

## Chapter udf

# Turing Machine Computations

### mac.1 Introduction

tur:mac:int:  
sec

What does it mean for a function, say, from  $\mathbb{N}$  to  $\mathbb{N}$  to be *computable*? Among the first answers, and the most well known one, is that a function is computable if it can be computed by a Turing machine. This notion was set out by Alan Turing in 1936. Turing machines are an example of *a model of computation*—they are a mathematically precise way of defining the idea of a “computational procedure.” What exactly that means is debated, but it is widely agreed that Turing machines are one way of specifying computational procedures. Even though the term “Turing machine” evokes the image of a physical machine with moving parts, strictly speaking a Turing machine is a purely mathematical construct, and as such it idealizes the idea of a computational procedure. For instance, we place no restriction on either the time or memory requirements of a Turing machine: Turing machines can compute something even if the computation would require more storage space or more steps than there are atoms in the universe.

It is perhaps best to think of a Turing machine as a program for a special kind of imaginary mechanism. This mechanism consists of a *tape* and a *read-write head*. In our version of Turing machines, the tape is infinite in one direction (to the right), and it is divided into *squares*, each of which may contain a symbol from a finite *alphabet*. Such alphabets can contain any number of different symbols, but we will mainly make do with three:  $\triangleright$ , 0, and 1. When the mechanism is started, the tape is empty (i.e., each square contains the symbol 0) except for the leftmost square, which contains  $\triangleright$ , and a finite number of squares which contain the *input*. At any time, the mechanism is in one of a finite number of *states*. At the outset, the head scans the leftmost square and in a specified *initial state*. At each step of the mechanism’s run, the content of the square currently scanned together with the state the mechanism is in and the Turing machine program determine what happens next. The Turing machine program is given by a partial function which takes as input a state  $q$  and a symbol  $\sigma$  and outputs a triple  $\langle q', \sigma', D \rangle$ . Whenever the mechanism is in

explanation

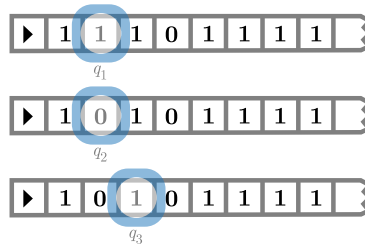


Figure mac.1: A Turing machine executing its program.

tur:mac:int:  
fig:tm

state  $q$  and reads symbol  $\sigma$ , it replaces the symbol on the current square with  $\sigma'$ , the head moves left, right, or stays put according to whether  $D$  is  $L$ ,  $R$ , or  $N$ , and the mechanism goes into state  $q'$ .

For instance, consider the situation in **Figure mac.1**. The visible part of the tape of the Turing machine contains the end-of-tape symbol  $\triangleright$  on the leftmost square, followed by three 1's, a 0, and four more 1's. The head is reading the third square from the left, which contains a 1, and is in state  $q_1$ —we say “the machine is reading a 1 in state  $q_1$ .” If the program of the Turing machine returns, for input  $\langle q_1, 1 \rangle$ , the triple  $\langle q_2, 0, N \rangle$ , the machine would now replace the 1 on the third square with a 0, leave the read/write head where it is, and switch to state  $q_2$ . If then the program returns  $\langle q_3, 0, R \rangle$  for input  $\langle q_2, 0 \rangle$ , the machine would now overwrite the 0 with another 0 (effectively, leaving the content of the tape under the read/write head unchanged), move one square to the right, and enter state  $q_3$ . And so on.

We say that the machine *halts* when it encounters some state,  $q_n$ , and symbol,  $\sigma$  such that there is no instruction for  $\langle q_n, \sigma \rangle$ , i.e., the transition function for input  $\langle q_n, \sigma \rangle$  is undefined. In other words, the machine has no instruction to carry out, and at that point, it ceases operation. Halting is sometimes represented by a specific halt state  $h$ . This will be demonstrated in more detail later on.

digression

The beauty of Turing’s paper, “On computable numbers,” is that he presents not only a formal definition, but also an argument that the definition captures the intuitive notion of computability. From the definition, it should be clear that any function computable by a Turing machine is computable in the intuitive sense. Turing offers three types of argument that the converse is true, i.e., that any function that we would naturally regard as computable is computable by such a machine. They are (in Turing’s words):

1. A direct appeal to intuition.
2. A proof of the equivalence of two definitions (in case the new definition has a greater intuitive appeal).
3. Giving examples of large classes of numbers which are computable.

Our goal is to try to define the notion of computability “in principle,” i.e., without taking into account practical limitations of time and space. Of course, with the broadest definition of computability in place, one can then go on to consider computation with bounded resources; this forms the heart of the subject known as “computational complexity.”

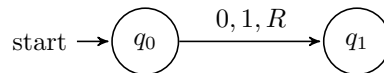
**Historical Remarks** Alan Turing invented Turing machines in 1936. While his interest at the time was the decidability of first-order logic, the paper has been described as a definitive paper on the foundations of computer design. In the paper, Turing focuses on computable real numbers, i.e., real numbers whose decimal expansions are computable; but he notes that it is not hard to adapt his notions to computable functions on the natural numbers, and so on. Notice that this was a full five years before the first working general purpose computer was built in 1941 (by the German Konrad Zuse in his parent’s living room), seven years before Turing and his colleagues at Bletchley Park built the code-breaking Colossus (1943), nine years before the American ENIAC (1945), twelve years before the first British general purpose computer—the Manchester Small-Scale Experimental Machine—was built in Manchester (1948), and thirteen years before the Americans first tested the BINAC (1949). The Manchester SSEM has the distinction of being the first stored-program computer—previous machines had to be rewired by hand for each new task.

