Turing Machines Part One

What problems can we solve with a computer?

All Languages

That same drawing, to scale.

All Languages

The Problem

- Finite automata accept precisely the regular languages.
- We may need unbounded memory to recognize context-free languages.
	- e.g. $\{a^n b^n \mid n \in \mathbb{N}\}$ requires unbounded counting.
- How do we build an automaton with fnitely many states but unbounded memory?

A Brief History Lesson

A Simple Turing Machine *q*0 $q_{\rm acc}$ $q_{\rm rei}$ *q*1 *start* $a \rightarrow \Box$, R $a \rightarrow \Box$, R $\square \rightarrow \square$, R $\square \rightarrow \square$, R nite state control. It issues q *finite state control.* It issues commands that drive the peration of the machine. This is the Turing machine's *finite state control*. It issues commands that drive the operation of the machine.

A Simple Turing Machine *q*0 $q_{\rm acc}$ $q_{\rm rej}$ *q*1 *start* $a \rightarrow \Box$, R $a \rightarrow \Box$, R $\square \rightarrow \square$, R $\square \rightarrow \square$, R This is the TM's *infinite tape*. Each tape cell holds a *tipe i symbol*. Initially, all (infinitely many) tape symbols are blank. This is the TM's *infinite tape*. Each tape cell holds a *tipe i stymbnt*. Initially, all (infnitely many) tape symbols are blank.

A Simple Turing Machine *q*0 q_{acc} *q*1 *start* $a \rightarrow \Box$, R $\square \rightarrow \square$, R

a → , R ☐ Lach transition has the form Each transition has the form

> *<u><i><u>oud</u>* → write, and</u> *read → write, dir*

grad means "if symbol *read* is under the tape head, replace it with *write* and write the tape head in direction $\frac{d}{dx}$ (Ler D). The \Box symbol denotes a blank sell nove the tape head in direction *dif* (L or R). The □ symbol denotes a blank cell. and means "if symbol *read* is under the tape head, replace it with *write* and move the tape head in direction *dir* (L or R). The □ symbol denotes a blank cell.

$\square \rightarrow \square, R$ A Simple Turing Machine *q*0 q_{acc} *q*1 *start* $a \rightarrow \Box$, R

a → , R ☐ Lach transition has the form Each transition has the form

<u><i><u>oud</u> → write, and</u> *read → write, dir*

grad means "if symbol *read* is under the tape head, replace it with *write* and write the tape head in direction $\frac{d}{dx}$ (Ler D). The \Box symbol denotes a blank sell nove the tape head in direction *dif* (L or R). The □ symbol denotes a blank cell. and means "if symbol *read* is under the tape head, replace it with *write* and move the tape head in direction *dir* (L or R). The □ symbol denotes a blank cell.

$\square \rightarrow \square, R$ A Simple Turing Machine *q*0 q_{acc} *q*1 *start* **a** → □, R

pn has the form

write, dir

he tape head, 1

a → , R ☐ Lach transition has the form Each transition has the form

<u><i><u>oud</u> → write, and</u> *read → write, dir*

and means "if symbol *read* is under the tape head, replace it with *write* and we the tape head in direction $\frac{d}{dx}$ (Ler D). The \Box symbol denotes a blank sell nove the tape head in direction *dif* (L or R). The □ symbol denotes a blank cell. and means "if symbol *read* is under the tape head, replace it with *write* and move the tape head in direction *dir* (L or R). The □ symbol denotes a blank cell.

$\square \rightarrow \square, R$ A Simple Turing Machine *q*0 q_{acc} *q*1 *start* **a** → □, R

pn has the form

write, dir

he tape head, 1

a → , R ☐ Lach transition has the form Each transition has the form

> *<u><i><u>oud</u>* → write, and</u> *read → write, dir*

and means "if symbol *read* is under the tape head, replace it with *write* and we the tape head in direction $\frac{d}{dx}$ (Ler D). The \Box symbol denotes a blank sell nove the tape head in direction *dif* (L or R). The □ symbol denotes a blank cell. and means "if symbol *read* is under the tape head, replace it with *write* and move the tape head in direction *dir* (L or R). The □ symbol denotes a blank cell.

