Loops and infinity

Boaz Barak

**Disclaimer:** The MS-Word version of this text has significant formatting issues compared to the PDF and HTML versions.

# Loops and infinity

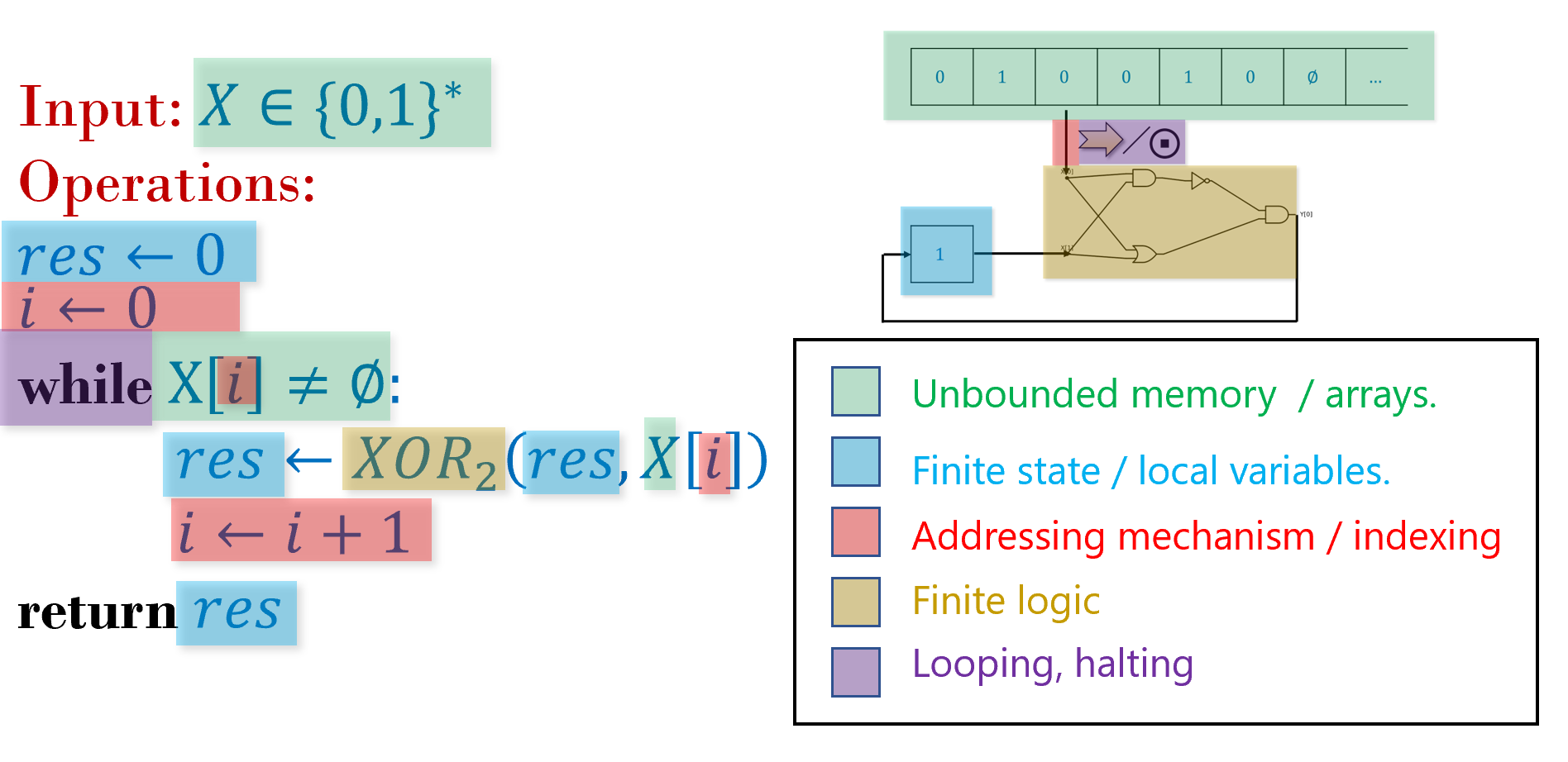
### 

* Learn the model of *Turing machines*, which can compute functions of *arbitrary input lengths*.
* See a programming-language description of Turing machines, using NAND-TM programs, which add *loops* and *arrays* to NAND-CIRC.
* See some basic syntactic sugar and equivalence of variants of Turing machines and NAND-TM programs.

*“The bounds of arithmetic were however outstepped the moment the idea of applying the [punched] cards had occurred; and the Analytical Engine does not occupy common ground with mere ‘calculating machines.’… In enabling mechanism to combine together general symbols, in successions of unlimited variety and extent, a uniting link is established between the operations of matter and the abstract mental processes of the most abstract branch of mathematical science.”* , Ada Augusta, countess of Lovelace, 1843

As the quote of chapinfinite says, an algorithm is “a finite answer to an infinite number of questions”. To express an algorithm, we need to write down a finite set of instructions that will enable us to compute on arbitrarily long inputs. To describe and execute an algorithm we need the following components (see algcomponentfig):

* The finite set of instructions to be performed.
* Some “local variables” or finite state used in the execution.
* A potentially unbounded working memory to store the input and any other values we may require later.
* While the memory is unbounded, at every single step we can only read and write to a finite part of it, and we need a way to *address* which are the parts we want to read from and write to.
* If we only have a finite set of instructions but our input can be arbitrarily long, we will need to *repeat* instructions (i.e., *loop* back). We need a mechanism to decide when we will loop and when we will halt.

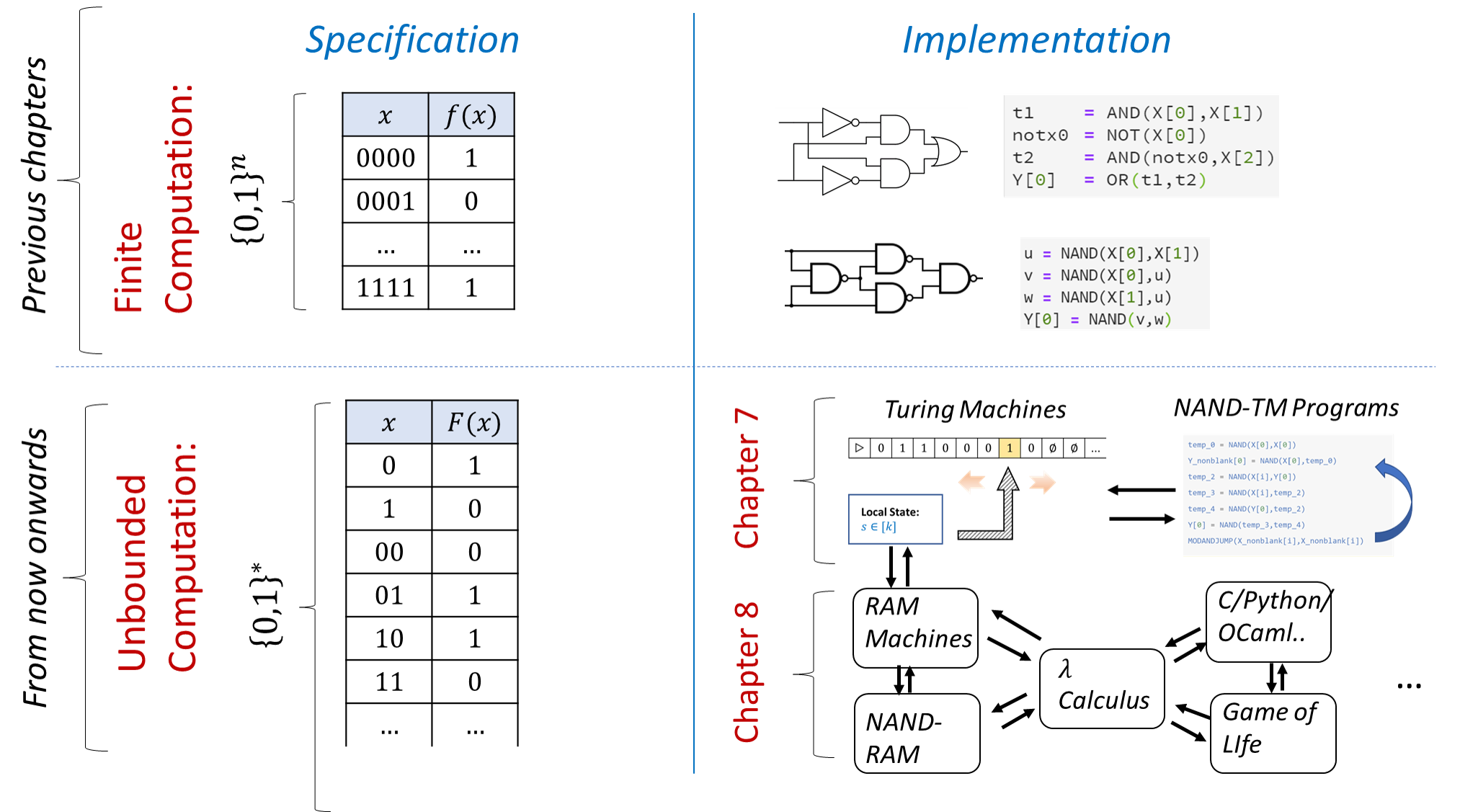


An algorithm is a finite recipe to compute on arbitrarily long inputs. The components of an algorithm include the instructions to be performed, finite state or “local variables”, the memory to store the input and intermediate computations, as well as mechanisms to decide which part of the memory to access, and when to repeat instructions and when to halt.