## mac.2 Representing Turing Machines

tur:mac:rep:  
sec

Turing machines can be represented visually by *state diagrams*. The diagrams are composed of state cells connected by arrows. Unsurprisingly, each state cell represents a state of the machine. Each arrow represents an instruction that can be carried out from that state, with the specifics of the instruction written above or below the appropriate arrow. Consider the following machine, which has only two internal states,  $q_0$  and  $q_1$ , and one instruction:

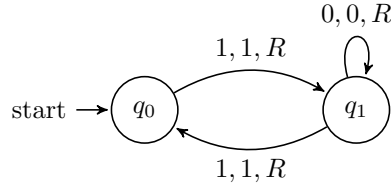
explanation



Recall that the Turing machine has a read/write head and a tape with the input written on it. The instruction can be read as *if reading a 0 in state  $q_0$ , write a 1, move right, and move to state  $q_1$* . This is equivalent to the transition function mapping  $\langle q_0, 0 \rangle$  to  $\langle q_1, 1, R \rangle$ .

**Example mac.1.** *Even Machine:* The following Turing machine halts if, and only if, there are an even number of 1’s on the tape (under the assumption

that all 1's come before the first 0 on the tape).



The state diagram corresponds to the following transition function:

$$\begin{aligned}\delta(q_0, 1) &= \langle q_1, 1, R \rangle, \\ \delta(q_1, 1) &= \langle q_0, 1, R \rangle, \\ \delta(q_1, 0) &= \langle q_1, 0, R \rangle\end{aligned}$$

explanation

The above machine halts only when the input is an even number of strokes. Otherwise, the machine (theoretically) continues to operate indefinitely. For any machine and input, it is possible to trace through the *configurations* of the machine in order to determine the output. We will give a formal definition of configurations later. For now, we can intuitively think of configurations as a series of diagrams showing the state of the machine at any point in time during operation. Configurations show the content of the tape, the state of the machine and the location of the read/write head.

Let us trace through the configurations of the even machine if it is started with an input of four 1's. In this case, we expect that the machine will halt. We will then run the machine on an input of three 1's, where the machine will run forever.

The machine starts in state  $q_0$ , scanning the leftmost 1. We can represent the initial state of the machine as follows:

$$\triangleright_0 1110 \dots$$

The above configuration is straightforward. As can be seen, the machine starts in state one, scanning the leftmost 1. This is represented by a subscript of the state name on the first 1. The applicable instruction at this point is  $\delta(q_0, 1) = \langle q_1, 1, R \rangle$ , and so the machine moves right on the tape and changes to state  $q_1$ .

$$\triangleright_{11} 110 \dots$$

Since the machine is now in state  $q_1$  scanning a 1, we have to “follow” the instruction  $\delta(q_1, 1) = \langle q_0, 1, R \rangle$ . This results in the configuration

$$\triangleright_{1110} 10 \dots$$

As the machine continues, the rules are applied again in the same order, resulting in the following two configurations:

$$\triangleright_{11110} 10 \dots$$

▷11110<sub>0</sub>...

The machine is now in state  $q_0$  scanning a 0. Based on the transition diagram, we can easily see that there is no instruction to be carried out, and thus the machine has halted. This means that the input has been accepted.

Suppose next we start the machine with an input of three 1's. The first few configurations are similar, as the same instructions are carried out, with only a small difference of the tape input:

▷1<sub>0</sub>110...

▷11<sub>1</sub>10...

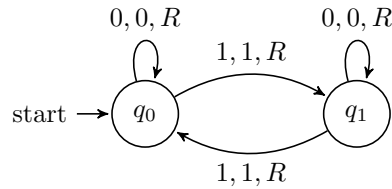
▷111<sub>0</sub>0...

▷1110<sub>1</sub>...

The machine has now traversed past all the 1's, and is reading a 0 in state  $q_1$ . As shown in the diagram, there is an instruction of the form  $\delta(q_1, 0) = \langle q_1, 0, R \rangle$ . Since the tape is filled with 0 indefinitely to the right, the machine will continue to execute this instruction *forever*, staying in state  $q_1$  and moving ever further to the right. The machine will never halt, and does not accept the input.

It is important to note that not all machines will halt. If halting means that the machine runs out of instructions to execute, then we can create a machine [explanation](#) that never halts simply by ensuring that there is an outgoing arrow for each symbol at each state. The even machine can be modified to run indefinitely by adding an instruction for scanning a 0 at  $q_0$ .

### Example mac.2.



Machine tables are another way of representing Turing machines. Machine [explanation](#) tables have the tape alphabet displayed on the  $x$ -axis, and the set of machine states across the  $y$ -axis. Inside the table, at the intersection of each state and symbol, is written the rest of the instruction—the new state, new symbol, and direction of movement. Machine tables make it easy to determine in what state, and for what symbol, the machine halts. Whenever there is a gap in the table is a possible point for the machine to halt. Unlike state diagrams and instruction sets, where the points at which the machine halts are not always immediately obvious, any halting points are quickly identified by finding the gaps in the machine table.

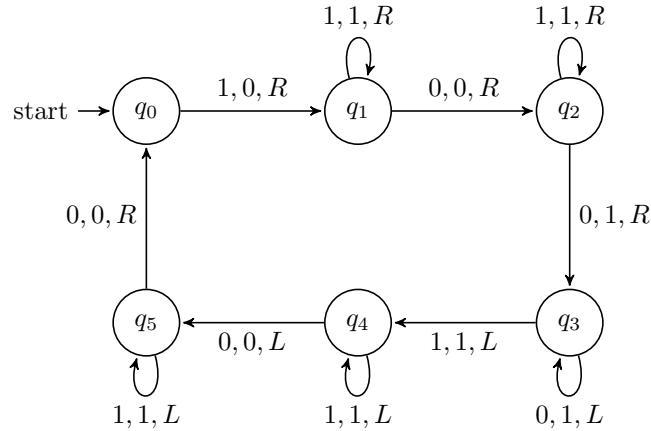


Figure mac.2: A doubler machine

**Example mac.3.** The machine table for the even machine is:

	0	1	$\triangleright$
$q_0$		$1, q_1, R$	
$q_1$	$0, q_1, R$	$1, q_0, R$	

As we can see, the machine halts when scanning a 0 in state  $q_0$ .

explanation

So far we have only considered machines that read and accept input. However, Turing machines have the capacity to both read and write. An example of such a machine (although there are many, many examples) is a *doubler*. A doubler, when started with a block of  $n$  1's on the tape, outputs a block of  $2n$  1's.

