

Function & DFA

Zeyu Mi

2022-12-9

Adapted From:

《数理逻辑与集合论》11.1,11.2

IALC 1.1,1.5,2.2

Stanford CS103 Turing Machine



Binary Relation: Recap

- **Binary Relation Basics**
- **Equivalence Relation and Partitions**
 - Reflexivity, Symmetry and Transitivity.
- **Compatible Relation and Covers**
 - Reflexivity and Symmetry
- **Partial Order Relation**
 - Reflexivity, Anti-symmetry and Transitivity
- **Closure**

How we investigate computability?

2.1 Discussion on Problems

Part1. Intro & Set theory(I) : Basics & Formal Language.

Part2. Set Theory(II): Axiom system & Cardinality.

Part3. Capture Structures: Binary Relation & Function

2.2 Discussion on Computation

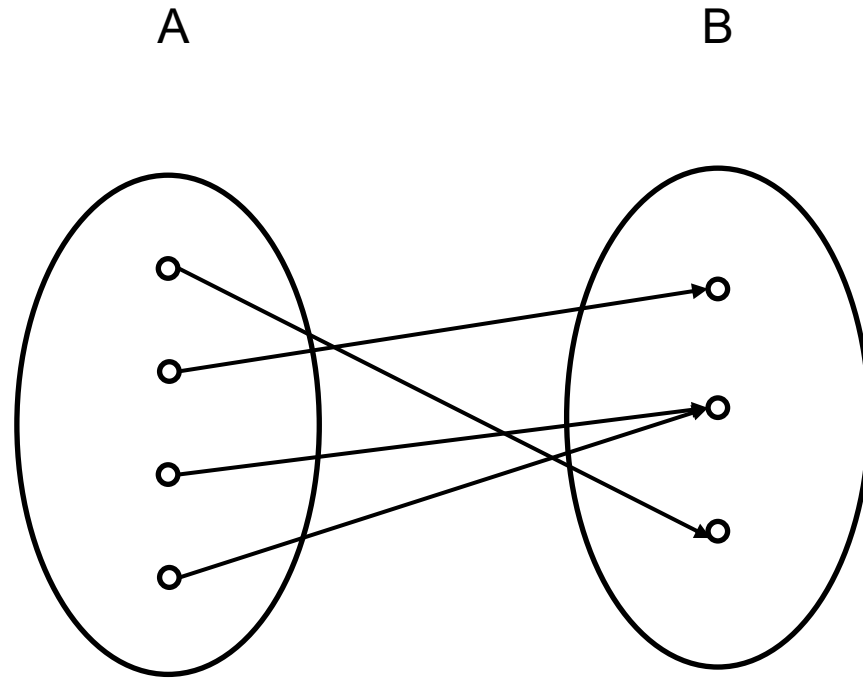
Part1. Turing Machine Basics.

Part2. Variants of Turing Machine. Church-Turing Thesis.

2.3 Discussion on computability

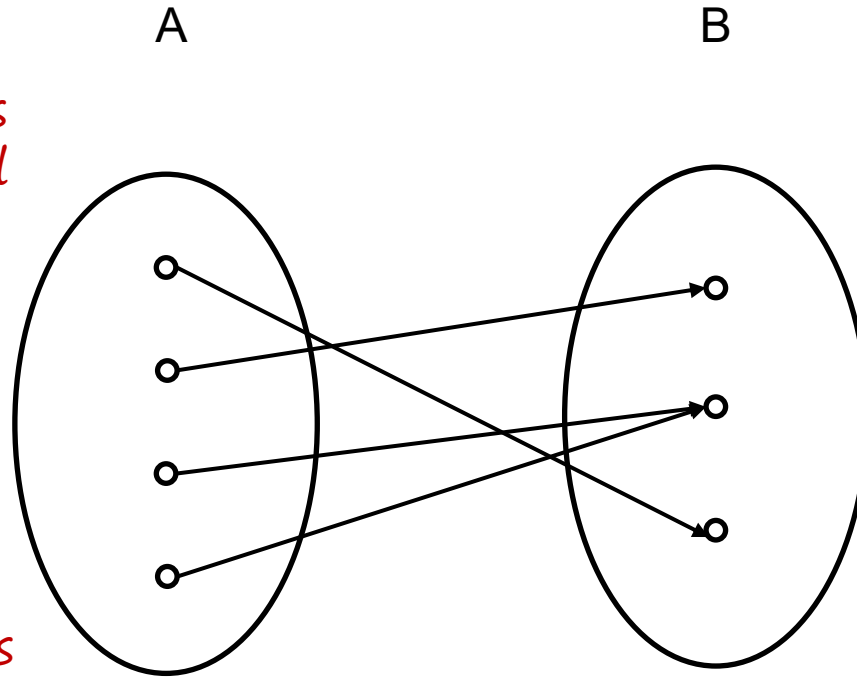
Part1. The Language of Turing Machine. R & RE

Part2. Undecidability.