$\square \rightarrow \square, R$ A Simple Turing Machine *q*0 q_{acc} *q*1 *start* $\frac{a \rightarrow \Box, R}{c}$

on has the form

write, dir

he tape head, i

a → , R ☐ Lach transition has the form Each transition has the form

> *<u><i><u>oud</u>* → write, and</u> *read → write, dir*

and means "if symbol *read* is under the tape head, replace it with *write* and we the tape head in direction $\frac{d}{dx}$ (Ler D). The \Box symbol denotes a blank sell nove the tape head in direction *dif* (L or R). The □ symbol denotes a blank cell. and means "if symbol *read* is under the tape head, replace it with *write* and move the tape head in direction *dir* (L or R). The □ symbol denotes a blank cell.

… a a … $\square \rightarrow \square$, R a A Simple Turing Machine *q*0 $q_{\rm acc}$ $q_{\rm rel}$ *q*1 *start* $a \rightarrow \Box$, R $\square \rightarrow \square$, R $a \rightarrow \Box, R$
 $a \rightarrow \Box, R$

$\square \rightarrow \square$, R A Simple Turing Machine *q*0 *q*1 *start* $a \rightarrow \Box$, R $a \rightarrow \Box, R \quad \boxed{\Box \rightarrow \Box, R}$ *q*acc

The Turing Machine

- A Turing machine consists of three parts:
	- A *finite-state control* that issues commands,
	- an *infinite tape* for input and scratch space, and
	- a **tape head** that can read and write a single tape cell.
- At each step, the Turing machine
	- writes a symbol to the tape cell under the tape head,
	- changes state, and
	- moves the tape head to the left or to the right.

Input and Tape Alphabets

- A Turing machine has two alphabets:
	- An *input alphabet* Σ . All input strings are written in the input alphabet.
	- A *tape alphabet* Γ, where $\Sigma \subsetneq \Gamma$. The tape alphabet contains all symbols that can be written onto the tape.
- The tape alphabet Γ can contain any number of symbols, but always contains at least one **blank** *symbol*, denoted \Box . You are quaranteed $\Box \notin \Sigma$.
- At startup, the Turing machine begins with an infinite tape of \square symbols with the input written at some location. The tape head is positioned at the start of the input.

Accepting and Rejecting States

- Unlike DFAs, Turing machines do not stop processing the input when they finish reading it.
- Turing machines decide when (and if!) they will accept or reject their input.
- Turing machines can enter infinite loops and never accept or reject; more on that later...

Determinism

- Turing machines are *deterministic*: for every combination of a (non-accepting, non-rejecting) state *q* and a tape symbol *a* ∈ Γ, there must be exactly one transition defned for that combination of *q* and *a*.
- Any transitions that are missing implicitly go straight to a rejecting state. We'll use this later to simplify our designs.

Determinism

- Turing machines are *deterministic*: for every combination of a (non-accepting, non-rejecting) state *q* and a tape symbol *a* ∈ Γ, there must be exactly one transition defned for that combination of *q* and *a*.
- Any transitions that are missing implicitly go straight to a rejecting state. We'll use this later to simplify our designs.

E. None of these, or two or more of these.

Answer at **PollEv.com/cs103** or text **CS103** to **22333** once to join, then *A*, *B*, *C*, *D*, or *E*. Answer at **PollEv.com/cs103** or text **CS103** to **22333** once to join, then *A*, *B*, *C*, *D*, or *E*.

Answer at **PollEv.com/cs103** or text **CS103** to **22333** once to join, then *A*, *B*, *C*, or *D*. Answer at **PollEv.com/cs103** or text **CS103** to **22333** once to join, then *A*, *B*, *C*, or *D*.

Although the tape ends with **bba** written on it, the original input string was **aab**. This shows that the TM accepts **aab**, not **bba**. Although the tape ends with **bba** written on it, the original input string was **aab**. This shows that the TM accepts **aab**, not **bba**.

> $SO \mathscr{L}(M) = \{ W \in \{a, b\}^n \mid W \text{ ends in } b \}$ So $\mathcal{L}(M) = \{ w \in \{a, b\}^* \mid w \text{ ends in } b \}$

Designing Turing Machines

- Despite their simplicity, Turing machines are very powerful computing devices.
- Today's lecture explores how to design Turing machines for various languages.

Designing Turing Machines

- Let $\Sigma = \{0, 1\}$ and consider the language $L = \{ 0^n 1^n \mid n \in \mathbb{N} \}$.
- We know that *L* is context-free.
- How might we build a Turing machine for it?