**Example mac.4.** Before building a doubler machine, it is important to come up with a *strategy* for solving the problem. Since the machine (as we have formulated it) cannot remember how many 1's it has read, we need to come up with a way to keep track of all the 1's on the tape. One such way is to separate the output from the input with a 0. The machine can then erase the first 1 from the input, traverse over the rest of the input, leave a 0, and write two new 1's. The machine will then go back and find the second 1 in the input, and double that one as well. For each one 1 of input, it will write two 1's of output. By erasing the input as the machine goes, we can guarantee that no 1 is missed or doubled twice. When the entire input is erased, there will be  $2n$  1's left on the tape. The state diagram of the resulting Turing machine is depicted in [Figure mac.2](#).

tur:mac:rep:  
fig:doubler

tur:mac:rep:  
ex:doubler

**Problem mac.1.** Choose an arbitrary input and trace through the configurations of the doubler machine in [Example mac.4](#).

**Problem mac.2.** Design a Turing-machine with alphabet  $\{\triangleright, 0, A, B\}$  that accepts, i.e., halts on, any string of  $A$ 's and  $B$ 's where the number of  $A$ 's is the same as the number of  $B$ 's *and* all the  $A$ 's precede all the  $B$ 's, and rejects, i.e., does not halt on, any string where the number of  $A$ 's is not equal to the number of  $B$ 's or the  $A$ 's do not precede all the  $B$ 's. (E.g., the machine should accept  $AABB$ , and  $AAABBB$ , but reject both  $AAB$  and  $AABBAABB$ .)

**Problem mac.3.** Design a Turing-machine with alphabet  $\{\triangleright, 0, A, B\}$  that takes as input any string  $\alpha$  of  $A$ 's and  $B$ 's and duplicates them to produce an output of the form  $\alpha\alpha$ . (E.g. input  $ABBA$  should result in output  $ABBAABBA$ ).

**Problem mac.4.** *Alphabetical?*: Design a Turing-machine with alphabet  $\{\triangleright, 0, A, B\}$  that when given as input a finite sequence of  $A$ 's and  $B$ 's checks to see if all the  $A$ 's appear to the left of all the  $B$ 's or not. The machine should leave the input string on the tape, and either halt if the string is “alphabetical”, or loop forever if the string is not.

**Problem mac.5.** *Alphabetizer*: Design a Turing-machine with alphabet  $\{\triangleright, 0, A, B\}$  that takes as input a finite sequence of  $A$ 's and  $B$ 's rearranges them so that all the  $A$ 's are to the left of all the  $B$ 's. (e.g., the sequence  $BABAA$  should become the sequence  $AAABB$ , and the sequence  $ABBABB$  should become the sequence  $AABBBB$ ).

### mac.3 Turing Machines

tur:mac:tur: sec The formal definition of what constitutes a Turing machine looks abstract, but explanation is actually simple: it merely packs into one mathematical structure all the information needed to specify the workings of a Turing machine. This includes (1) which states the machine can be in, (2) which symbols are allowed to be on the tape, (3) which state the machine should start in, and (4) what the instruction set of the machine is.

**Definition mac.5 (Turing machine).** A *Turing machine*  $M$  is a tuple  $\langle Q, \Sigma, q_0, \delta \rangle$  consisting of

1. a finite set of *states*  $Q$ ,
2. a finite *alphabet*  $\Sigma$  which includes  $\triangleright$  and  $0$ ,
3. an *initial state*  $q_0 \in Q$ ,
4. a finite *instruction set*  $\delta: Q \times \Sigma \rightarrow Q \times \Sigma \times \{L, R, N\}$ .

The partial function  $\delta$  is also called the *transition function* of  $M$ .

We assume that the tape is infinite in one direction only. For this reason explanation it is useful to designate a special symbol  $\triangleright$  as a marker for the left end of the tape. This makes it easier for Turing machine programs to tell when they're

“in danger” of running off the tape. We could assume that this symbol is never overwritten, i.e., that  $\delta(q, \triangleright) = \langle q', \triangleright, x \rangle$  if  $\delta(q, \triangleright)$  is defined. Some textbooks do this, we do not. You can simply be careful when constructing your Turing machine that it never overwrites  $\triangleright$ . Moreover, there are cases where allowing such overwriting provides some convenient flexibility.

**Example mac.6.** *Even Machine:* The even machine is formally the quadruple  $\langle Q, \Sigma, q_0, \delta \rangle$  where

$$\begin{aligned} Q &= \{q_0, q_1\} \\ \Sigma &= \{\triangleright, 0, 1\}, \\ \delta(q_0, 1) &= \langle q_1, 1, R \rangle, \\ \delta(q_1, 1) &= \langle q_0, 1, R \rangle, \\ \delta(q_1, 0) &= \langle q_1, 0, R \rangle. \end{aligned}$$

## tur.4 Configurations and Computations

explanation

Recall tracing through the configurations of the even machine earlier. The imaginary mechanism consisting of tape, read/write head, and Turing machine program is really just an intuitive way of visualizing what a Turing machine computation is. Formally, we can define the computation of a Turing machine on a given input as a sequence of *configurations*—and a configuration in turn is a sequence of symbols (corresponding to the contents of the tape at a given point in the computation), a number indicating the position of the read/write head, and a state. Using these, we can define what the Turing machine  $M$  computes on a given input.

cmp:tur:con:  
sec

**Definition tur.7 (Configuration).** A *configuration* of Turing machine  $M = \langle Q, \Sigma, q_0, \delta \rangle$  is a triple  $\langle C, m, q \rangle$  where