A special relation from A to B

Every elements
in A are paired



Every elements
in A are paired
only once

A special relation from A to B

Function

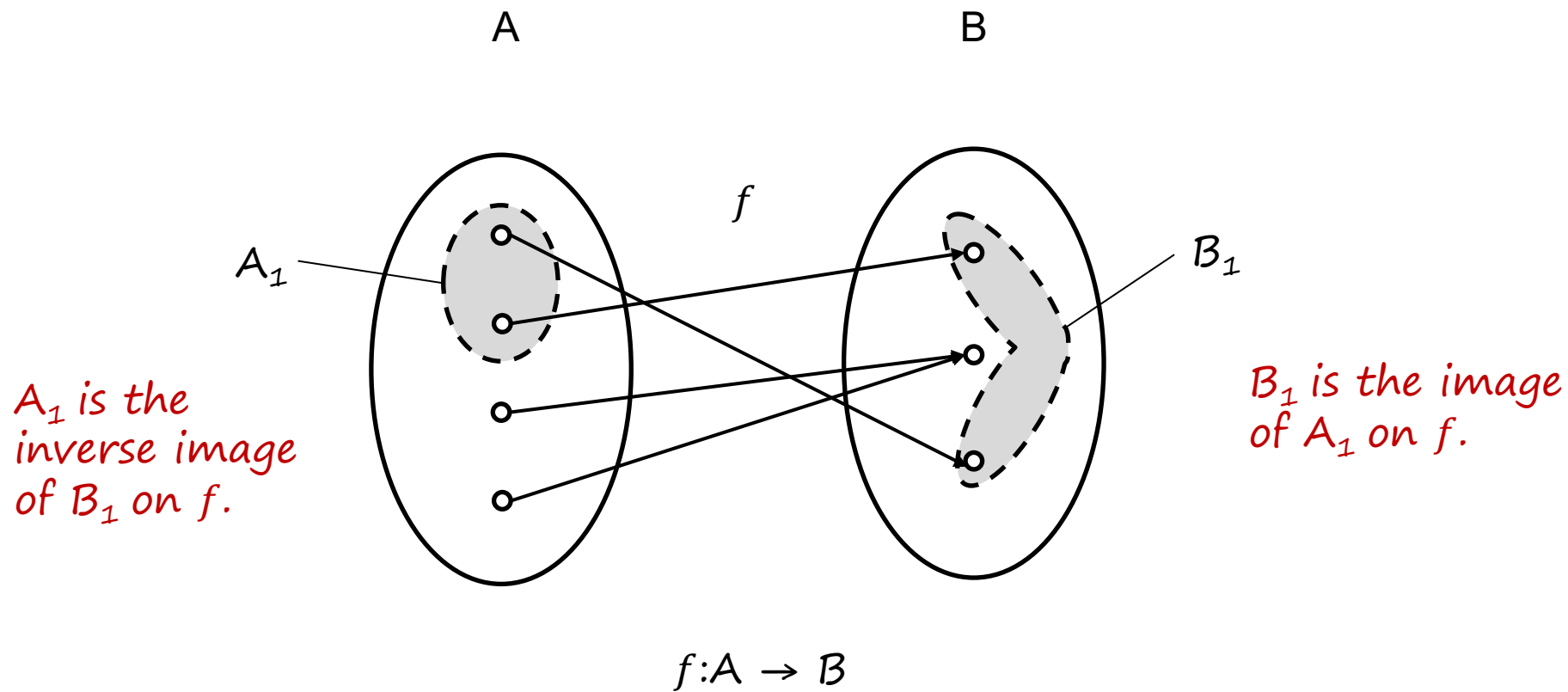
- A relation f from set A to set B is called a **function** if:
 - For any $x \in \text{dom}(f)$, there exists **unique** y such that xfy
$$(\forall x)(\forall y_1)(\forall y_2)((xfy_1 \wedge xfy_2) \rightarrow y_1 = y_2)$$
 - $\text{dom}(f) = A$
$$(\forall x)(x \in A \rightarrow (\exists y)(y \in B \wedge xfy))$$
- A function f from set A to set B is denoted as $f: A \rightarrow B$.
- xfy written as $f: x \mapsto y$ or $f(x) = y$

Function

- All the function from set A to set B is A_B
$$A_B = \{f | f: A \rightarrow B\}$$
- If $|A| = n$, $|B| = m$
 - then $|A_B| = m^n$

Function

- For relation R on A , for a set B
 - the **restriction**(限制) of R on B , denote as $R \upharpoonright B$ is:
$$R \upharpoonright B = \{\langle x, y \rangle \mid \langle x, y \rangle \in R \wedge x \in B\}$$
 - The **image**(象) of R on B , denote as $R[B]$ is:
$$R[B] = \{y \mid \exists x (\langle x, y \rangle \in R \wedge x \in B)\}$$
- Similarly for a function $f: A \rightarrow B$, $\forall A_1 \subseteq A$ we can define $f[A_1]$ to be the **image**(象) of A_1 on function f .
- $\forall B_1 \subseteq B$, $f^{-1}[B_1]$ is the **inverse image or preimage**(原象) of B_1 on f .



Common Functions

- $f: A \rightarrow B$, then f is called **constant function** (常函数) if:
$$(\exists y)(\forall x)(x \in A \rightarrow f(x) = y)$$
- $f: A \rightarrow A$, then f is called **identity function** (恒等函数) if:
$$(\forall x)(x \in A \rightarrow f(x) = x) \text{ or } f = I_A$$
- $f: A^n \rightarrow A$ ($n \in \mathbb{N}$), then f is called a **n-ary operator** (n元运算)
- $f: \mathbb{R} \rightarrow \mathbb{R}$
 - If $(x \leq y) \rightarrow (f(x) \leq f(y))$, we call f is **monotonically increasing** (单调递增)
 - If $(x < y) \rightarrow (f(x) < f(y))$, we call f is **strictly monotonically increasing** (严格单调递增)
 - If $(x \leq y) \rightarrow (f(x) \geq f(y))$, we call f is **monotonically decreasing** (单调递减)
 - If $(x < y) \rightarrow (f(x) > f(y))$, we call f is **strictly monotonically decreasing** (严格单调递减)

Common Functions

- $F: A \rightarrow B_C$ is called a **functional**(noun, 泛函)
 - We define that $F(y) = (f(x) = x + y)$, where $f(x)$ is a certain function.
 - Then $F(1)$ is the function $f(x) = x + 1$
- Let E be the universal set, for any $A \subseteq E$, the **indicator function, or characteristic function**(特征函数) $\chi_A: E \rightarrow \{0, 1\}$ is defined by:

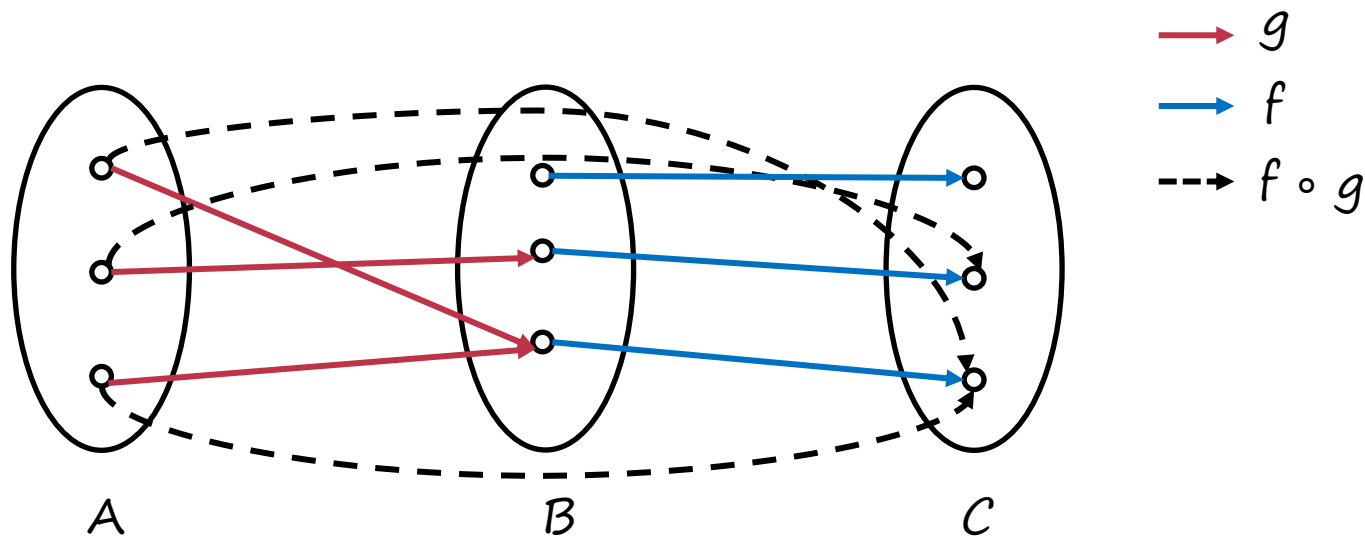
$$\chi_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

