input y ∈ D*, whether or not y ∈ L_Q.
- P ≤ Q provided there is an algorithm (TM) f: S* → D* such that x ∈ L_P iff f(x) ∈ L_Q.

Double Utility of Reduction

- Reduction can apply to language acceptance problems as well as language recognition problems.
- Recognition: P ≤ Q implies L_P is recognized by a TM provided L_Q is recognized by a TM.
  - i.e. L_P recursively enumerable → L_Q recursively enumerable
- Acceptance: P ≤ Q implies L_P is accepted by a TM provided L_Q is accepted by a TM.
  - i.e. L_P recursively enumerable → L_Q recursively enumerable

Reduction Examples

- P is the halting problem, Q is the blank-tape halting problem:
  - L_P = {x_j | T_j halts on x_j}
  - L_Q = {x_j | T_j halts on []}
  - f(x_j) = description of the machine such that writes x_j on its tape and then behaves like T_j.
  - P ≤ Q
  - Q solvable → P solvable
  - P not solvable → Q not solvable

Reduction Examples

- Q is the blank-tape halting problem, R is the some-tape halting problem:
  - L_Q = {x_j | T_j halts on []}
  - L_R = {x_j | T_j halts on some tape}
  - f(x_j) = description of the machine such that erases its tape, then x_j on its tape and then behaves like T_Q.
  - Q ≤ R
  - R solvable → Q solvable
  - Q not solvable → R not solvable

Pairing Functions

- Often we need to be able to talk about TMs implementing functions of more than one argument.
- An easy way to handle this is to require that the arguments be adjacent on the input tape, separated by a special symbol, say # (blank would work too, but could cause confusion).

Pairing Functions

- We will sometimes use <x, y> to mean a single input representing the pair of inputs.
- It is required that we be able to compute <x, y> from x and y and also be able to compute both x and y from <x, y>.
- Therefore <> as a function must be 1-1, and ideally onto.
The “Recognizes” language
- \textbf{Recognizes} = \{<T_i, x_j> \mid T_i \text{ recognizes } x_j\}
- Both of the following problems are reducible to the \textbf{Recognizes} language:
  - The halting problem.
  - The blank-tape halting problem.
- Therefore the \textbf{Recognizes} language is not recursive.

The “Diverges” language
- \textbf{Diverges} = \{<T_i, x_j> \mid T_i \text{ diverges on } x_j\}
- Both of the following problems are reducible to the \textbf{Diverges} language:
  - \{x_j \mid T_j \text{ diverges on } x_j\} (= “\textbf{SelfDiverge}”)
  - \{x_j \mid T_j \text{ diverges on } []\}.
- \textbf{SelfDiverge} was proved to be not recursively-enumerable.
- Therefore \textbf{Diverges} is not recursively enumerable.

The “Accepts” language
- \textbf{Accepts} = \{<T_i, x_j> \mid T_i \text{ accepts on } x_j\}
- \textbf{Accepts} is recursively enumerable.
- \textbf{Accepts} is the language accepted by any Universal Turing Machine.
- \textbf{Accepts} is not recursive; its complement is not recursively enumerable.

Table of “Solvabilities”
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SelfDiverges</td>
<td>No</td>
<td>No</td>
<td>diagonalization</td>
</tr>
<tr>
<td>SelfAccepts</td>
<td>No</td>
<td>Yes</td>
<td>red to Accepts</td>
</tr>
<tr>
<td>Accepts</td>
<td>No</td>
<td>Yes</td>
<td>UTM, comp</td>
</tr>
<tr>
<td>Diverges</td>
<td>No</td>
<td>No</td>
<td>red from SelfDiverges</td>
</tr>
<tr>
<td>Halts</td>
<td>No</td>
<td>Yes</td>
<td>red from Accepts</td>
</tr>
<tr>
<td>AcceptsEmpty</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AcceptsSome</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AcceptsAll</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AcceptsNone</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Rice’s Theorem
- Could be considered a “meta-theorem” due to its sweeping character.
- Applies to \textbf{functional} questions: languages and functions in the abstract; not to TM \textbf{structural} questions.

Functional vs. Structural
- \textbf{Structural questions}:
  - Is there an algorithm for determining whether a TM has more than 500 control states?
  - Is there an algorithm for determining whether a TM ever uses more than 500 cells of tape on a given input?
- \textbf{Functional questions}:
  - Is there an algorithm for determining whether a TM accepts the empty language?
  - Is there an algorithm for determining whether a TM accepts a finite language?
  - Is there an algorithm for determining whether a TM accepts a regular language?
Encoding Functional Properties as Languages

- Consider some functional property $P$ of Turing machines.
- We can define a property precisely as the language of descriptions (or indices) of TMs that have the property:
- Define $L_P = \{x_j \mid L(T_j) \text{ has property } P\}$.
- Example: $L_{\text{AcceptsEmpty}} = \{x_j \mid \epsilon \in L(T_j)\}$.