1.  $C \in \Sigma^*$  is a finite sequence of symbols from  $\Sigma$ ,
2.  $m \in \mathbb{N}$  is a number  $< \text{len}(C)$ , and
3.  $q \in Q$

Intuitively, the sequence  $C$  is the content of the tape (symbols of all squares from the leftmost square to the last non-blank or previously visited square),  $m$  is the number of the square the read/write head is scanning (beginning with 0 being the number of the leftmost square), and  $q$  is the current state of the machine.

explanation

The potential input for a Turing machine is a sequence of symbols, usually a sequence that encodes a number in some form. The initial configuration of the Turing machine is that configuration in which we start the Turing machine to work on that input: the tape contains the tape end marker immediately followed by the input written on the squares to the right, the read/write head



is scanning the leftmost square of the input (i.e., the square to the right of the left end marker), and the mechanism is in the designated start state  $q_0$ .

**Definition tur.8 (Initial configuration).** The *initial configuration* of  $M$  for input  $I \in \Sigma^*$  is

$$\langle \triangleright \frown I, 1, q_0 \rangle.$$

The  $\frown$  symbol is for *concatenation*—the input string begins immediately to the left end marker. explanation

**Definition tur.9.** We say that a configuration  $\langle C, m, q \rangle$  *yields the configuration*  $\langle C', m', q' \rangle$  *in one step* (according to  $M$ ), iff

1. the  $m$ -th symbol of  $C$  is  $\sigma$ ,
2. the instruction set of  $M$  specifies  $\delta(q, \sigma) = \langle q', \sigma', D \rangle$ ,
3. the  $m$ -th symbol of  $C'$  is  $\sigma'$ , and
4. a)  $D = L$  and  $m' = m - 1$  if  $m > 0$ , otherwise  $m' = 0$ , or  
b)  $D = R$  and  $m' = m + 1$ , or  
c)  $D = N$  and  $m' = m$ ,
5. if  $m' = \text{len}(C)$ , then  $\text{len}(C') = \text{len}(C) + 1$  and the  $m'$ -th symbol of  $C'$  is 0. Otherwise  $\text{len}(C') = \text{len}(C)$ .
6. for all  $i$  such that  $i < \text{len}(C)$  and  $i \neq m$ ,  $C'(i) = C(i)$ ,

cmp:tur:con:  
defn:run-output

**Definition tur.10.** A *run of  $M$  on input  $I$*  is a sequence  $C_i$  of configurations of  $M$ , where  $C_0$  is the initial configuration of  $M$  for input  $I$ , and each  $C_i$  yields  $C_{i+1}$  in one step.

We say that  $M$  *halts on input  $I$  after  $k$  steps* if  $C_k = \langle C, m, q \rangle$ , the  $m$ th symbol of  $C$  is  $\sigma$ , and  $\delta(q, \sigma)$  is undefined. In that case, the *output* of  $M$  for input  $I$  is  $O$ , where  $O$  is a string of symbols not ending in 0 such that  $C = \triangleright \frown O \frown 0^j$  for some  $i, j \in \mathbb{N}$ .

According to this definition, the output  $O$  of  $M$  always ends in a symbol other than 0, or, if at time  $k$  the entire tape is filled with 0 (except for the leftmost  $\triangleright$ ),  $O$  is the empty string. explanation

## mac.5 Unary Representation of Numbers

tur:mac:una:  
sec

Turing machines work on sequences of symbols written on their tape. Depending on the alphabet a Turing machine uses, these sequences of symbols can represent various inputs and outputs. Of particular interest, of course, are Turing machines which compute *arithmetical* functions, i.e., functions of natural numbers. A simple way to represent positive integers is by coding them as sequences of a single symbol 1. If  $n \in \mathbb{N}$ , let  $1^n$  be the empty sequence if  $n = 0$ , and otherwise the sequence consisting of exactly  $n$  1's. explanation

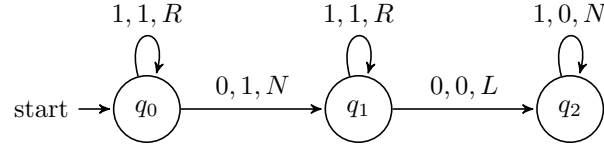


Figure mac.3: A machine computing  $f(x, y) = x + y$

**Definition mac.11 (Computation).** A Turing machine  $M$  computes the function  $f: \mathbb{N}^k \rightarrow \mathbb{N}$  iff  $M$  halts on input

$$1^{n_1}01^{n_2}0 \dots 01^{n_k}$$

with output  $1^{f(n_1, \dots, n_k)}$ .

**Problem mac.6.** Give a definition for when a Turing machine  $M$  computes the function  $f: \mathbb{N}^k \rightarrow \mathbb{N}^m$ .

**Example mac.12.** *Addition:* Let's build a machine that computes the function  $f(n, m) = n + m$ . This requires a machine that starts with two blocks of 1's of length  $n$  and  $m$  on the tape, and halts with one block consisting of  $n + m$  1's. The two input blocks of 1's are separated by a 0, so one method would be to write a stroke on the square containing the 0, and erase the last 1.

**Problem mac.7.** Trace through the configurations of the machine from **Example mac.12** for input  $\langle 3, 2 \rangle$ . What happens if the machine computes  $0 + 0$ ?

explanation

In **Example mac.4**, we gave an example of a Turing machine that takes as input a sequence of 1's and halts with a sequence of twice as many 1's on the tape—the doubler machine. However, because the output contains 0's to the left of the doubled block of 1's, it does not actually compute the function  $f(x) = 2x$ , as you might have assumed. We'll describe two ways of fixing that.