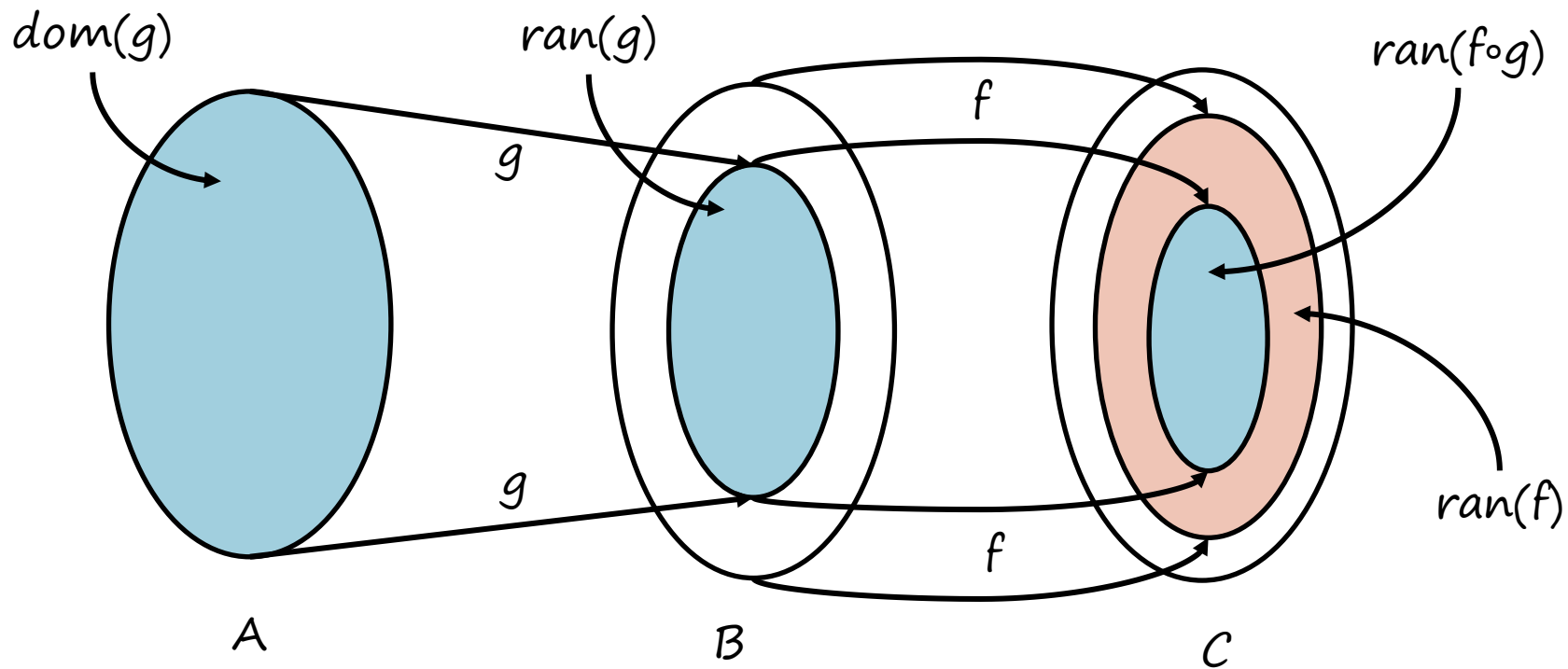
- Let R be an equivalence relation on A , g is a **canonical map or natural map**(典型映射或自然映射) from A to quotient set A/R if:

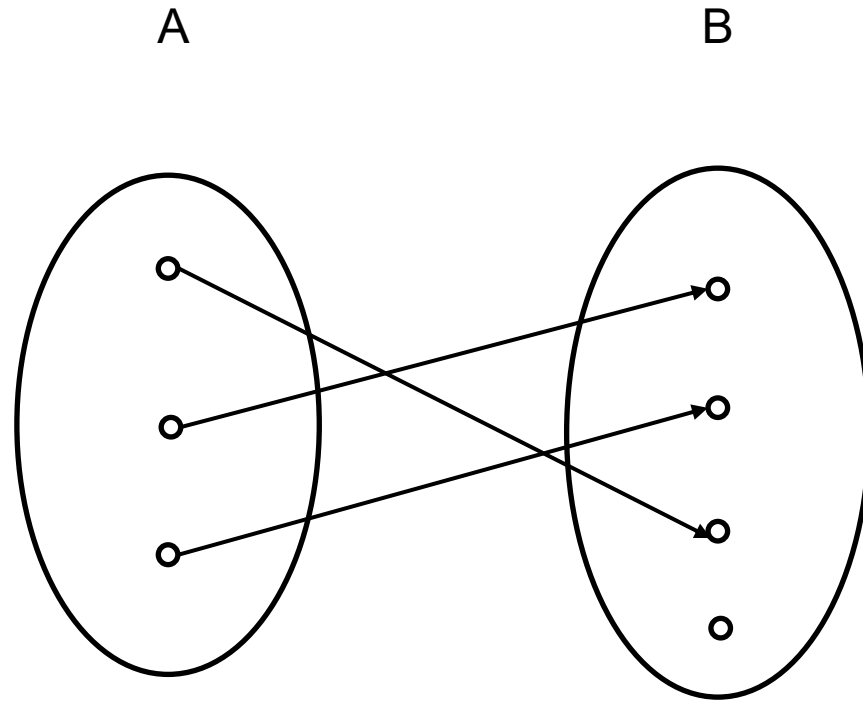
$$g: A \rightarrow A/R, g(x) = [x]_R$$

Compose of Function

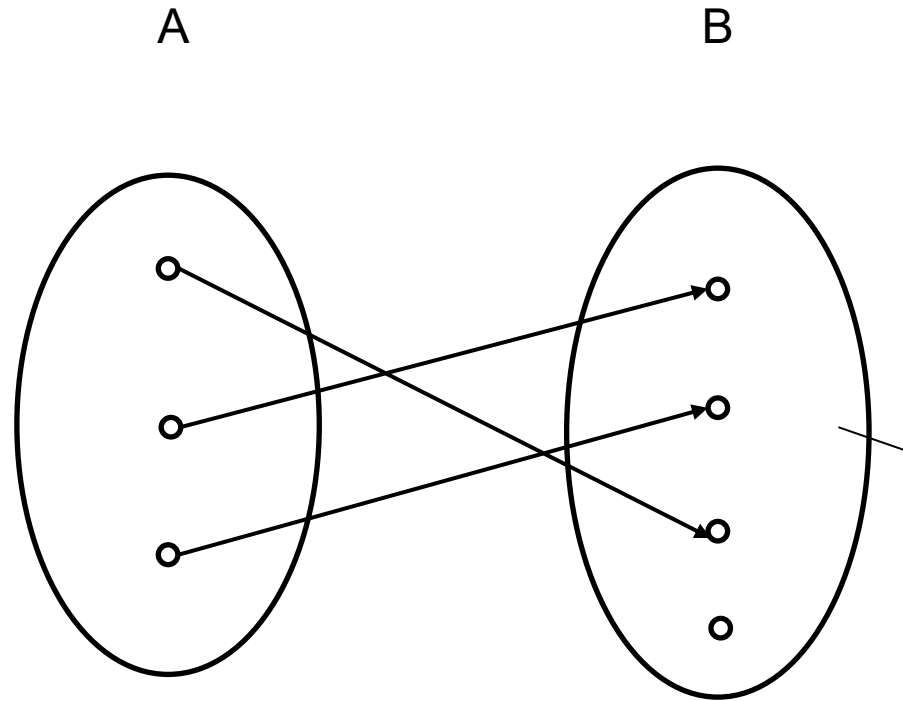
- Let $g: A \rightarrow B$, $f: B \rightarrow C$, then:
 - $f \circ g$ is a function $f \circ g: A \rightarrow C$
 - For any $x \in A$, $(f \circ g)(x) = f(g(x))$







A special function



No elements in B
are paired twice

A special function

Injection(单射)

- $f: A \rightarrow B$ is an **injection** or **injective** if for any $x_1, x_2 \in A, x_1 \neq x_2$, then $f(x_1) \neq f(x_2)$.

$$(\forall x_1)(\forall x_2)(x_1 \in A \wedge x_2 \in A \wedge x_1 \neq x_2 \rightarrow f(x_1) \neq f(x_2))$$

Different inputs has different outputs

- Or if $f(x_1) = f(x_2)$, then $x_1 = x_2$

$$(\forall x_1)(\forall x_2)(x_1 \in A \wedge x_2 \in A \wedge f(x_1) = f(x_2) \rightarrow x_1 = x_2)$$

Same outputs have the same input

Injection

- $f(x) = x \ (\mathbb{R} \rightarrow \mathbb{R})$
 - injective
 - For any $x_1 \neq x_2$, $f(x_1) = x_1 \neq x_2 = f(x_2)$
- $f(x) = x^2 \ (\mathbb{R} \rightarrow \mathbb{R})$
 - not injective.
 - $f(-1) = f(1) = 1$

Injection and Composition

Let $g: A \rightarrow B$, $f: B \rightarrow C$, if f , g are injective, then $f \circ g$ is also injective.

Proof:

For any $x_1, x_2 \in A$, $x_1 \neq x_2$, we have $g(x_1) \neq g(x_2)$ since g is injective.

Now $g(x_1), g(x_2) \in B$, and $g(x_1) \neq g(x_2)$, then we have $f(g(x_1)) \neq f(g(x_2))$ since f is injective.

$(f \circ g)(x) = f(g(x))$ by definition, so $(f \circ g)(x_1) \neq (f \circ g)(x_2)$

Q.E.D

Injection and Composition

Let $g: A \rightarrow B$, $f: B \rightarrow C$, if $f \circ g$ is injective, then g is also injective.

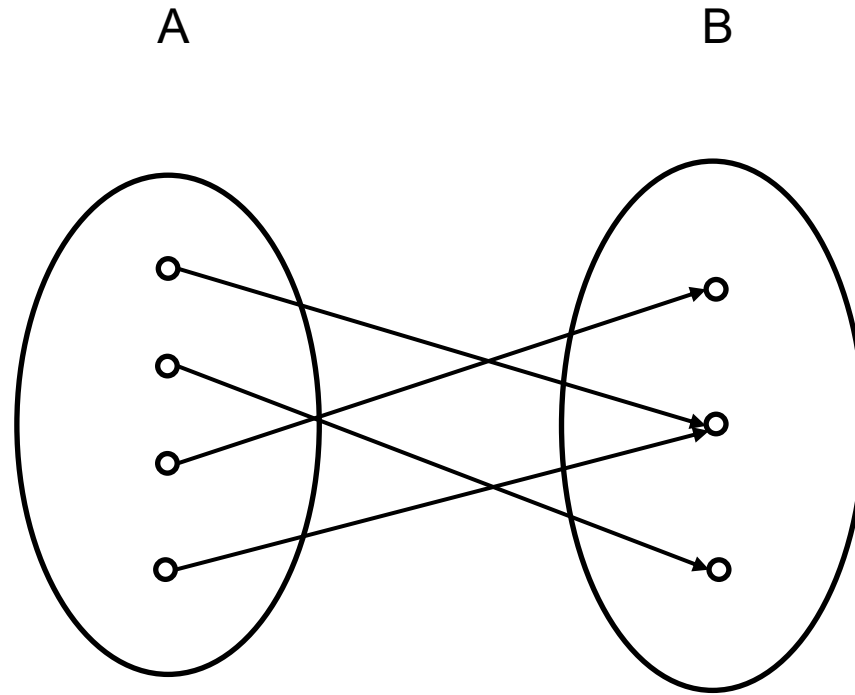
Proof:

Assume that g is not injective, then there exists $x_1, x_2 \in A$, $x_1 \neq x_2$ and $g(x_1) = g(x_2)$.