In this chapter, we give a general model of an algorithm, which (unlike Boolean circuits) is not restricted to a fixed input length, and (unlike finite automata) is not restricted to a finite amount of working memory. We will see two ways to model algorithms:

* *Turing machines*, invented by Alan Turing in 1936, are hypothetical abstract devices that yield finite descriptions of algorithms that can handle arbitrarily long inputs.
* The *NAND-TM Programming language* extends NAND-CIRC with the notion of *loops* and *arrays* to obtain finite programs that can compute a function with arbitrarily long inputs.

It turns out that these two models are *equivalent*. In fact, they are equivalent to many other computational models, including programming languages such as C, Lisp, Python, JavaScript, etc. This notion, known as *Turing equivalence* or *Turing completeness*, will be discussed in chapequivalentmodels. See chaploopoverviewfig for an overview of the models presented in this chapter and chapequivalentmodels.



Overview of our models for finite and unbounded computation. In the previous chapters we study the computation of *finite functions*, which are functions for some fixed , and modeled computing these functions using circuits or straight-line programs. In this chapter we study computing *unbounded* functions of the form or . We model computing these functions using *Turing Machines* or (equivalently) NAND-TM programs, which add the notion of *loops* to the NAND-CIRC programming language. In chapequivalentmodels we will show that these models are equivalent to many other models, including RAM machines, the calculus, and all the common programming languages including C, Python, Java, JavaScript, etc.

## Turing Machines

*“Computing is normally done by writing certain symbols on paper. We may suppose that this paper is divided into squares like a child’s arithmetic book.. The behavior of the [human] computer at any moment is determined by the symbols which he is observing, and of his’ state of mind’ at that moment… We may suppose that in a simple operation not more than one symbol is altered.”* ,  
*“We compare a man in the process of computing … to a machine which is only capable of a finite number of configurations… The machine is supplied with a ‘tape’ (the analogue of paper) … divided into sections (called ‘squares’) each capable of bearing a ‘symbol’”* , Alan Turing, 1936

*“What is the difference between a Turing machine and the modern computer? It’s the same as that between Hillary’s ascent of Everest and the establishment of a Hilton hotel on its peak.”* , Alan Perlis, 1982.



Aside from his many other achievements, Alan Turing was an excellent long-distance runner who just fell shy of making England’s Olympic team. A fellow runner once asked him why he punished himself so much in training. Alan said “I have such a stressful job that the only way I can get it out of my mind is by running hard; it’s the only way I can get some release.”

The “granddaddy” of all models of computation is the *Turing machine*. Turing machines were defined in 1936 by Alan Turing in an attempt to formally capture all the functions that can be computed by human “computers” (see humancomputersfig) that follow a well-defined set of rules, such as the standard algorithms for addition or multiplication.



Until the advent of electronic computers, the word “computer” was used to describe a person that performed calculations. Most of these “human computers” were women, and they were absolutely essential to many achievements, including mapping the stars, breaking the Enigma cipher, and the NASA space mission; see also the bibliographical notes. Photo from [National Photo Company Collection](https://www.loc.gov/pictures/item/2016838906/); see also [@sobel2017the].

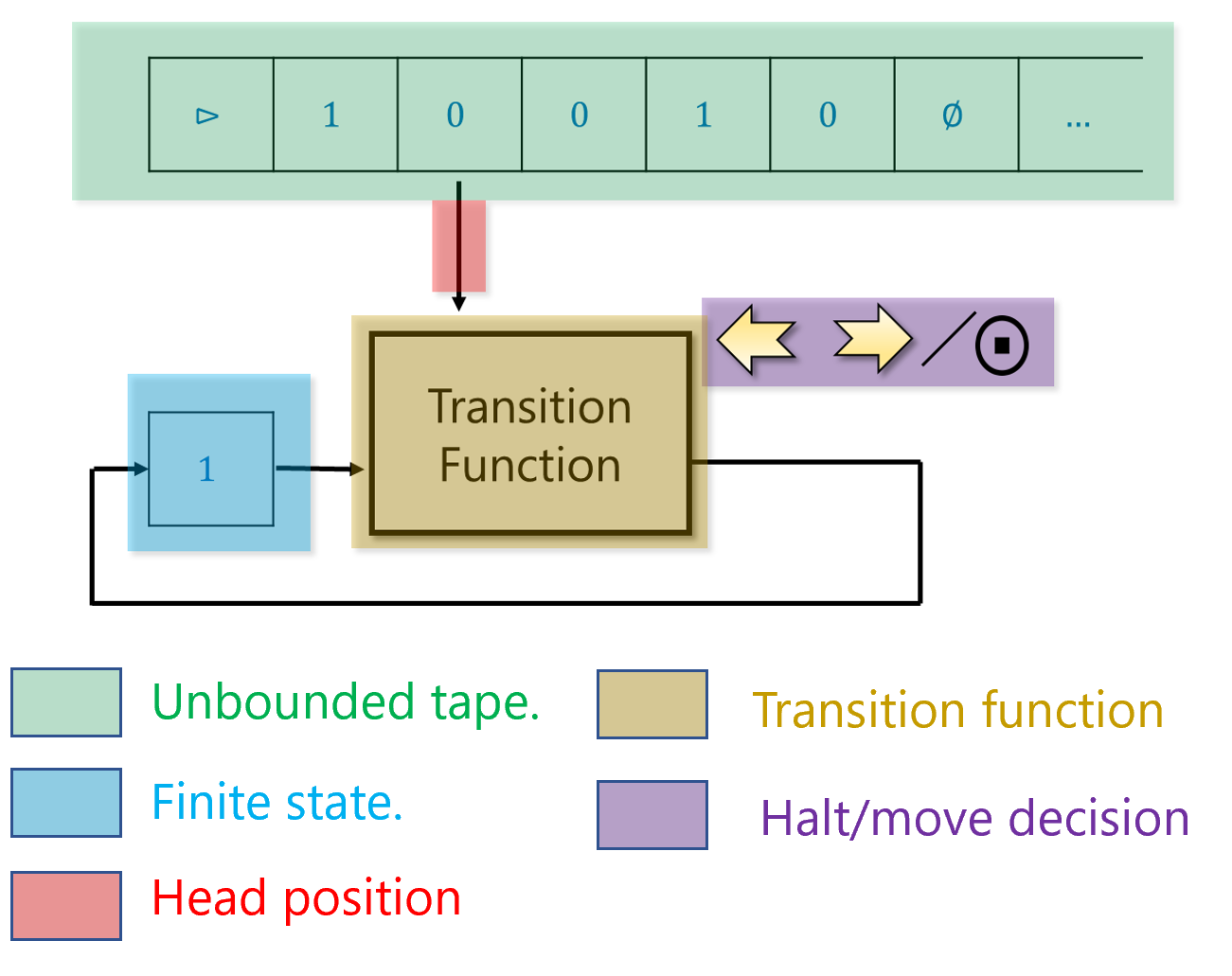
Turing thought of such a person as having access to as much “scratch paper” as they need. For simplicity, we can think of this scratch paper as a one dimensional piece of graph paper (or *tape*, as it is commonly referred to). The paper is divided into “cells”, where each “cell” can hold a single symbol (e.g., one digit or letter, and more generally, some element of a finite *alphabet*). At any point in time, the person can read from and write to a single cell of the paper. Based on the contents of this cell, the person can update their finite mental state, and/or move to the cell immediately to the left or right of the current one.



Steam-powered Turing machine mural, painted by CSE grad students at the University of Washington on the night before spring qualifying examinations, 1987. Image from <https://www.cs.washington.edu/building/art/SPTM>.