**Example mac.13.** The machine in **Figure mac.4** computes the function  $f(x) = 2x$ . Instead of erasing the input and writing two 1's at the far right for every 1 in the input as the machine from **Example mac.4** does, this machine adds a single 1 to the right for every 1 in the input. It has to keep track of where the input ends, so it leaves a 0 between the input and the added strokes, which it fills with a 1 at the very end. And we have to “remember” where we are in the input, so we temporarily replace a 1 in the input block by a 0.

**Example mac.14.** A second possibility for computing  $f(x) = 2x$  is to keep the original doubler machine, but add states and instructions at the end which move the doubled block of strokes to the far left of the tape. The machine in

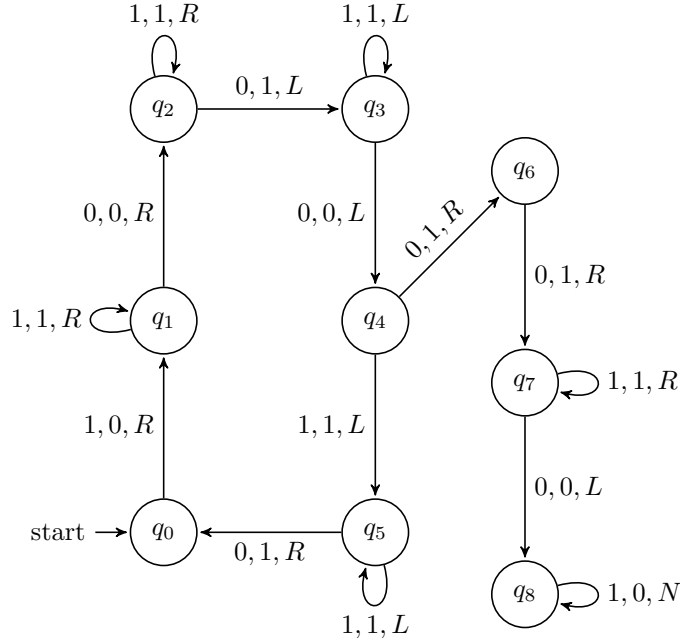


Figure mac.4: A machine computing  $f(x) = 2x$

tur:mac:una:  
fig:doubler-disc

**Figure mac.5** does just this last part: started on a tape consisting of a block of 0's followed by a block of 1's (and the head positioned anywhere in the block of 0's), it erases the 1's one at a time and writes them at the beginning of the tape. In order to be able to tell when it is done, it first marks the end of the block of 1's with a  $\triangleright$  symbol, which gets deleted at the end. We've started numbering the states at  $q_6$ , so they can be added to the doubler machine. All you'll need is an additional instruction  $\delta(q_5, 0) = \langle q_6, 0, N \rangle$ , i.e., an arrow from  $q_5$  to  $q_6$  labelled  $0, 0, N$ . (There is one subtle problem: the resulting machine does not work for input  $x = 0$ . We'll leave this as an exercise.)

**Problem mac.8.** In **Example mac.14** we described a machine consisting of a combination of the doubler machine from **Figure mac.4** and the mover machine from **Figure mac.5**. What happens if you start this combined machine on input  $x = 0$ , i.e., on an empty tape? How would you fix the machine so that in this case the machine halts with output  $2x = 0$ ? (You should be able to do this by adding one state and one transition.)

**Problem mac.9. Subtraction:** Design a Turing machine that when given an input of two non-empty strings of strokes of length  $n$  and  $m$ , where  $n > m$ , computes the function  $f(n, m) = n - m$ .

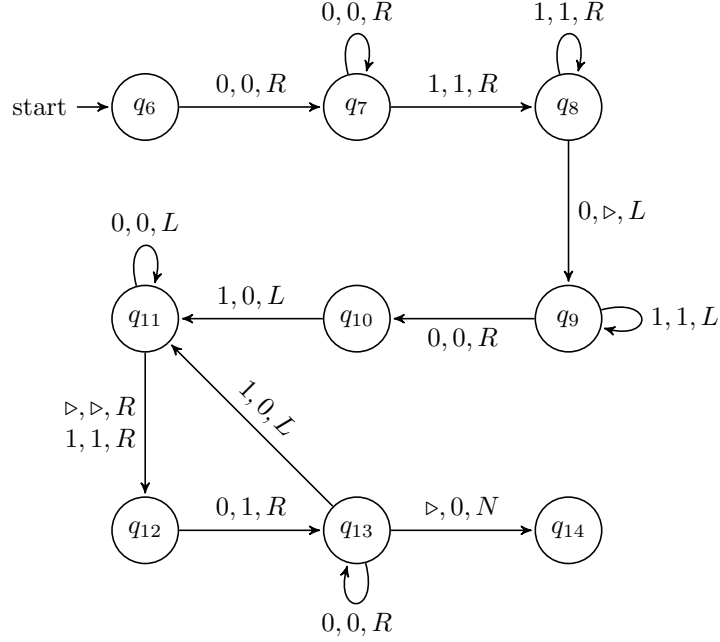


Figure mac.5: Moving a block of 1's to the left

**Problem mac.10. Equality:** Design a Turing machine to compute the following function:

$$\text{equality}(n, m) = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases}$$

where  $n$  and  $m \in \mathbb{Z}^+$ .

**Problem mac.11.** Design a Turing machine to compute the function  $\min(x, y)$  where  $x$  and  $y$  are positive integers represented on the tape by strings of 1's separated by a 0. You may use additional symbols in the alphabet of the machine.

The function  $\min$  selects the smallest value from its arguments, so  $\min(3, 5) = 3$ ,  $\min(20, 16) = 16$ , and  $\min(4, 4) = 4$ , and so on.

**Definition mac.15.** A Turing machine  $M$  computes the partial function  $f: \mathbb{N}^k \rightarrow \mathbb{N}$  iff,

1.  $M$  halts on input  $1^{n_1} \frown 0 \frown \dots \frown 0 \frown 1^{n_k}$  with output  $1^m$  if  $f(n_1, \dots, n_k) = m$ .
2.  $M$  does not halt at all, or with an output that is not a single block of 1's if  $f(n_1, \dots, n_k)$  is undefined.