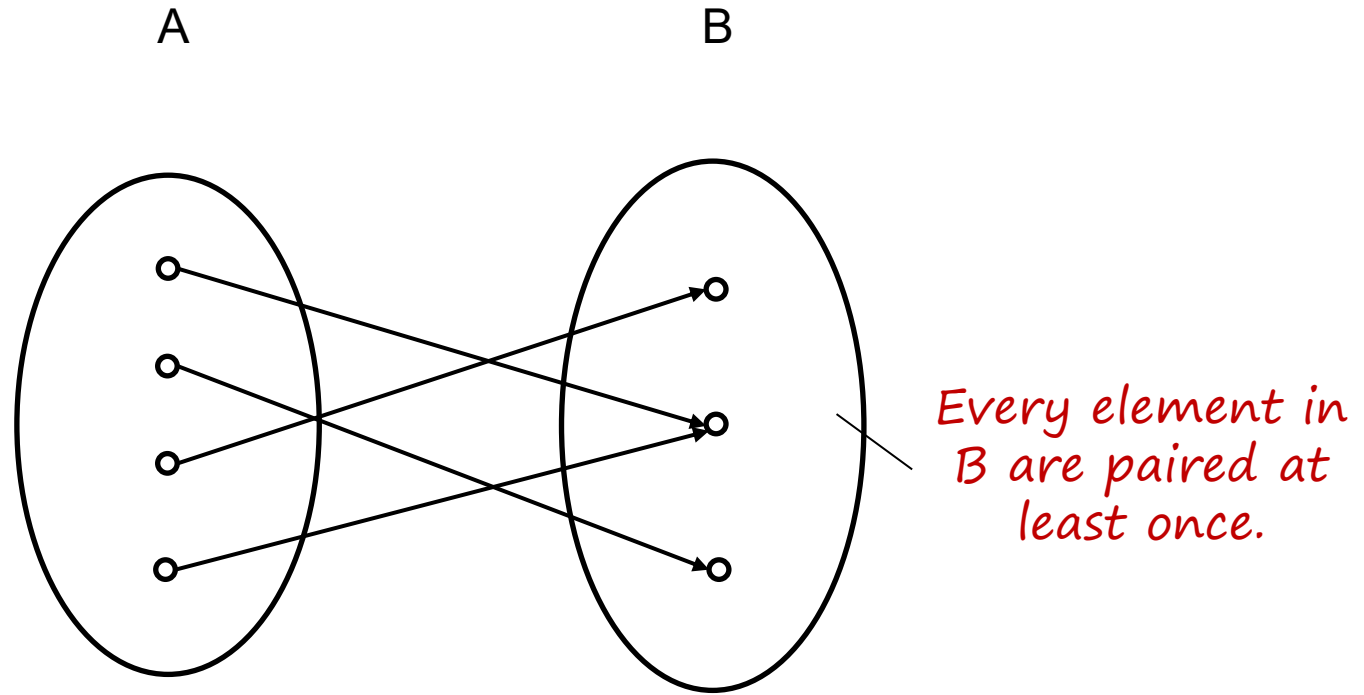
Now $g(x_1), g(x_2) \in B$, and $g(x_1) = g(x_2)$, then we have $f(g(x_1)) = f(g(x_2))$.

But $f \circ g$ is injective, it means $f(g(x_1)) \neq f(g(x_2))$. Contradiction

Q.E.D



A special function



A special function

Surjection(满射)

- $f: A \rightarrow B$ is a **surjection** or **surjective** if $\text{ran}(f) = B$, or for any $y \in B$, there exists $x \in A$ such that $f(x) = y$

$$(\forall y)(\exists x)(y \in B \rightarrow (x \in A \wedge f(x) = y))$$

Every possible output has an input

Surjection

- $f(x) = x + 5$ ($\mathbb{R} \rightarrow \mathbb{R}$)
 - surjective.
 - For any $x_0 \in \mathbb{R}$, $f(x_0 - 5) = x_0$
- $f(x) = x^2$ ($\mathbb{R} \rightarrow \mathbb{R}$)
 - not surjective.
 - There's no element $x_0 \in \mathbb{R}$ such that $f(x_0) = x_0^2 = -1$

Surjection and Composition

Let $g: A \rightarrow B$, $f: B \rightarrow C$, if f, g are surjective, then $f \circ g$ is also surjective.

Proof:

For any $y_1 \in C$, we have $x_1 \in B$ such that $f(x_1) = y_1$ since f is surjective.

Now $x_1 \in B$, so we have $x_0 \in A$ such that $g(x_0) = x_1$ since g is surjective.

So for any $y_1 \in C$, we have $x_0 \in A$ such that $f(g(x_0)) = y_1$, or $(f \circ g)(x_0) = y_1$, it means that $f \circ g$ is surjective

Q.E.D

Surjection and Composition

Let $g: A \rightarrow B$, $f: B \rightarrow C$, if $f \circ g$ is surjective, then f is also surjective.

Proof:

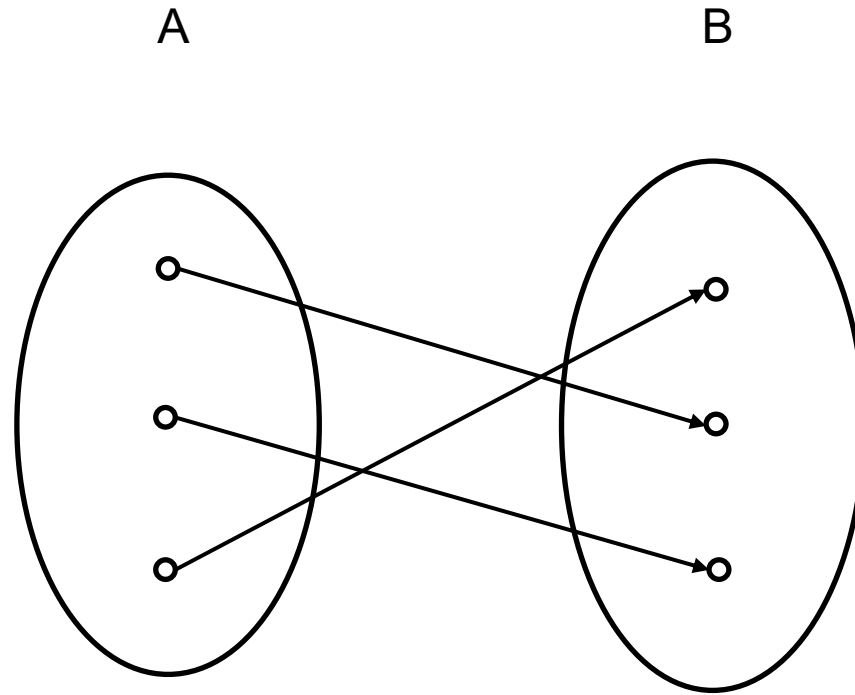
For any $y \in C$, we have $x \in A$ such that $(f \circ g)(x) = y$ as $f \circ g$ is surjective.

Let $x_0 = g(x) \in B$, so we have $(f \circ g)(x) = f(g(x)) = f(x_0) = y$.