Turing modeled such a computation by a “machine” that maintains one of states. At each point in time the machine reads from its “work tape” a single symbol from a finite alphabet and uses that to update its state, write to tape, and possibly move to an adjacent cell (see turing-machine-fig). To compute a function using this machine, we initialize the tape with the input and our goal is to ensure that the tape will contain the value at the end of the computation. Specifically, a computation of a Turing machine with states and alphabet on input proceeds as follows:

* Initially the machine is at state (known as the “starting state”) and the tape is initialized to . We use the symbol to denote the beginning of the tape, and the symbol to denote an empty cell. We will always assume that the alphabet is a (potentially strict) superset of .
* The location to which the machine points to is set to .
* At each step, the machine reads the symbol that is in the location of the tape. Based on this symbol and its state , the machine decides on:
  + What symbol to write on the tape
  + Whether to move **L**eft (i.e., ), **R**ight (i.e., ), **S**tay in place, or **H**alt the computation.
  + What is going to be the new state
* The set of rules the Turing machine follows is known as its *transition function*.
* When the machine halts, its output is the binary string obtained by reading the tape from the beginning until the first location in which it contains a symbol, and then outputting all and symbols in sequence, dropping the initial symbol if it exists, as well as the final symbol.



The components of a Turing Machine. Note how they correspond to the general components of algorithms as described in algcomponentfig.

### Extended example: A Turing machine for palindromes

Let (for *palindromes*) be the function that on input , outputs if and only if is an (even length) *palindrome*, in the sense that for some and .

We now show a Turing machine that computes . To specify we need to specify **(i)** ’s tape alphabet which should contain at least the symbols ,, and , and **(ii)** ’s *transition function* which determines what action takes when it reads a given symbol while it is in a particular state.

In our case, will use the alphabet and will have states. Though the states are simply numbers between and , we will give them the following labels for convenience:

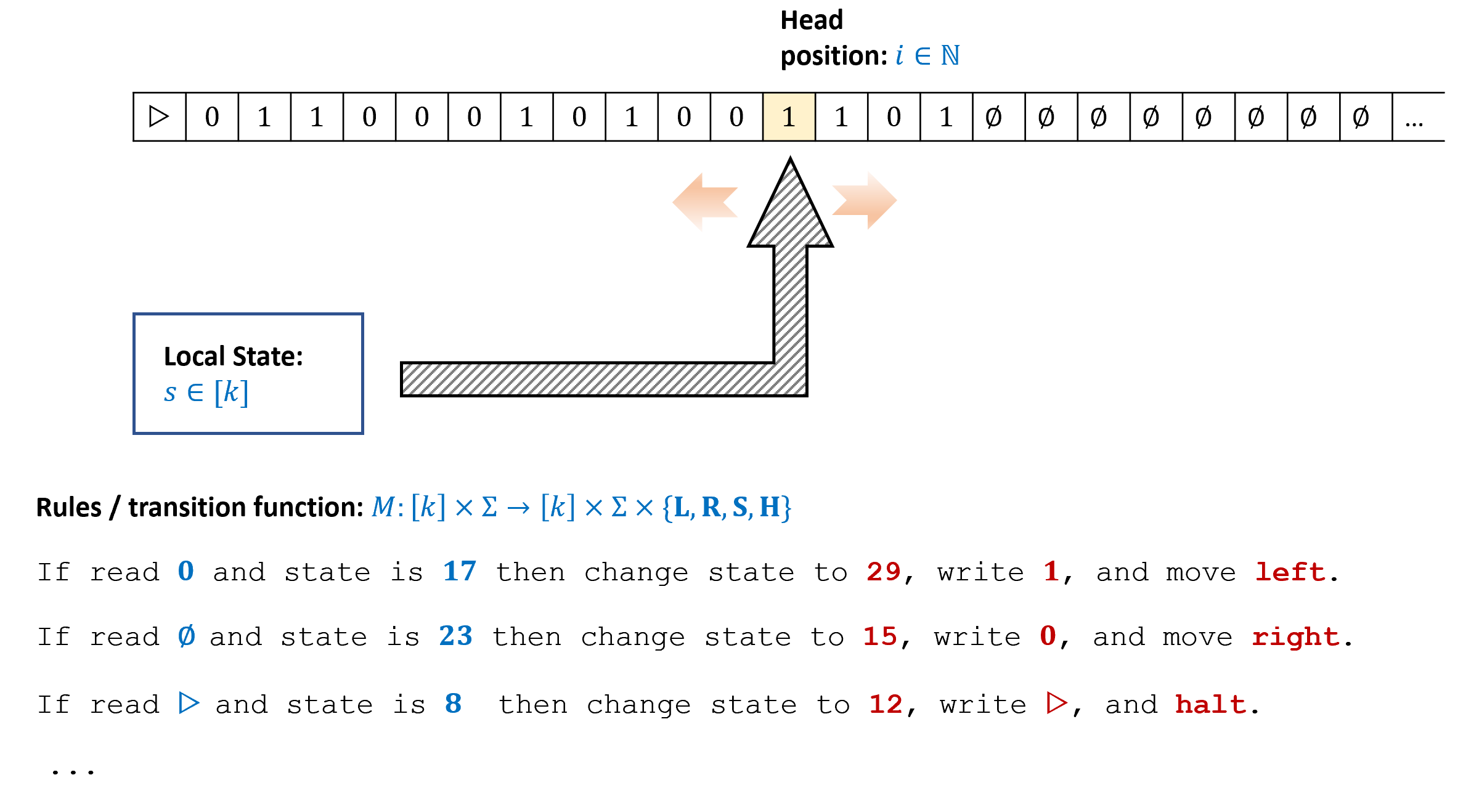
---  
caption: ''  
alignment: ''  
table-width: ''  
id: ''  
---  
State, Label  
0, `START`  
1,`RIGHT\_0`  
2,`RIGHT\_1`  
3,`LOOK\_FOR\_0`  
4,`LOOK\_FOR\_1`  
5,`RETURN`  
6,`OUTPUT\_0`  
7,`OUTPUT\_1`  
8,`0\_AND\_BLANK`  
9,`1\_AND\_BLANK`  
10,`BLANK\_AND\_STOP`

We describe the operation of our Turing machine in words:

* starts in state START and goes right, looking for the first symbol that is or . If it finds before it hits such a symbol then it moves to the OUTPUT\_1 state described below.
* Once finds such a symbol , deletes from the tape by writing the symbol, it enters either the RIGHT\_ mode and starts moving rightwards until it hits the first or symbol.
* Once finds this symbol, it goes into the state LOOK\_FOR\_0 or LOOK\_FOR\_1 depending on whether it was in the state RIGHT\_0 or RIGHT\_1 and makes one left move.
* In the state LOOK\_FOR\_, checks whether the value on the tape is . If it is, then deletes it by changing its value to , and moves to the state RETURN. Otherwise, it changes to the OUTPUT\_0 state.
* The RETURN state means that goes back to the beginning. Specifically, moves leftward until it hits the first symbol that is not or , in which case it changes its state to START.
* The OUTPUT\_ states mean that will eventually output the value . In both the OUTPUT\_0 and OUTPUT\_1 states, goes left until it hits . Once it does so, it makes a right step, and changes to the 1\_AND\_BLANK or 0\_AND\_BLANK states respectively. In the latter states, writes the corresponding value, moves right and changes to the BLANK\_AND\_STOP state, in which it writes to the tape and halts.

The above description can be turned into a table describing for each one of the combination of state and symbol, what the Turing machine will do when it is in that state and it reads that symbol. This table is known as the *transition function* of the Turing machine.

### Turing machines: a formal definition



A Turing machine has access to a *tape* of unbounded length. At each point in the execution, the machine can read a single symbol of the tape, and based on that and its current state, write a new symbol, update the tape, decide whether to move left, right, stay, or halt.

The formal definition of Turing machines is as follows:

A (one tape) *Turing machine* with states and alphabet is represented by a *transition function* .

For every , the *output* of on input , denoted by , is the result of the following process:

* We initialize to be the sequence , where . (That is, , for , and for .)
* We also initialize and .
* We then repeat the following process:
  1. Let .
  2. Set , .
  3. If then set , if then set . (If then we keep the same.)
  4. If , then halt.
* If the process above halts, then ’s output, denoted by , is the string obtained by concatenating all the symbols in in positions where is the first location in the tape containing .
* If The Turing machine does not halt then we denote .

You should make sure you see why this formal definition corresponds to our informal description of a Turing machine. To get more intuition on Turing machines, you can explore some of the online available simulators such as [Martin Ugarte’s](https://turingmachinesimulator.com/), [Anthony Morphett’s](http://morphett.info/turing/turing.html), or [Paul Rendell’s](http://rendell-attic.org/gol/TMapplet/index.htm).

One should not confuse the *transition function* of a Turing machine with the function that the machine computes. The transition function is a *finite* function, with inputs and outputs. (Can you see why?) The machine can compute an *infinite* function that takes as input a string of arbitrary length and might also produce an arbitrary length string as output.