Perspective on Functional Properties

- If a language is accepted by some TM, then it is accepted by infinitely many TMs. (Why?)
- For a given functional property, we must get the same categorization of an input, regardless of which of the many equivalent TM's is used.
- So a functional property truly is a property of a language or function, rather than just a particular implementation of the language or function.
- However, the only ways we know (so far) to represent some languages constructively is by machines or grammars.

Trivial Functional Properties

- A functional property is called "trivial" if its language is either $\emptyset$ or $\Sigma^*$.
- In other words, a trivial property holds for no Turing machine or it holds for all of them.
- A non-trivial property, then, holds for some machines, but not all.

Rice’s Theorem

- Any non-trivial functional property of Turing machines is not recursive.
- Put another way:
  No non-trivial functional property of Turing machines is recursive.
- In other words, there is no TM that will determine whether an arbitrary TM:
  - Accepts $\emptyset$
  - Accepts a finite set.
  - Accepts an empty set.
  - Accepts a regular set.
  - etc.

Proof of Rice’s Theorem (1 of 3)

- Suppose $P$ is a non-trivial property.
- The empty language $\emptyset$ either has property $P$ or does not.
- Proceed assuming that $\emptyset$ does not have property $P$.
  This assumption is critical.
- If instead $\emptyset$ does have property $P$, use the following argument on the complementary property, which $\emptyset$ does not have.
- Let $L_P$ be the language of descriptions of TMs that have property $P$.

Proof of Rice’s Theorem (2 of 3)

- The plan is to reduce a non-recursive language $L_u$ to $L_P$, which will imply that $L_P$ is also non-recursive.
- $L_u$ is the language $\{<x, x> \mid T \text{ halts on } x_j\}$, essentially the language accepted by the universal TM ($L_u$ is R.E. but not recursive).
- Let $L$ be an arbitrary language with property $P$ (which must exist, because $P$ is non-trivial).
- Let $M_L$ be a machine accepting $L$.
- The reduction works as follows: For any $<x, x>$, construct a machine $M(x, x)$ with the specifications on the following page:
Proof of Rice’s Theorem (3 of 3)

• M’(x_i, x_j): with input w, temporarily set aside w and simulate T_i on x_j.
  • If the above simulation does not terminate, then L(M’) = ∅, in which case M’ does not have property P.
  • If the above simulation terminates, then simulate M_L on the original input w. If and when M_L terminates, M’ accepts w iff M_L accepts w. In this case, M’ does have property P, because M_L was selected to have it.

• So L_u has been reduced to L_P:
  • T_i halts on x_j iff M’ does not have property P.

Practice expressing the proof as a reduction

• We want a function f:Tapes Æ Tapes such that
  • f is the function that with input <x_i, x_j> constructs machine M’(x_i, x_j) described in the previous slide (M’ with input w ...).

Reflection on Rice’s Theorem

• We were insistent on P being a functional property.
• Where did this assumption get used in the proof?
• Could we have used a different starting point for reduction, e.g. the blank-tape halting problem?

Note on the Definition of R.E.

• Various equivalent definitions are possible for a language L being R.E.
  • The original definition given is that L is the domain of a partial recursive function (i.e. that of the TM that accepts it).
  • The most suggestive definition is that L is either ∅ or is the range of a total recursive function (the function that enumerates it).
  • The Martin book uses another style of enumeration, which I find less elegant than the preceding.

Martin’s (and others) version of Enumeration

• To enumerate an infinite set, a TM must continue forever, writing elements of the set on its tape one after the other, separated by markers.
  • When a marker is written, the TM is not allowed to go back and change any string to the left of the marker, under construction.
The Uniform Halting Problem

- Consider the set
  \( \text{HaltsAll} = \{ x_i | T_i \text{ halts on all inputs} \} \)
  - By now we can surmise that this set is not recursive.
  - But is it even recursively-enumerable?

HaltsAll is Not R.E.

- If HaltsAll = \( \{ x_i | T_i \text{ halts on all inputs} \} \) were R.E. then there would be an always-halting Turing machine \( T^* \) that enumerates it.
  - That is \( \{ T^*(x_0), T^*(x_1), T^*(x_2), \ldots \} = \text{HaltsAll} \).
  - Now define
    \[
    T^{**}(x_j) = \begin{cases} \emptyset & \text{if } T^*(x_j) \neq \emptyset \\ 1 & \text{if } T^*(x_j) = \emptyset \end{cases} \quad (1 \text{ can be any string not } \emptyset)
    \]
    
  - Since \( T^* \) always halts, so does \( T^{**} \). But the description of \( T^{**} \) can't be HaltsAll, for if it is in the list as \( T^*(x_k) \), then by the definition above, \( T^{**}(x_j) \) gives a value other than \( T^*(x_k) \), which is absurd.