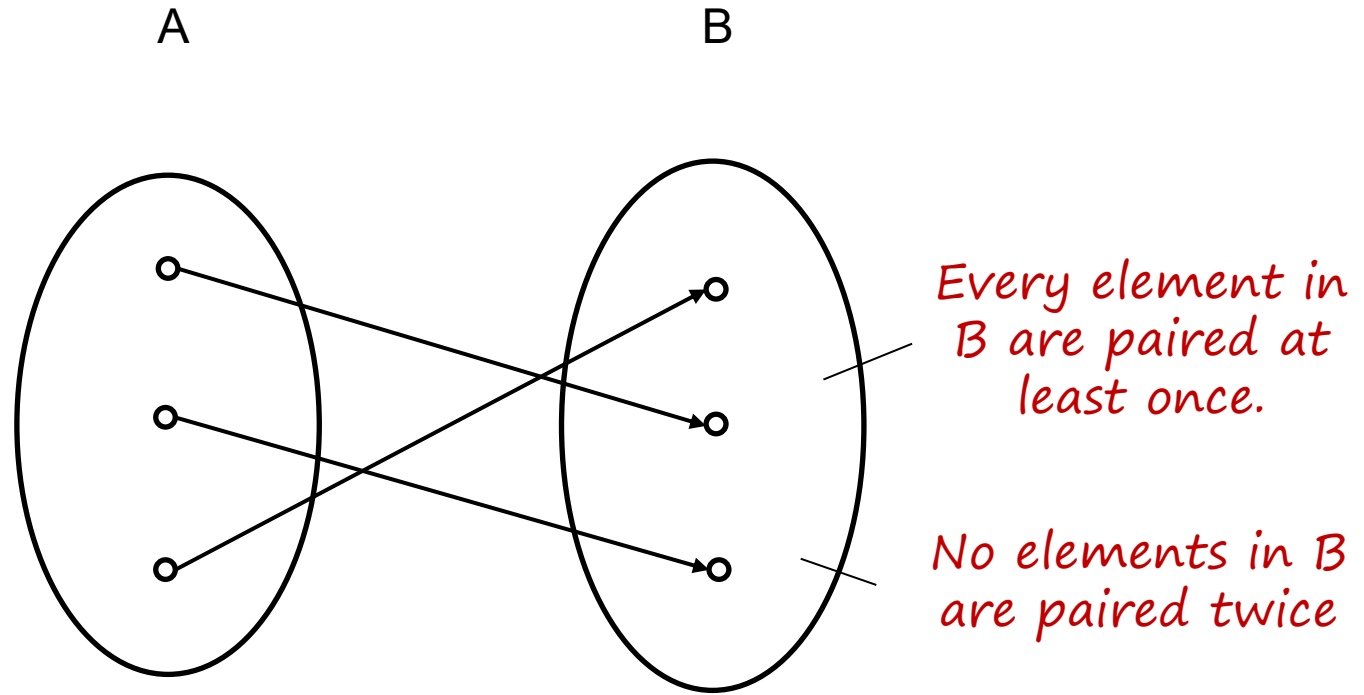
So for any $y \in C$, we have $x_0 \in B$ such that $f(x_0) = y$.

f is surjective.

Q.E.D



A special function



A special function

Bijection(双射)

- $f: A \rightarrow B$ is a **bijection** or **bijjective** if f is **both** injective and surjective.
- Bijections are sometimes called **one-to-one correspondences(一一对应)**.

Bijection

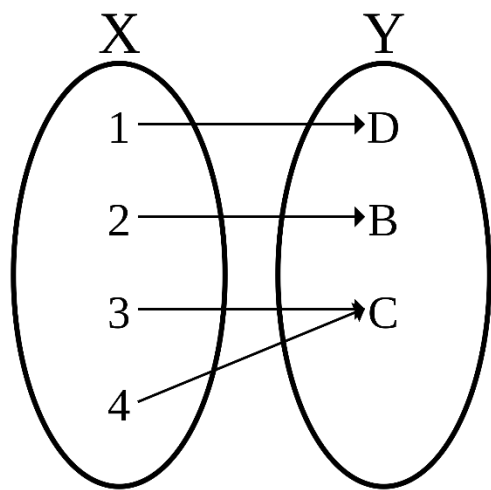
- $f(x) = x + 5$ ($\mathbb{R} \rightarrow \mathbb{R}$) is bijective.
- $f(x) = e^x$ ($\mathbb{R} \rightarrow \mathbb{R}$) is not bijective since it's not surjective.
- $f(x) = x^3 - x$ ($\mathbb{R} \rightarrow \mathbb{R}$) is not bijective since it's not injective.

Bijection and Composition

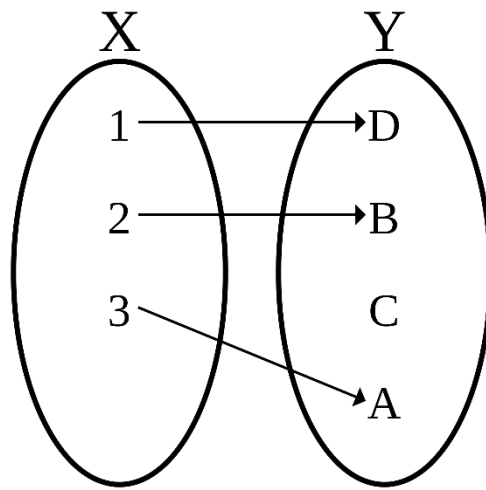
Let $g: A \rightarrow B$, $f: B \rightarrow C$, if f, g is bijective, then $f \circ g$ is also bijective.

Let $g: A \rightarrow B$, $f: B \rightarrow C$, if $f \circ g$ is bijective, then g is injective and f is surjective.

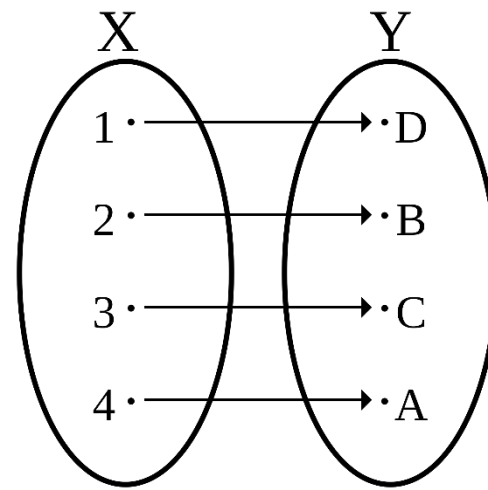
Bijection, injection and surjection



Surjection



Injection



Bijection

Comparing Cardinalities

For set A and B , A and B are **equinumerous** (等势) if there is a way to pair their elements off without leaving any elements uncovered, marked as $A \approx B$. Otherwise marked as $\neg A \approx B$.

Comparing Cardinalities

~~For set A and B , A and B are **equianumerous** (等势) if there is a way to pair their elements off without leaving any elements uncovered, marked as $A \approx B$. Otherwise marked as $\neg A \approx B$.~~

For set A and B , A and B are **equianumerous** (等势) if there is a **bijection from A to B** , marked as $A \approx B$. Otherwise marked as $\neg A \approx B$.