In our formal definition, we identified the machine with its transition function since the transition function tells us everything we need to know about the Turing machine. However, this choice of representation is somewhat arbitrary, and is based on our convention that the state space is always the numbers with as the starting state. Other texts use different conventions, and so their mathematical definition of a Turing machine might look superficially different. However, these definitions describe the same computational process and have the same computational powers. Hence they are equivalent despite their superficial differences. See chaploopnotes for a comparison between TM-def and the way Turing Machines are defined in texts such as Sipser [@SipserBook].

### Computable functions

We now turn to make one of the most important definitions in this book: *computable functions*.

Let be a (total) function and let be a Turing machine. We say that *computes* if for every , .

We say that a function is *computable* if there exists a Turing machine that computes it.

Defining a function “computable” if and only if it can be computed by a Turing machine might seem “reckless” but, as we’ll see in chapequivalentmodels, being computable in the sense of computablefuncdef is equivalent to being computable in virtually any reasonable model of computation. This statement is known as the *Church-Turing Thesis*. (Unlike the *extended* Church-Turing Thesis which we discussed in PECTTsec, the Church-Turing thesis itself is widely believed and there are no candidate devices that attack it.)

We can precisely define what it means for a function to be computable by *any possible algorithm*.

This is a good point to remind the reader that *functions* are *not* the same as *programs*:

A Turing machine (or program) can *compute* some function , but it is not the same as . In particular, there can be more than one program to compute the same function. Being computable is a property of *functions*, not of machines.

We will often pay special attention to functions that have a single bit of output. Hence we give a special name for the set of computable functions of this form.

### 

We define be the set of all *computable* functions .

As discussed in languagessec, many texts use the terminology of “languages” rather than functions to refer to computational tasks. A Turing machine *decides* a language if for every input , outputs if and only if . This is equivalent to computing the Boolean function defined as iff . A language is *decidable* if there is a Turing machine that decides it. For historical reasons, some texts also call such languages *recursive* , which is the reason that the letter is often used to denote the set of computable Boolean functions / decidable languages defined in classRdef.

In this book we stick to the terminology of *functions* rather than languages, but all definitions and results can be easily translated back and forth by using the equivalence between the function and the language .

### Infinite loops and partial functions

One crucial difference between circuits/straight-line programs and Turing machines is the following. Looking at a NAND-CIRC program , we can always tell how many inputs and how many outputs has by simply looking at the X and Y variables. Furthermore, we are guaranteed that if we invoke on any input, then *some* output will be produced.

In contrast, given a Turing machine , we cannot determine a priori the length of ’s output. In fact, we don’t even know if an output would be produced at all! For example, it is straightforward to come up with a Turing machine whose transition function never outputs and hence never halts.

If a machine fails to stop and produce an output on some input , then it cannot compute any total function , since clearly on input , will fail to output . However, can still compute a *partial function*.[[1]](#footnote-52)

For example, consider the partial function that on input a pair of natural numbers, outputs if , and is undefined otherwise. We can define a Turing machine that computes on input by outputting the first such that . If and then the machine will never halt, but this is OK, since is undefined on such inputs. If and , the machine will output , which is also OK, since we don’t care about what the program outputs on inputs on which is undefined. Formally, we define computability of partial functions as follows:

Let be either a total or partial function mapping to and let be a Turing machine. We say that *computes* if for every on which is defined, . We say that a (partial or total) function is *computable* if there is a Turing machine that computes it.

Note that if is a total function, then it is defined on every and hence in this case, computablepartialfuncdef is identical to computablefuncdef.

We often use as our special “failure symbol”. If a Turing machine fails to halt on some input then we denote this by . This *does not* mean that outputs some encoding of the symbol but rather that enters into an infinite loop when given as input.

If a partial function is undefined on then we can also write . Therefore one might think that computablepartialfuncdef can be simplified to requiring that for every , which would imply that for every , halts on if and only if is defined on . However, this is not the case: for a Turing machine to compute a partial function it is not *necessary* for to enter an infinite loop on inputs on which is not defined. All that is needed is for to output on values of on which is defined: on other inputs it is OK for to output an arbitrary value such as , , or anything else, or not to halt at all. To borrow a term from the C programming language, on inputs on which is not defined, what does is “undefined behavior”.

## Turing machines as programming languages

The name “Turing machine”, with its “tape” and “head” evokes a physical object, while in contrast we think of a *program* as a piece of text. But we can think of a Turing machine as a program as well. For example, consider the Turing machine of turingmachinepalindrome that computes the function such that iff is a palindrome. We can also describe this machine as a *program* using the Python-like pseudocode of the form below

# Gets an array Tape initialized to  
# [">", x\_0 , x\_1 , .... , x\_(n-1), "∅", "∅", ...]  
# At the end of the execution, Tape[1] is equal to 1  
# if x is a palindrome and is equal to 0 otherwise  
def PAL(Tape):  
 head = 0  
 state = 0 # START  
 while (state != 12):  
 if (state == 0 && Tape[head]=='0'):  
 state = 3 # LOOK\_FOR\_0  
 Tape[head] = 'x'  
 head += 1 # move right  
 if (state==0 && Tape[head]=='1')  
 state = 4 # LOOK\_FOR\_1  
 Tape[head] = 'x'  
 head += 1 # move right  
 ... # more if statements here

The precise details of this program are not important. What matters is that we can describe Turing machines as *programs*. Moreover, note that when translating a Turing machine into a program, the *tape* becomes a *list* or *array* that can hold values from the finite set .[[2]](#footnote-56) The *head position* can be thought of as an integer-valued variable that holds integers of unbounded size. The *state* is a *local register* that can hold one of a fixed number of values in .

More generally we can think of every Turing machine as equivalent to a program similar to the following:

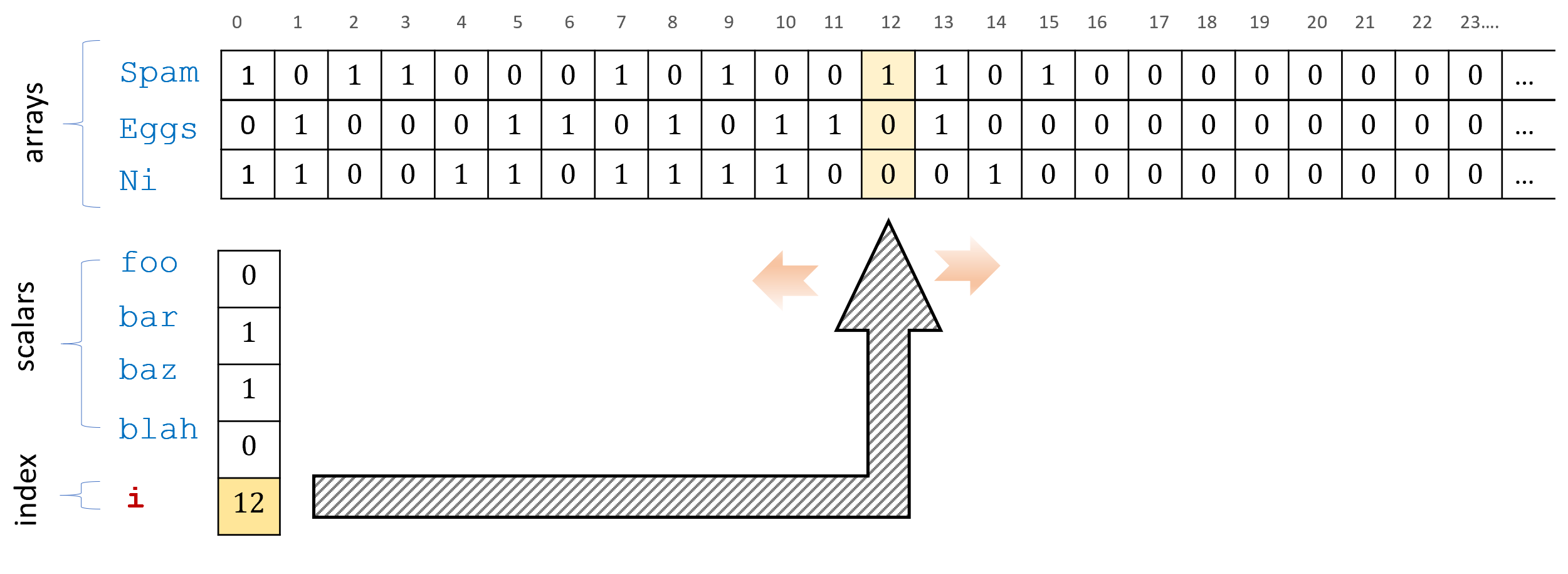
# Gets an array Tape initialized to  
# [">", x\_0 , x\_1 , .... , x\_(n-1), "∅", "∅", ...]  
def M(Tape):  
 state = 0  
 i = 0 # holds head location  
 while (True):  
 # Move head, modify state, write to tape  
 # based on current state and cell at head  
 # below are just examples for how program looks for a particular transition function  
 if Tape[i]=="0" and state==7: # δ\_M(7,"0")=(19,"1","R")  
 Tape[i]="1"  
 i += 1  
  
 state = 19  
 elif Tape[i]==">" and state == 13: # δ\_M(13,">")=(15,"0","S")  
 Tape[i]="0"  
 state = 15  
 elif ...  
 ...  
 elif Tape[i]==">" and state == 29: # δ\_M(29,">")=(.,.,"H")  
 break # Halt

If we wanted to use only *Boolean* (i.e., /-valued) variables, then we can encode the state variables using bits. Similarly, we can represent each element of the alphabet using bits and hence we can replace the -valued array Tape[] with Boolean-valued arrays Tape0[],, Tape[].