A Recursive Approach

- The string ε is in *L*.
- The string **0***w***1** is in *L* if *w* is in *L*.
- Any string starting with **1** is not in *L*.
- Any string ending with **0** is not in *L*.

A Sketch of the TM

A Sketch of the TM

A Sketch of the TM

A Sketch of the TM

A Sketch of the TM

A Sketch of the TM $\overline{\mathbf{0}}$ 1 $\mathbf{1}$

 \mathbf{r} , \mathbf{r} , \mathbf{r}

 \mathbf{r} , \mathbf{r} , \mathbf{r}

Another TM Design

- We've designed a TM for $\{0^n1^n \mid n \in \mathbb{N}\}.$
- Consider this language over $\Sigma = \{0, 1\}$:

 $L = \{ w \in \Sigma^* \mid w \text{ has the same number } \}$ of **0**s and **1**s }

- This language is also not regular, but it is context-free.
- How might we design a TM for it?

How do we know that How do we know that this blank isn't one of this blank isn't one of the infnitely many the infnitely many blanks after our input blanks after our input string?

How do we know that How do we know that this blank isn't one of this blank isn't one of the infnitely many the infnitely many blanks after our input blanks after our input string?

How do we know that How do we know that this blank isn't one of this blank isn't one of the infnitely many the infnitely many blanks after our input blanks after our input string?

… **× 0 × 1 1 1 0 0** …

… **× × × × 1 1 0 0** …

… **× × × × × 1 × 0** …

… **× × × × × × × ×** …

Remember that all Remember that all missing transitions missing transitions implicitly reject. implicitly reject.

Constant Storage

- Sometimes, a TM needs to remember some additional information that can't be put on the tape.
- In this case, you can use similar techniques from DFAs and introduce extra states into the TM's fnite-state control.
- The finite-state control can only remember one of fnitely many things, but that might be all that you need!

Time-Out for Announcements!

Second Midterm Exam

- You're done with the second midterm exam! Woohoo!
- We'll be grading the exam this weekend. Unfortunately, we will not be able to get grades back before Friday.
- Have questions? Feel free to ask in ofice hours or on Piazza!

Problem Set Seven

- Problem Set Seven is due this Friday at 2:30PM.
- As always, if you have questions, feel free to stop by ofice hours or ask on Piazza!

Problem Set Six Scores

Back to CS103!

Another TM Design

• We just designed a TM for this language over Σ = {**0**, **1**}:

$$
L = \{ w \in \Sigma^* \mid w \text{ has the same number} \text{ of } \mathbf{0} \text{s and } \mathbf{1} \text{s } \}
$$

• Let's do a quick review of how it worked.

How do we know that How do we know that this blank isn't one of this blank isn't one of the infnitely many the infnitely many blanks after our input blanks after our input string?

A Diferent Idea

… **0 0 0 1 1 1 1 0** … A Diferent Strategy

Could we sort Could we sort The characters ot | this string? this string?


```
Observation 1: A string 
Observation 1: A string 
 of 0s and 1s is sorted 
of 0s and 1s is sorted 
if it matches the regex 
if it matches the regex 
                   0^*1^*.
```


```
Observation 1: A string 
Observation 1: A string 
 of 0s and 1s is sorted 
of 0s and 1s is sorted 
if it matches the regex 
if it matches the regex 
                   0^*1^*.
```



```
Observation 1: A string 
Observation 1: A string 
 of 0s and 1s is sorted 
of 0s and 1s is sorted 
if it matches the regex 
if it matches the regex 
                   0^*1^*.
```


Observation 1: A string **Observation 1:** A string of 0s and 1s is sorted of 0s and 1s is sorted if it matches the regex if it matches the regex 0^*1^* .

Let's Build It!

 $\boldsymbol{0}$

 ~ 100

 $\overline{0}$

 Ω

 ~ 100

 ~ 100

 ~ 100

TM Subroutines

- A **TM subroutine** is a Turing machine that, instead of accepting or rejecting an input, does some sort of processing job.
- TM subroutines let us compose larger TMs out of smaller TMs, just as you'd write a larger program using lots of smaller helper functions.
- Here, we saw a TM subroutine that sorts a sequence of 0s and 1s into ascending order.

TM Subroutines

- Typically, when a subroutine is done running, you have it enter a state marked "done" with a dashed line around it.
- When we're composing multiple subroutines together – which we'll do in a bit – the idea is that we'll snap in some real state for the "done" state.

What other subroutines can we make?