## mac.6 Halting States

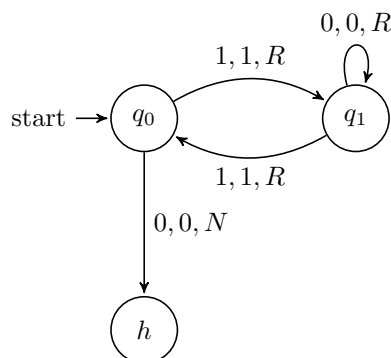
tur:mac:hal:  
sec

Although we have defined our machines to halt only when there is no instruction to carry out, common representations of Turing machines have a dedicated *halting state*  $h$ , such that  $h \in Q$ .

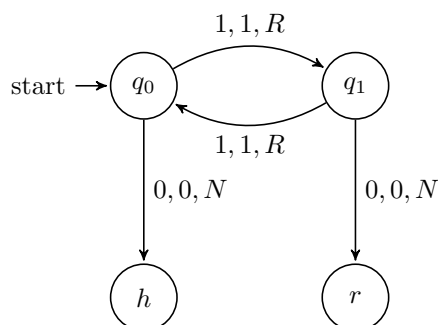
[explanation](#)

The idea behind a halting state is simple: when the machine has finished operation (it is ready to accept input, or has finished writing the output), it goes into a state  $h$  where it halts. Some machines have two halting states, one that accepts input and one that rejects input.

**Example mac.16.** *Halting States.* To elucidate this concept, let us begin with an alteration of the even machine. Instead of having the machine halt in state  $q_0$  if the input is even, we can add an instruction to send the machine into a halting state.



Let us further expand the example. When the machine determines that the input is odd, it never halts. We can alter the machine to include a *reject* state by replacing the looping instruction with an instruction to go to a reject state  $r$ .



Adding a dedicated halting state can be advantageous in cases like this, where it makes explicit when the machine accepts/rejects certain inputs. However, it is important to note that no computing power is gained by adding a dedicated halting state. Similarly, a less formal notion of halting has its

[explanation](#)

own advantages. The definition of halting used so far in this chapter makes the proof of the *Halting Problem* intuitive and easy to demonstrate. For this reason, we continue with our original definition.

## mac.7 Disciplined Machines

explanation In section [section mac.6](#), we considered Turing machines that have a single, designated halting state  $h$ —such machines are guaranteed to halt, if they halt at all, in state  $h$ . In this way, machines with a single halting state are more “disciplined” than we allow Turing machines in general to be. There are other restrictions we might impose on the behavior of Turing machines. For instance, we also have not prohibited Turing machines from ever erasing the tape-end marker on square 0, or to attempt to move left from square 0. (Our definition states that the head simply stays on square 0 in this case; other definitions have the machine halt.) It is likewise sometimes desirable to be able to assume that a Turing machine, if it halts at all, halts on square 1. tur:mac:dis:sec

**Definition mac.17.** A Turing machine  $M$  is *disciplined* iff

1. it has a designated single halting state  $h$ ,
2. it halts, if it halts at all, while scanning square 1,
3. it never erases the  $\triangleright$  symbol on square 0, and
4. it never attempts to move left from square 0.

tur:mac:dis:prop:disciplined

explanation We have already discussed that any Turing machine can be changed into one with the same behavior but with a designated halting state. This is done simply by adding a new state  $h$ , and adding an instruction  $\delta(q, \sigma) = \langle h, \sigma, N \rangle$  for any pair  $\langle q, \sigma \rangle$  where the original  $\delta$  is undefined. It is true, although tedious to prove, that any Turing machine  $M$  can be turned into a disciplined Turing machine  $M'$  which halts on the same inputs and produces the same output. For instance, if the Turing machine halts and is not on square 1, we can add some instructions to make the head move left until it finds the tape-end marker, then move one square to the right, then halt. We’ll leave you to think about how the other conditions can be dealt with.

**Example mac.18.** In [Figure mac.6](#), we turn the addition machine from [Example mac.12](#) into a disciplined machine.

**Proposition mac.19.** *For every Turing machine  $M$ , there is a disciplined Turing machine  $M'$  which halts with output  $O$  if  $M$  halts with output  $O$ , and does not halt if  $M$  does not halt. In particular, any function  $f: \mathbb{N}^n \rightarrow \mathbb{N}$  computable by a Turing machine is also computable by a disciplined Turing machine.* tur:mac:dis:prop:disciplined

**Problem mac.12.** Give a disciplined machine that computes  $f(x) = x + 1$ .

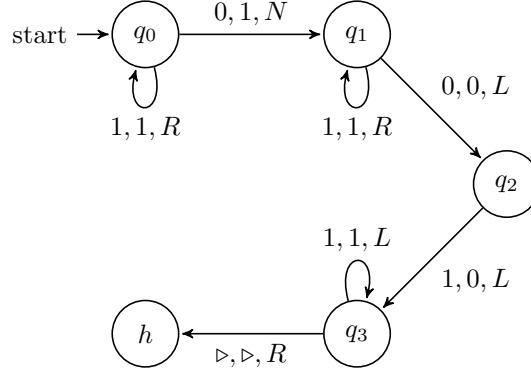


Figure mac.6: A disciplined addition machine

tur:mac:dis:  
fig:adder-disc

**Problem mac.13.** Find a disciplined machine which, when started on input  $1^n$  produces output  $1^n \frown 0 \frown 1^n$ .