### The NAND-TM Programming language

We now introduce the *NAND-TM programming language*, which captures the power of a Turing machine with a programming-language formalism. Like the difference between Boolean circuits and Turing machines, the main difference between NAND-TM and NAND-CIRC is that NAND-TM models a *single uniform algorithm* that can compute a function that takes inputs of *arbitrary lengths*. To do so, we extend the NAND-CIRC programming language with two constructs:

* *Loops*: NAND-CIRC is a *straight-line* programming language- a NAND-CIRC program of lines takes exactly steps of computation and hence in particular, cannot even touch more than variables. *Loops* allow us to use a fixed-length program to encode the instructions for a computation that can take an arbitrary amount of time.
* *Arrays*: A NAND-CIRC program of lines touches at most variables. While we can use variables with names such as Foo\_17 or Bar[22] in NAND-CIRC, they are not true arrays, since the number in the identifier is a constant that is “hardwired” into the program. NAND-TM contains actual arrays that can have a length that is not a priori bounded.



A NAND-TM program has *scalar* variables that can take a Boolean value, *array* variables that hold a sequence of Boolean values, and a special *index* variable i that can be used to index the array variables. We refer to the i-th value of the array variable Spam using Spam[i]. At each iteration of the program the index variable can be incremented or decremented by one step using the MODANDJUMP operation.

Thus a good way to remember NAND-TM is using the following informal equation:

### 

As we will see, adding loops and arrays to NAND-CIRC is enough to capture the full power of all programming languages! Hence we could replace “NAND-TM” with any of *Python*, *C*, *Javascript*, *OCaml*, etc. in the left-hand side of eqnandloops. But we’re getting ahead of ourselves: this issue will be discussed in chapequivalentmodels.

Concretely, the NAND-TM programming language adds the following features on top of NAND-CIRC (see nandtmfig):

* We add a special *integer valued* variable i. All other variables in NAND-TM are *Boolean valued* (as in NAND-CIRC).
* Apart from i NAND-TM has two kinds of variables: *scalars* and *arrays*. *Scalar* variables hold one bit (just as in NAND-CIRC). *Array* variables hold an unbounded number of bits. At any point in the computation we can access the array variables at the location indexed by i using Foo[i]. We cannot access the arrays at locations other than the one pointed to by i.
* We use the convention that *arrays* always start with a capital letter, and *scalar variables* (which are never indexed with i) start with lowercase letters. Hence Foo is an array and bar is a scalar variable.
* The input and output X and Y are now considered *arrays* with values of zeroes and ones. (There are also two other special arrays X\_nonblank and Y\_nonblank, see below.)
* We add a special MODANDJUMP instruction that takes two Boolean variables as input and does the following:
  + If and then MODANDJUMP() increments i by one and jumps to the first line of the program.
  + If and then MODANDJUMP() decrements i by one and jumps to the first line of the program. (If i is already equal to then it stays at .)
  + If and then MODANDJUMP() jumps to the first line of the program without modifying i.
  + If then MODANDJUMP() halts execution of the program.
* The MODANDJUMP instruction always appears in the last line of a NAND-TM program and nowhere else.

**Default values.** We need one more convention to handle “default values”. Turing machines have the special symbol to indicate that tape location is “blank” or “uninitialized”. In NAND-TM there is no such symbol, and all variables are *Boolean*, containing either or . All variables and locations of arrays default to if they have not been initialized to another value. To keep track of whether a in an array corresponds to a true zero or to an uninitialized cell, a programmer can always add to an array Foo a “companion array” Foo\_nonblank and set Foo\_nonblank[i] to whenever the i th location is initialized. In particular, we will use this convention for the input and output arrays X and Y. A NAND-TM program has *four* special arrays X, X\_nonblank, Y, and Y\_nonblank. When a NAND-TM program is executed on input of length , the first cells of the array X are initialized to and the first cells of the array X\_nonblank are initialized to . (All uninitialized cells default to .) The output of a NAND-TM program is the string Y[], , Y[] where is the smallest integer such that Y\_nonblank[]. A NAND-TM program gets called with X and X\_nonblank initialized to contain the input, and writes to Y and Y\_nonblank to produce the output.

Formally, NAND-TM programs are defined as follows:

A *NAND-TM program* consists of a sequence of lines of the form foo = NAND(bar,blah) ending with a line of the form MODANDJUMP(foo,bar), where foo,bar,blah are either *scalar variables* (sequences of letters, digits, and underscores) or *array variables* of the form Foo[i] (starting with capital letters and indexed by i). The program has the array variables X, X\_nonblank, Y, Y\_nonblank and the index variable i built in, and can use additional array and scalar variables.

If is a NAND-TM program and is an input then an execution of on is the following process:

1. The arrays X and X\_nonblank are initialized by X[] and X\_nonblank[] for all . All other variables and cells are initialized to . The index variable i is also initialized to .
2. The program is executed line by line. When the last line MODANDJUMP(foo,bar) is executed we do as follows:
   1. If foo and bar, jump to the first line without modifying the value of i.
   2. If foo and bar, increment i by one and jump to the first line.
   3. If foo and bar, decrement i by one (unless it is already zero) and jump to the first line.
   4. If foo and bar, halt and output Y[], , Y[] where is the smallest integer such that Y\_nonblank[].

### Sneak peak: NAND-TM vs Turing machines

As the name implies, NAND-TM programs are a direct implementation of Turing machines in programming language form. We will show the equivalence below, but you can already see how the components of Turing machines and NAND-TM programs correspond to one another:

---  
caption: 'Turing Machine and NAND-TM analogs'  
alignment: 'LL'  
table-width: '1/1'  
id: TMvsNANDTMtable  
---  
\*\*Turing Machine\*\* | \*\*NAND-TM program\*\*  
\*State:\* single register that takes values in $[k]$ | \*Scalar variables:\* Several variables such as `foo`, `bar` etc.. each taking values in $\{0,1\}$.  
\*Tape:\* One tape containing values in a finite set $\Sigma$. Potentially infinite but $T[t]$ defaults to $\varnothing$ for all locations $t$ that have not been accessed. | \*Arrays:\* Several arrays such as `Foo`, `Bar` etc.. for each such array `Arr` and index $j$, the value of `Arr` at position $j$ is either $0$ or $1$. The value defaults to $0$ for position that have not been written to.  
\*Head location:\* A number $i\in \mathbb{N}$ that encodes the position of the head. | \*Index variable:\* The variable `i` that can be used to access the arrays.  
\*Accessing memory:\* At every step the Turing machine has access to its local state, but can only access the tape at the position of the current head location. | \*Accessing memory:\* At every step a NAND-TM program has access to all the scalar variables, but can only access the arrays at the location `i` of the index variable  
\*Control of location:\* In each step the machine can move the head location by at most one position. | \*Control of index variable:\* In each iteration of its main loop the program can modify the index `i` by at most one.

### Examples

We now present some examples of NAND-TM programs.

The following is a NAND-TM program to compute the *increment function*. That is, such that for every , is the bit long string such that if is the number represented by , then is the (least-significant digit first) binary representation of the number .