Comparing Cardinalities

For set A and B , $\text{card}(A) \leq \text{card}(B)$ if there is a way to pair their elements off without leaving **any elements of A** uncovered.

Comparing Cardinalities

~~For set A and B , $\text{card}(A) \leq \text{card}(B)$ if there is a way to pair their elements off without leaving **any elements of A** uncovered.~~

For set A and B , $\text{card}(A) \leq \text{card}(B)$ if there **is an injection from A to B .**

Comparing Cardinalities

- $f: A \rightarrow B$
 - If f is injective, then $|A| \leq |B|$
 - If f is surjective, then $|A| \geq |B|$
 - If f is bijective, then $|A| = |B|$

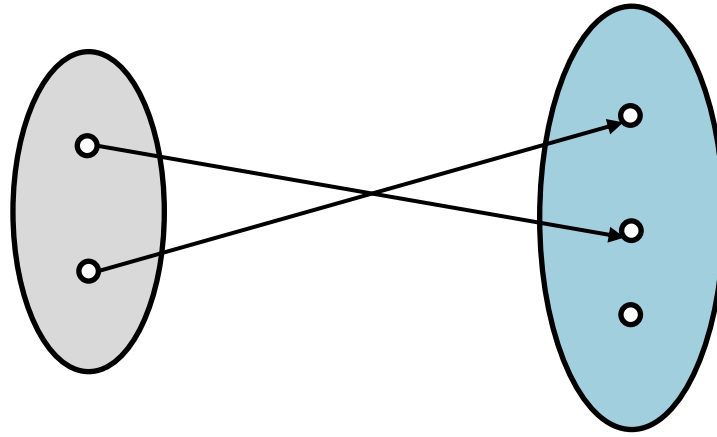
Axiom of Choice

ZF.10 Axiom of choice(选择公理)

For any relation R , there exists a function F , such that F is a subset of R and the domain of F and R are equal.

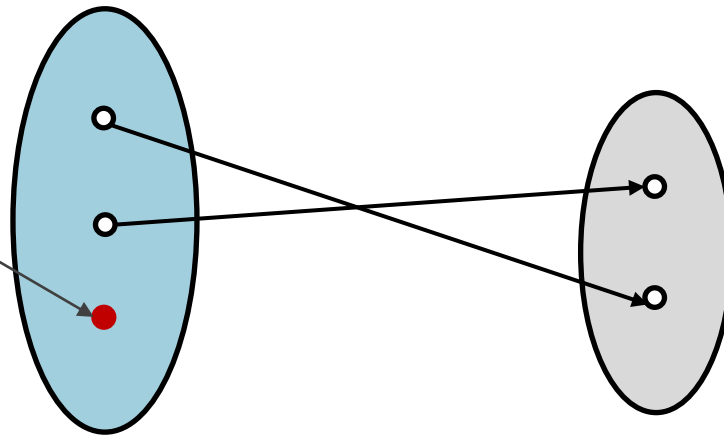
$$(\forall \text{ Relation } R)(\exists \text{ Function } F)(F \subseteq R \wedge \text{dom}(R) = \text{dom}(F))$$

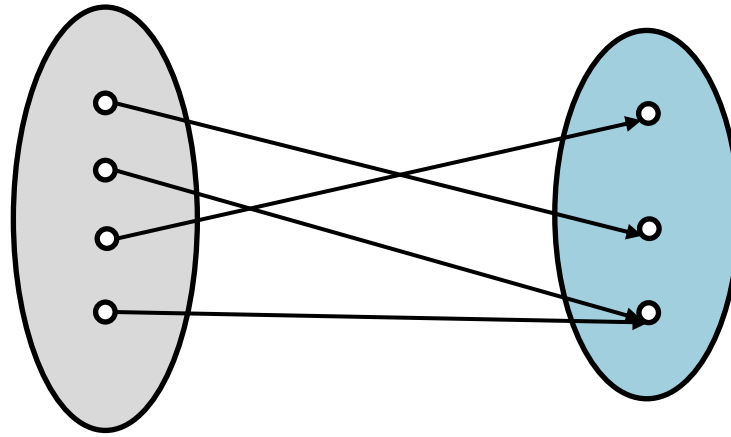
- $R = \{\langle 1, a \rangle, \langle 1, b \rangle, \langle 2, a \rangle\}$
- $f_1 = \{\langle 1, a \rangle, \langle 2, a \rangle\}$
- $f_2 = \{\langle 1, b \rangle, \langle 2, a \rangle\}$



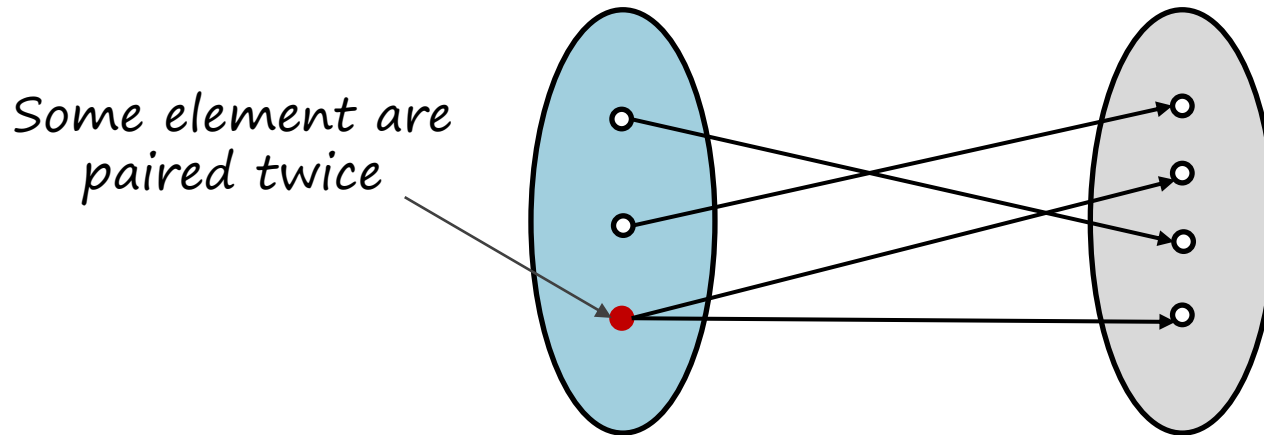
↓ Inverse

Some element are
not paired





↓ Inverse



Inverse of Function

- Let $f: A \rightarrow B$
 - The inverse (defined by relation) of f may not be a function
 - If the inverse of f is a function, we call f is **invertible** (可逆的)

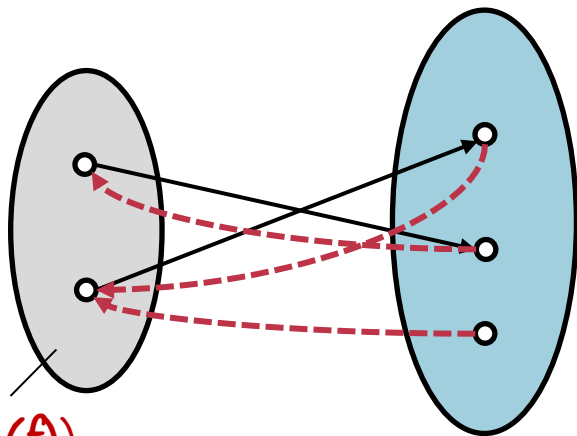
$f: A \rightarrow B$, if there exists $g: B \rightarrow A$, such that $g \circ f = I_A$ and $f \circ g = I_B$, we call that f is **invertible**, and g is the **inverse function**(反函数) of f .