## mac.8 Combining Turing Machines

tur:mac:cmb:  
sec

The examples of Turing machines we have seen so far have been fairly simple in nature. But in fact, any problem that can be solved with any modern programming language can also be solved with Turing machines. To build more complex Turing machines, it is important to convince ourselves that we can combine them, so we can build machines to solve more complex problems by breaking the procedure into simpler parts. If we can find a natural way to break a complex problem down into constituent parts, we can tackle the problem in several stages, creating several simple Turing machines and combining them into one machine that can solve the problem. This point is especially important when tackling the Halting Problem in the next section.

explanation

How do we combine Turing machines  $M = \langle Q, \Sigma, q_0, \delta \rangle$  and  $M' = \langle Q', \Sigma', q'_0, \delta' \rangle$ ? We now use the configuration of the tape after  $M$  has halted as the input configuration of a run of machine  $M'$ . To get a single Turing machine  $M \frown M'$  that does this, do the following:

1. Renumber (or relabel) all the states  $Q'$  of  $M'$  so that  $M$  and  $M'$  have no states in common ( $Q \cap Q' = \emptyset$ ).
2. The states of  $M \frown M'$  are  $Q \cup Q'$ .
3. The tape alphabet is  $\Sigma \cup \Sigma'$ .
4. The start state is  $q_0$ .

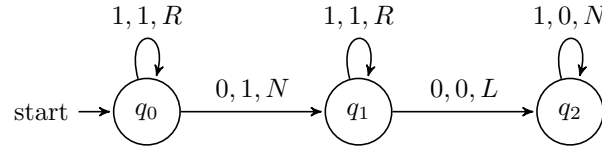
5. The transition function is the function  $\delta''$  given by:

$$\delta''(q, \sigma) = \begin{cases} \delta(q, \sigma) & \text{if } q \in Q \\ \delta'(q, \sigma) & \text{if } q \in Q' \\ \langle q'_0, \sigma, N \rangle & \text{if } q \in Q \text{ and } \delta(q, \sigma) \text{ is undefined} \end{cases}$$

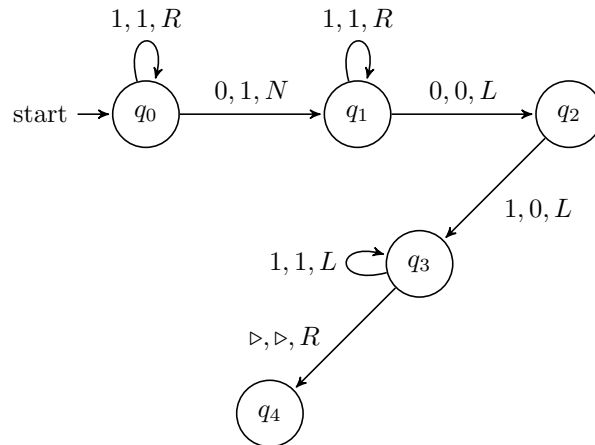
The resulting machine uses the instructions of  $M$  when it is in a state  $q \in Q$ , the instructions of  $M'$  when it is in a state  $q \in Q'$ . When it is in a state  $q \in Q$  and is scanning a symbol  $\sigma$  for which  $M$  has no transition (i.e.,  $M$  would have halted), it enters the start state of  $M'$  (and leaves the tape contents and head position as it is).

Note that unless the machine  $M$  is disciplined, we don't know where the tape head is when  $M$  halts, so the halting configuration of  $M$  need not have the head scanning square 1. When combining machines, it's important to keep this in mind.

**Example mac.20. Combining Machines:** We'll design a machine which, when started on input consisting of two blocks of 1's of length  $n$  and  $m$ , halts with a single block of  $2(m+n)$  1's on the tape. In order to build this machine, we can combine two machines we are already familiar with: the addition machine, and the doubler. We begin by drawing a state diagram for the addition machine.



Instead of halting in state  $q_2$ , we want to continue operation in order to double the output. Recall that the doubler machine erases the first stroke in the input and writes two strokes in a separate output. Let's add an instruction to make sure the tape head is reading the first stroke of the output of the addition machine.





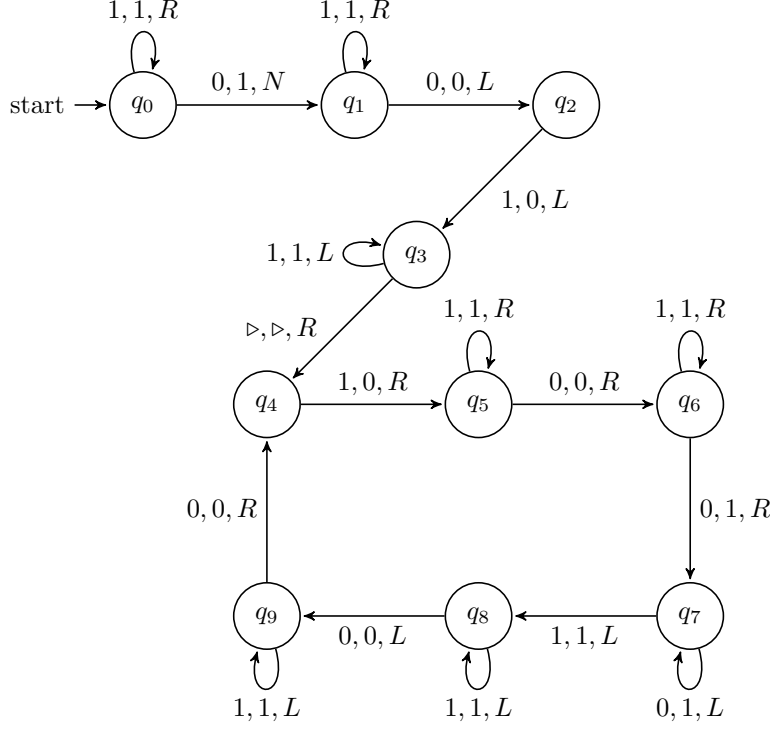


Figure mac.7: Combining adder and doubler machines

tur:mac:cmb:  
fig:combined

It is now easy to double the input—all we have to do is connect the doubler machine onto state  $q_4$ . This requires renaming the states of the doubler machine so that they start at  $q_4$  instead of  $q_0$ —this way we don't end up with two starting states. The final diagram should look as in [Figure mac.7](#).