We start by describing the program using “syntactic sugar” for NAND-CIRC for the IF, XOR and AND functions (as well as the constant one function, and the function COPY that just maps a bit to itself).

carry = IF(started,carry,one(started))  
started = one(started)  
Y[i] = XOR(X[i],carry)  
carry = AND(X[i],carry)  
Y\_nonblank[i] = one(started)  
MODANDJUMP(X\_nonblank[i],X\_nonblank[i])

Since we used syntactic sugar, the above is not, strictly speaking, a valid NAND-TM program. However, by “opening up” all the syntactic sugar, we get the following “sugar free” valid program to compute the same function.

temp\_0 = NAND(started,started)  
temp\_1 = NAND(started,temp\_0)  
temp\_2 = NAND(started,started)  
temp\_3 = NAND(temp\_1,temp\_2)  
temp\_4 = NAND(carry,started)  
carry = NAND(temp\_3,temp\_4)  
temp\_6 = NAND(started,started)  
started = NAND(started,temp\_6)  
temp\_8 = NAND(X[i],carry)  
temp\_9 = NAND(X[i],temp\_8)  
temp\_10 = NAND(carry,temp\_8)  
Y[i] = NAND(temp\_9,temp\_10)  
temp\_12 = NAND(X[i],carry)  
carry = NAND(temp\_12,temp\_12)  
temp\_14 = NAND(started,started)  
Y\_nonblank[i] = NAND(started,temp\_14)  
MODANDJUMP(X\_nonblank[i],X\_nonblank[i])

The following is a NAND-TM program to compute the XOR function on inputs of arbitrary length. That is such that for every . Once again, we use a certain “syntactic sugar”. Specifically, we access the arrays X and Y at their zero-th entry, while NAND-TM only allows access to arrays in the coordinate of the variable i.

temp\_0 = NAND(X[0],X[0])  
Y\_nonblank[0] = NAND(X[0],temp\_0)  
temp\_2 = NAND(X[i],Y[0])  
temp\_3 = NAND(X[i],temp\_2)  
temp\_4 = NAND(Y[0],temp\_2)  
Y[0] = NAND(temp\_3,temp\_4)  
MODANDJUMP(X\_nonblank[i],X\_nonblank[i])

To transform the program above to a valid NAND-TM program, we can transform references such as X[0] and Y[0] to scalar variables x\_0 and y\_0 (similarly we can transform any reference of the form Foo[17] or Bar[15] to scalars such as foo\_17 and bar\_15). We then need to add code to load the value of X[0] to x\_0 and similarly to write to Y[0] the value of y\_0, but this is not hard to do. Using the fact that variables are initialized to zero by default, we can create a variable init which will be set to at the end of the first iteration and not changed since then. We can then add an array Atzero and code that will modify Atzero[i] to if init is and otherwise leave it as it is. This will ensure that Atzero[i] is equal to if and only if i is set to zero, and allow the program to know when we are at the zeroth location. Thus we can add code to read and write to the corresponding scalars x\_0, y\_0 when we are at the zeroth location, and also code to move i to zero and then halt at the end. Working this out fully is somewhat tedious, but can be a good exercise.

Working out the above two examples can go a long way towards understanding the NAND-TM language. See our [GitHub repository](https://github.com/boazbk/tcscode) for a full specification of the NAND-TM language.

## Equivalence of Turing machines and NAND-TM programs

Given the above discussion, it might not be surprising that Turing machines turn out to be equivalent to NAND-TM programs. Indeed, we designed the NAND-TM language to have this property. Nevertheless, this is a significant result, and the first of many other such equivalence results we will see in this book.

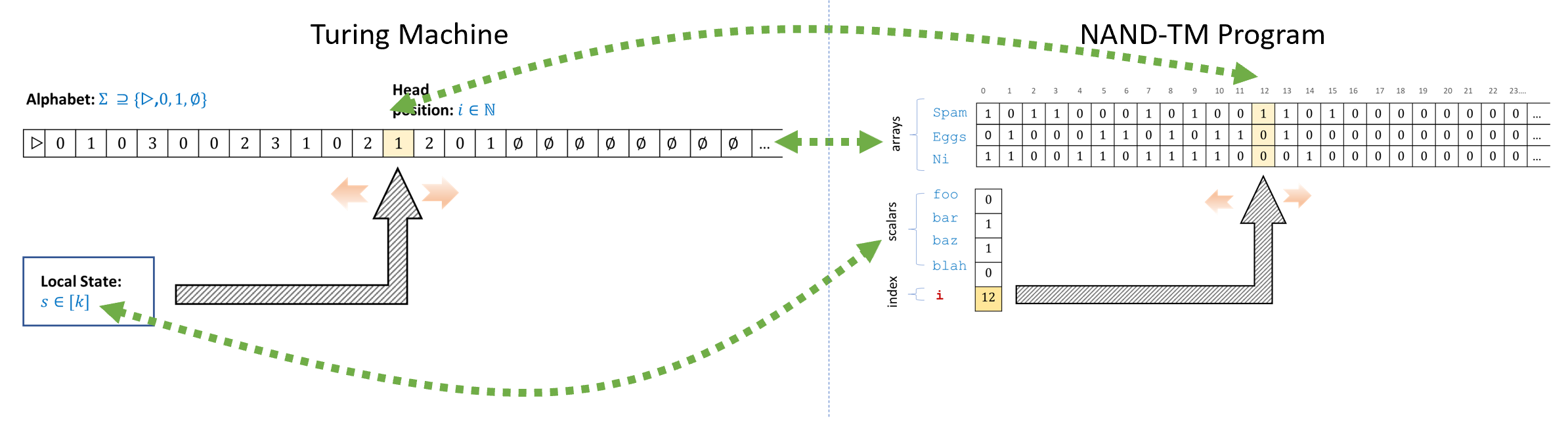
### 

For every , is computable by a NAND-TM program if and only if there is a Turing machine that computes .

To prove such an equivalence theorem, we need to show two directions. We need to be able to **(1)** transform a Turing machine to a NAND-TM program that computes the same function as and **(2)** transform a NAND-TM program into a Turing machine that computes the same function as .

The idea of the proof is illustrated in tmvsnandppfig. To show **(1)**, given a Turing machine , we will create a NAND-TM program that will have an array Tape for the tape of and scalar (i.e., non-array) variable(s) state for the state of . Specifically, since the state of a Turing machine is not in but rather in a larger set , we will use variables state\_ , , state\_ variables to store the representation of the state. Similarly, to encode the larger alphabet of the tape, we will use arrays Tape\_ , , Tape\_, such that the location of these arrays encodes the symbol in the tape for every tape. Using the fact that *every* function can be computed by a NAND-CIRC program, we will be able to compute the transition function of , replacing moving left and right by decrementing and incrementing i respectively.

We show **(2)** using very similar ideas. Given a program that uses array variables and scalar variables, we will create a Turing machine with about states to encode the values of scalar variables, and an alphabet of about so we can encode the arrays using our tape. (The reason the sizes are only “about” and is that we need to add some symbols and steps for bookkeeping purposes.) The Turing machine simulates each iteration of the program by updating its state and tape accordingly.



Comparing a Turing machine to a NAND-TM program. Both have an unbounded memory component (the *tape* for a Turing machine, and the *arrays* for a NAND-TM program), as well as a constant local memory (*state* for a Turing machine, and *scalar variables* for a NAND-TM program). Both can only access at each step one location of the unbounded memory, this is the “head” location for a Turing machine, and the value of the index variable i for a NAND-TM program.

We start by proving the “if” direction of TM-equiv-thm. Namely we show that given a Turing machine , we can find a NAND-TM program such that for every input , if halts on input with output then . Since our goal is just to show such a program *exists*, we don’t need to write out the full code of line by line, and can take advantage of our various “syntactic sugar” in describing it.

The key observation is that by NAND-univ-thm we can compute *every* finite function using a NAND-CIRC program. In particular, consider the transition function of our Turing machine. We can encode its components as follows:

* We encode using and using , where and .
* We encode the set using . We will choose the encoding , , , . (This conveniently corresponds to the semantics of the MODANDJUMP operation.)

Hence we can identify with a function , mapping strings of length to strings of length . By NAND-univ-thm there exists a finite length NAND-CIRC program ComputeM that computes this function . The idea behind the NAND-TM program to simulate is to:

1. Use variables state\_ state\_ to encode ’s state.
2. Use arrays Tape\_[] Tape\_[] to encode ’s tape.
3. Use the fact that transition is finite and computable by NAND-CIRC program.

Given the above, we can write code of the form:

state\_ state\_, Tape\_[i] Tape\_[i], dir0,dir1 TRANSITION( state\_ state\_, Tape\_[i] Tape\_[i] )

MODANDJUMP(dir0,dir1)

Every step of the main loop of the above program perfectly mimics the computation of the Turing machine , and so the program carries out exactly the definition of computation by a Turing machine as per TM-def.

For the other direction, suppose that is a NAND-TM program with lines, scalar variables, and array variables. We will show that there exists a Turing machine with states and alphabet of size that computes the same functions as (where , are some constants to be determined later).

Specifically, consider the function that on input the contents of ’s scalar variables and the contents of the array variables at location i in the beginning of an iteration, outputs all the new values of these variables at the last line of the iteration, right before the MODANDJUMP instruction is executed.

If foo and bar are the two variables that are used as input to the MODANDJUMP instruction, then based on the values of these variables we can compute whether i will increase, decrease or stay the same, and whether the program will halt or jump back to the beginning. Hence a Turing machine can simulate an execution of in one iteration using a finite function applied to its alphabet. The overall operation of the Turing machine will be as follows:

1. The machine encodes the contents of the array variables of in its tape and the contents of the scalar variables in (part of) its state. Specifically, if has local variables and arrays, then the state space of will be large enough to encode all assignments to the local variables, and the alphabet of will be large enough to encode all assignments for the array variables at each location. The head location corresponds to the index variable i.
2. Recall that every line of the program corresponds to reading and writing either a scalar variable, or an array variable at the location i. In one iteration of the value of i remains fixed, and so the machine can simulate this iteration by reading the values of all array variables at i (which are encoded by the single symbol in the alphabet located at the i-th cell of the tape) , reading the values of all scalar variables (which are encoded by the state), and updating both. The transition function of can output depending on whether the values given to the MODANDJUMP operation are , or respectively.
3. When the program halts (i.e., MODANDJUMP gets ) then the Turing machine will enter into a special loop to copy the results of the Y array into the output and then halt. We can achieve this by adding a few more states.

The above is not a full formal description of a Turing machine, but our goal is just to show that such a machine exists. One can see that simulates every step of , and hence computes the same function as .

If we examine the proof of TM-equiv-thm then we can see that every iteration of the loop of a NAND-TM program corresponds to one step in the execution of the Turing machine. We will come back to this question of measuring the number of computation steps later in this course. For now, the main take away point is that NAND-TM programs and Turing machines are essentially equivalent in power even when taking running time into account.

### Specification vs implementation (again)

Once you understand the definitions of both NAND-TM programs and Turing machines, TM-equiv-thm is straightforward. Indeed, NAND-TM programs are not as much a different model from Turing machines as they are simply a reformulation of the same model using programming language notation. You can think of the difference between a Turing machine and a NAND-TM program as the difference between representing a number using decimal or binary notation. In contrast, the difference between a *function* and a Turing machine that computes is much more profound: it is like the difference between the equation , and the number that is a solution for this equation. For this reason, while we take special care in distinguishing *functions* from *programs* or *machines*, we will often identify the two latter concepts. We will move freely between describing an algorithm as a Turing machine or as a NAND-TM program (as well as some of the other equivalent computational models we will see in chapequivalentmodels and beyond).

---  
caption: 'Specification vs Implementation formalisms'  
alignment: 'LL'  
table-width: ''  
id: specvsimp  
---  
\*Setting\* ; \*Specification\* ; \*Implementation\*  
\_Finite computation\_ ; \_\_Functions\_\_ mapping $\{0,1\}^n$ to $\{0,1\}^m$ ; \_\_Circuits\_\_, \_\_Straightline programs\_\_  
\_Infinite computation\_ ; \_\_Functions\_\_ mapping $\{0,1\}^\*$ to $\{0,1\}$ or to $\{0,1\}^\*$. ; \_\_Algorithms\_\_, \_\_Turing Machines\_\_, \_\_Programs\_\_

## NAND-TM syntactic sugar

Just like we did with NAND-CIRC in finiteuniversalchap, we can use “syntactic sugar” to make NAND-TM programs easier to write. For starters, we can use all of the syntactic sugar of NAND-CIRC, such as macro definitions and conditionals (i.e., if/then). However, we can go beyond this and achieve (for example):

* Inner loops such as the while and for operations common to many programming languages.
* Multiple index variables (e.g., not just i but we can add j, k, etc.).
* Arrays with more than one dimension (e.g., Foo[i][j], Bar[i][j][k] etc.)

In all of these cases (and many others) we can implement the new feature as mere “syntactic sugar” on top of standard NAND-TM. This means that the set of functions computable by NAND-TM with this feature is the same as the set of functions computable by standard NAND-TM. Similarly, we can show that the set of functions computable by Turing machines that have more than one tape, or tapes of more dimensions than one, is the same as the set of functions computable by standard Turing machines.

### “GOTO” and inner loops

We can implement more advanced *looping constructs* than the simple MODANDJUMP. For example, we can implement GOTO. A GOTO statement corresponds to jumping to a specific line in the execution. For example, if we have code of the form

"start": do foo  
 GOTO("end")  
"skip": do bar  
"end": do blah

then the program will only do foo and blah as when it reaches the line GOTO("end") it will jump to the line labeled with "end". We can achieve the effect of GOTO in NAND-TM using conditionals. In the code below, we assume that we have a variable pc that can take strings of some constant length. This can be encoded using a finite number of Boolean variables pc\_0, pc\_1, , pc\_, and so when we write below pc = "label" what we mean is something like pc\_0 = 0,pc\_1 = 1, (where the bits correspond to the encoding of the finite string "label" as a string of length ). We also assume that we have access to conditional (i.e., if statements), which we can emulate using syntactic sugar in the same way as we did in NAND-CIRC.

To emulate a GOTO statement, we will first modify a program P of the form

do foo  
do bar  
do blah

to have the following form (using syntactic sugar for if):

pc = "line1"  
if (pc=="line1"):  
 do foo  
 pc = "line2"  
if (pc=="line2"):  
 do bar  
 pc = "line3"  
if (pc=="line3"):  
 do blah

These two programs do the same thing. The variable pc corresponds to the “program counter” and tells the program which line to execute next. We can see that if we wanted to emulate a GOTO("line3") then we could simply modify the instruction pc = "line2" to be pc = "line3".

In NAND-CIRC we could only have GOTOs that go forward in the code, but since in NAND-TM everything is encompassed within a large outer loop, we can use the same ideas to implement GOTOs that can go backward, as well as conditional loops.

**Other loops.** Once we have GOTO, we can emulate all the standard loop constructs such as while, do .. until or for in NAND-TM as well. For example, we can replace the code

while foo:  
 do blah  
do bar

with

"loop":  
 if NOT(foo): GOTO("next")  
 do blah  
 GOTO("loop")  
"next":  
 do bar

The GOTO statement was a staple of most early programming languages, but has largely fallen out of favor and is not included in many modern languages such as *Python*, *Java*, *Javascript*. In 1968, Edsger Dijsktra wrote a famous letter titled “[Go to statement considered harmful.](https://goo.gl/bnNsjo)” (see also xkcdgotofig). The main trouble with GOTO is that it makes analysis of programs more difficult by making it harder to argue about *invariants* of the program.

When a program contains a loop of the form:

for j in range(100):  
 do something  
  
do blah

you know that the line of code do blah can only be reached if the loop ended, in which case you know that j is equal to , and might also be able to argue other properties of the state of the program. In contrast, if the program might jump to do blah from any other point in the code, then it’s very hard for you as the programmer to know what you can rely upon in this code. As Dijkstra said, such invariants are important because \_ “our intellectual powers are rather geared to master static relations and .. our powers to visualize processes evolving in time are relatively poorly developed”\_ and so *" we should … do …our utmost best to shorten the conceptual gap between the static program and the dynamic process."*

That said, GOTO is still a major part of lower level languages where it is used to implement higher-level looping constructs such as while and for loops. For example, even though *Java* doesn’t have a GOTO statement, the Java Bytecode (which is a lower-level representation of Java) does have such a statement. Similarly, Python bytecode has instructions such as POP\_JUMP\_IF\_TRUE that implement the GOTO functionality, and similar instructions are included in many assembly languages. The way we use GOTO to implement a higher-level functionality in NAND-TM is reminiscent of the way these various jump instructions are used to implement higher-level looping constructs.



XKCD’s take on the GOTO statement.

## Uniformity, and NAND vs NAND-TM (discussion)

While NAND-TM adds extra operations over NAND-CIRC, it is not exactly accurate to say that NAND-TM programs or Turing machines are “more powerful” than NAND-CIRC programs or Boolean circuits. NAND-CIRC programs, having no loops, are simply not applicable for computing functions with an unbounded number of inputs. Thus, to compute a function using NAND-CIRC (or equivalently, Boolean circuits) we need a *collection* of programs/circuits: one for every input length.

The key difference between NAND-CIRC and NAND-TM is that NAND-TM allows us to express the fact that the algorithm for computing parities of length- strings is really the same one as the algorithm for computing parities of length- strings (or similarly the fact that the algorithm for adding -bit numbers is the same for every , etc.). That is, one can think of the NAND-TM program for general parity as the “seed” out of which we can grow NAND-CIRC programs for length , length , or length parities as needed.

This notion of a single algorithm that can compute functions of all input lengths is known as *uniformity* of computation. Hence we think of Turing machines / NAND-TM as *uniform* models of computation, as opposed to Boolean circuits or NAND-CIRC, which are *non-uniform* models, in which we have to specify a different program for every input length.

Looking ahead, we will see that this uniformity leads to another crucial difference between Turing machines and circuits. Turing machines can have inputs and outputs that are longer than the description of the machine as a string, and in particular there exists a Turing machine that can “self replicate” in the sense that it can print its own code. The notion of “self replication”, and the related notion of “self reference” are crucial to many aspects of computation, and beyond that to life itself, whether in the form of digital or biological programs.