**Proposition mac.21.** *If  $M$  and  $M'$  are disciplined and compute the functions  $f: \mathbb{N}^k \rightarrow \mathbb{N}$  and  $f': \mathbb{N} \rightarrow \mathbb{N}$ , respectively, then  $M \frown M'$  is disciplined and computes  $f' \circ f$ .*

*Proof.* Since  $M$  is disciplined, when it halts with output  $f(n_1, \dots, n_k) = m$ , the head is scanning square 1. If we now enter the start state of  $M'$ , the machine will halt with output  $f'(m)$ , again scanning square 1. The other conditions of [Definition mac.17](#) are also satisfied.  $\square$

**Problem mac.14.** Give a disciplined Turing machine computing  $f(x) = x + 2$  by taking the machine  $M$  from [Problem mac.12](#) and construct  $M \frown M$ .

## mac.9 Variants of Turing Machines

There are in fact many possible ways to define Turing machines, of which ours is only one. In some ways, our definition is more liberal than others. We allow arbitrary finite alphabets, a more restricted definition might allow only two tape symbols, 1 and 0. We allow the machine to write a symbol to the tape and move at the same time, other definitions allow either writing or moving. We allow the possibility of writing without moving the tape head, other definitions leave out the  $N$  “instruction.” In other ways, our definition is more restrictive. We assumed that the tape is infinite in one direction only, other definitions allow the tape to be infinite both to the left and the right. In fact, one can even allow any number of separate tapes, or even an infinite grid of squares. We represent the instruction set of the Turing machine by a transition function; other definitions use a transition relation where the machine has more than one possible instruction in any given situation.

tur:mac:var:  
sec

This last relaxation of the definition is particularly interesting. In our definition, when the machine is in state  $q$  reading symbol  $\sigma$ ,  $\delta(q, \sigma)$  determines what the new symbol, state, and tape head position is. But if we allow the instruction set to be a relation between current state-symbol pairs  $\langle q, \sigma \rangle$  and new state-symbol-direction triples  $\langle q', \sigma', D \rangle$ , the action of the Turing machine may not be uniquely determined—the instruction relation may contain both  $\langle q, \sigma, q', \sigma', D \rangle$  and  $\langle q, \sigma, q'', \sigma'', D' \rangle$ . In this case we have a *non-deterministic* Turing machine. These play an important role in computational complexity theory.

There are also different conventions for when a Turing machine halts: we say it halts when the transition function is undefined, other definitions require the machine to be in a special designated halting state. We have explained in [section mac.6](#) why requiring a designated halting state is not a restriction which impacts what Turing machines can compute. Since the tapes of our Turing machines are infinite in one direction only, there are cases where a Turing machine can’t properly carry out an instruction: if it reads the leftmost square and is supposed to move left. According to our definition, it just stays put instead of “falling off”, but we could have defined it so that it halts when that happens. This definition is also equivalent: we could simulate the behavior of a Turing machine that halts when it attempts to move left from square 0 by deleting every transition  $\delta(q, \triangleright) = \langle q', \sigma, L \rangle$ —then instead of attempting to move left on  $\triangleright$  the machine halts.<sup>1</sup>

There are also different ways of representing numbers (and hence the input-output function computed by a Turing machine): we use unary representation, but you can also use binary representation. This requires two symbols in addition to 0 and  $\triangleright$ .

Now here is an interesting fact: none of these variations matters as to which functions are Turing computable. *If a function is Turing computable according*

---

<sup>1</sup>This doesn’t *quite* work, since nothing prevents us from writing and reading  $\triangleright$  on squares other than square 0 (see [Example mac.14](#)). We can get around that by adding a second  $\triangleright'$  symbol to use instead for such a purpose.

to one definition, it is Turing computable according to all of them.

We won't go into the details of verifying this. Here's just one example: we gain no additional computing power by allowing a tape that is infinite in both directions, or multiple tapes. The reason is, roughly, that a Turing machine with a single one-way infinite tape can simulate multiple or two-way infinite tapes. E.g., using additional states and instructions, we can “translate” a program for a machine with multiple tapes or two-way infinite tape into one with a single one-way infinite tape. The translated machine can use the even squares for the squares of tape 1 (or the “positive” squares of a two-way infinite tape) and the odd squares for the squares of tape 2 (or the “negative” squares).

## mac.10 The Church–Turing Thesis

tur:mac:ctt:  
sec

Turing machines are supposed to be a precise replacement for the concept of an effective procedure. Turing thought that anyone who grasped both the concept of an effective procedure and the concept of a Turing machine would have the intuition that anything that could be done via an effective procedure could be done by Turing machine. This claim is given support by the fact that all the other proposed precise replacements for the concept of an effective procedure turn out to be extensionally equivalent to the concept of a Turing machine—that is, they can compute exactly the same set of functions. This claim is called the *Church–Turing thesis*.

**Definition mac.22 (Church–Turing thesis).** The *Church–Turing Thesis* states that anything computable via an effective procedure is Turing computable.

The Church–Turing thesis is appealed to in two ways. The first kind of use of the Church–Turing thesis is an excuse for laziness. Suppose we have a description of an effective procedure to compute something, say, in “pseudo-code.” Then we can invoke the Church–Turing thesis to justify the claim that the same function is computed by some Turing machine, even if we have not in fact constructed it.

The other use of the Church–Turing thesis is more philosophically interesting. It can be shown that there are functions which cannot be computed by Turing machines. From this, using the Church–Turing thesis, one can conclude that it cannot be effectively computed, using any procedure whatsoever. For if there were such a procedure, by the Church–Turing thesis, it would follow that there would be a Turing machine for it. So if we can prove that there is no Turing machine that computes it, there also can't be an effective procedure. In particular, the Church–Turing thesis is invoked to claim that the so-called halting problem not only cannot be solved by Turing machines, it cannot be

effectively solved at all.

## Photo Credits

## Bibliography