For now, what you ought to remember is the following differences between *uniform* and *non-uniform* computational models:

* **Non-uniform computational models:** Examples are *NAND-CIRC programs* and *Boolean circuits*. These are models where each individual program/circuit can compute a *finite* function . We have seen that *every* finite function can be computed by *some* program/circuit. To discuss computation of an *infinite* function we need to allow a *sequence* of programs/circuits (one for every input length), but this does not capture the notion of a *single algorithm* to compute the function .
* **Uniform computational models:** Examples are *Turing machines* and *NAND-TM programs*. These are models where a single program/machine can take inputs of *arbitrary length* and hence compute an *infinite* function . The number of steps that a program/machine takes on some input is not a priori bounded in advance and in particular there is a chance that it will enter into an *infinite loop*. Unlike the non-uniform case, we have *not* shown that every infinite function can be computed by some NAND-TM program/Turing machine. We will come back to this point in chapcomputable.

### 

* *Turing machines* capture the notion of a single algorithm that can evaluate functions of every input length.
* They are equivalent to *NAND-TM programs*, which add loops and arrays to NAND-CIRC.
* Unlike NAND-CIRC or Boolean circuits, the number of steps that a Turing machine takes on a given input is not fixed in advance. In fact, a Turing machine or a NAND-TM program can enter into an *infinite loop* on certain inputs, and not halt at all.

## Exercises

Produce the code of a (syntactic-sugar free) NAND-TM program that computes the (unbounded input length) *Majority* function where for every , if and only if . We say “produce” rather than “write” because you do not have to write the code of by hand, but rather can use the programming language of your choice to compute this code.

Prove that the following functions are computable. For all of these functions, you do not have to fully specify the Turing machine or the NAND-TM program that computes the function, but rather only prove that such a machine or program exists:

1. which takes as input a representation of a natural number and outputs the representation of .
2. which takes as input a representation of a pair of natural numbers and outputs the representation of .
3. , which takes a representation of a pair of natural numbers and outputs the representation of .
4. which takes as input the representation of a list of natural numbers and returns its sorted version such that for every there is some with and .

Define NAND-TM’ to be the variant of NAND-TM where there are *two* index variables i and j. Arrays can be indexed by either i or j. The operation MODANDJUMP takes four variables and uses the values of to decide whether to increment j, decrement j or keep it in the same value (corresponding to , , and respectively). Prove that for every function , is computable by a NAND-TM program if and only if is computable by a NAND-TM’ program.

Define a *two tape Turing machine* to be a Turing machine which has two separate tapes and two separate heads. At every step, the transition function gets as input the location of the cells in the two tapes, and can decide whether to move each head independently. Prove that for every function , is computable by a standard Turing machine if and only if is computable by a two-tape Turing machine.

Define NAND-TM" to be the variant of NAND-TM where just like NAND-TM’ defined in twoindexex there are two index variables i and j, but now the arrays are *two dimensional* and so we index an array Foo by Foo[i][j]. Prove that for every function , is computable by a NAND-TM program if and only if is computable by a NAND-TM’’ program.

Define a *two-dimensional Turing machine* to be a Turing machine in which the tape is *two dimensional*. At every step the machine can move p, own, eft, ight, or tay. Prove that for every function , is computable by a standard Turing machine if and only if is computable by a two-dimensional Turing machine.

Prove the following closure properties of the set defined in classRdef:

1. If then the function is in .
2. If then the function is in .
3. If then the function in in where is defined as follows: iff there exist some strings such that and for every .
4. If then the function
5. is in .

Define a Turing machine to be *oblivious* if its head movements are independent of its input. That is, we say that is oblivious if there exists an infinite sequence such that for every , the movements of when given input (up until the point it halts, if such point exists) are given by .

Prove that for every function , if is computable then it is computable by an oblivious Turing machine. See footnote for hint.[[3]](#footnote-88)

### 

Prove that for every , the function is computable if and only if the following function is computable, where is defined as follows:

Recall that is the set of all total functions from to that are computable by a Turing machine (see classRdef). Prove that is *countable*. That is, prove that there exists a one-to-one map . You can use the equivalence between Turing machines and NAND-TM programs.

Prove that the set of *all* total functions from is *not* countable. You can use the results of cantorsec. (We will see an *explicit* uncomputable function in chapcomputable.)

## Bibliographical notes

Augusta Ada Byron, countess of Lovelace (1815-1852) lived a short but turbulent life, though is today most well known for her collaboration with Charles Babbage (see [@stein1987ada] for a biography). Ada took an immense interest in Babbage’s *analytical engine*, which we mentioned in compchap. In 1842-3, she translated from Italian a paper of Menabrea on the engine, adding copious notes (longer than the paper itself). The quote in the chapter’s beginning is taken from Nota A in this text. Lovelace’s notes contain several examples of *programs* for the analytical engine, and because of this she has been called “the world’s first computer programmer” though it is not clear whether they were written by Lovelace or Babbage himself [@holt2001ada]. Regardless, Ada was clearly one of very few people (perhaps the only one outside of Babbage himself) to fully appreciate how important and revolutionary the idea of mechanizing computation truly is.

The books of Shetterly [@shetterly2016hidden] and Sobel [@sobel2017the] discuss the history of human computers (who were female, more often than not) and their important contributions to scientific discoveries in astronomy and space exploration.

Alan Turing was one of the intellectual giants of the 20th century. He was not only the first person to define the notion of computation, but also invented and used some of the world’s earliest computational devices as part of the effort to break the *Enigma* cipher during World War II, saving [millions of lives](https://goo.gl/KY1bJN). Tragically, Turing committed suicide in 1954, following his conviction in 1952 for homosexual acts and a court-mandated hormonal treatment. In 2009, British prime minister Gordon Brown made an official public apology to Turing, and in 2013 Queen Elizabeth II granted Turing a posthumous pardon. Turing’s life is the subject of a [great book](https://goo.gl/3GdFdp) and a [mediocre movie](https://goo.gl/EtQvSu).

Sipser’s text [@SipserBook] defines a Turing machine as a *seven tuple* consisting of the state space, input alphabet, tape alphabet, transition function, starting state, accepting state, and rejecting state. Superficially this looks like a very different definition than TM-def but it is simply a different representation of the same concept, just as a graph can be represented in either adjacency list or adjacency matrix form.

One difference is that Sipser considers a general set of states that is not necessarily of the form for some natural number . Sipser also restricts his attention to Turing machines that output only a single bit and therefore designates two special *halting states*: the “ halting state” (often known as the *rejecting state*) and the other as the “ halting state” (often known as the *accepting state*). Thus instead of writing or on an output tape, the machine will enter into one of these states and halt. This again makes no difference to the computational power, though we prefer to consider the more general model of multi-bit outputs. (Sipser presents the basic task of a Turing machine as that of *deciding a language* as opposed to computing a function, but these are equivalent, see decidablelanguagesrem.)

Sipser considers also functions with input in for an arbitrary alphabet (and hence distinguishes between the *input alphabet* which he denotes as and the *tape alphabet* which he denotes as ), while we restrict attention to functions with binary strings as input. Again this is not a major issue, since we can always encode an element of using a binary string of length . Finally (and this is a very minor point) Sipser requires the machine to either move left or right in every step, without the tay operation, though staying in place is very easy to emulate by simply moving right and then back left.

Another definition used in the literature is that a Turing machine *recognizes* a language if for every , and for every , . A language is *recursively enumerable* if there exists a Turing machine that recognizes it, and the set of all recursively enumerable languages is often denoted by . We will not use this terminology in this book.

One of the first programming-language formulations of Turing machines was given by Wang [@Wang1957]. Our formulation of NAND-TM is aimed at making the connection with circuits more direct, with the eventual goal of using it for the Cook-Levin Theorem, as well as results such as and . The website [esolangs.org](https://esolangs.org) features a large variety of esoteric Turing-complete programming languages. One of the most famous of them is [Brainf\*ck](https://esolangs.org/wiki/Brainfuck).

1. A *partial function* from a set to a set is a function that is only defined on a *subset* of , (see functionsec). We can also think of such a function as mapping to where is a special “failure” symbol such that indicates the function is not defined on . [↑](#footnote-ref-52)
2. Most programming languages use arrays of fixed size, while a Turing machine’s tape is unbounded. But of course there is no need to store an infinite number of symbols. If you want, you can think of the tape as a list that starts off just long enough to store the input, but is dynamically grown in size as the Turing machine’s head explores new positions. [↑](#footnote-ref-56)
3. You can use the sequence , ,, , , , ,,, , , , . [↑](#footnote-ref-88)