Some times $g \circ f = I_A$ and $f \circ g = I_B$ may not hold simultaneously.

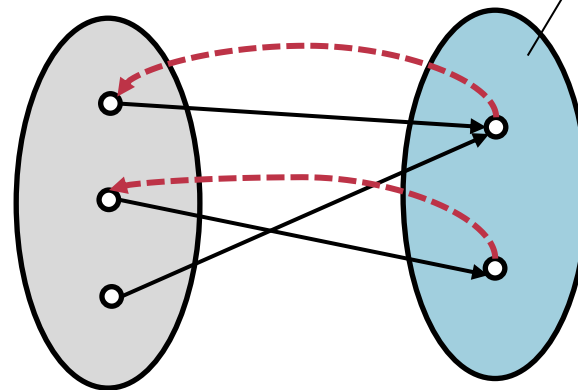
Left Inverse(左逆) and Right Inverse(右逆)

- Let $f: A \rightarrow B$, if there exists $g: B \rightarrow A$, such that
 - $g \circ f = I_A$, then we call g is the **left inverse** of f .
 - $f \circ g = I_B$, then we call g is the **right inverse** of f .

f $\longrightarrow$
 g $\dashrightarrow$



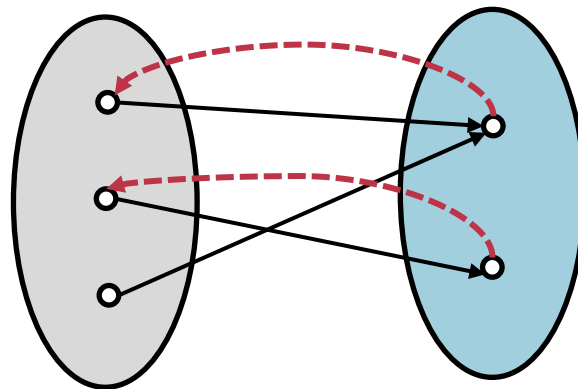
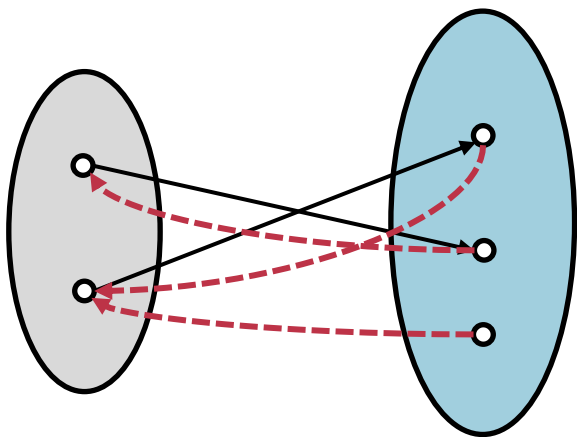
Left Inverse
 $g \circ f = I_A$



Right Inverse
 $f \circ g = I_B$

Left Inverse and Right Inverse

- Let $f: A \rightarrow B$, $A \neq \emptyset$
 - f has left inverse if and only if f is injective.
 - f has right inverse if and only if f is surjective.
 - f has both left and right inverse if and only if f is bijective.
 - f is bijective, then the left and right inverse of f are equal.



Left Inverse and Right Inverse

- Let $f: A \rightarrow B$, $A \neq \emptyset$, assume that g is the left inverse of f , and h is the right inverse of f .
 - We have $g \circ f = I_A$, $f \circ h = I_B$
 - Then $g = g \circ I_B = g \circ (f \circ h) = (g \circ f) \circ h = I_A \circ h = h$
 - So if f has both left and right inverse, then they are equal.

Inverse of Function

- Let $f: A \rightarrow B$
 - The inverse (defined by relation) of f may not be a function
 - If the inverse of f is a function, we call f is **invertible** (可逆的)

$f: A \rightarrow B$, if there exists $g: B \rightarrow A$, such that $g \circ f = I_A$ and $f \circ g = I_B$, we call that f is **invertible**, and g is the **inverse function**(反函数) of f .

- f has both left and right inverse if and only if f is bijective.



f is invertible if and only if f is bijective.

Inverse of Function

- Let $f: A \rightarrow B$ be a bijection
 - f^{-1} is a function and is the inverse function of f .
 - f^{-1} is bijective too.
 - $\forall x \in A, f^{-1}(f(x)) = x. \forall y \in B, f(f^{-1}(y)) = y$

How we investigate computability?

2.1 Discussion on Problems

Part1. Intro & Set theory(I) : Basics & Formal Language.

Part2. Set Theory(II): Axiom system & Cardinality.

Part3. Capture Structures: Binary Relation & Function

2.2 Discussion on Computation

Part1. Turing Machine Basics.

Part2. Variants of Turing Machine. Church-Turing Thesis.

2.3 Discussion on computability

Part1. The Language of Turing Machine. R & RE

Part2. Undecidability.

Ultimate Problem:

What problems can a computer solve?

What We Have Done

- **Computational Problem(计算问题)**
 - Well-defined input
 - Constraints that the output must satisfy.
- **A computational problem is computable(可计算) if there is a procedure (or algorithm) for it which:**
 - Rigorously follows some instructions, requires no ingenuity.
 - Always finishes (terminates) after a finite number of steps.

What We Have Done

- A computational problem can be converted into language recognizing problem.
 - Set basics & formal language theory
 - A language is naturally a set of string.
 - Given a language L and take w as input, determine whether $w \in L$.

~~What problems can a computer solve?~~

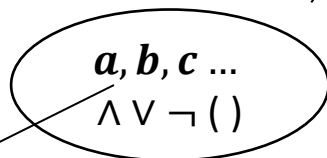


What languages can a computer recognize?

Recap: Formal Languages

*Alphabets are finite, non-empty
sets of characters*

Alphabet: Σ



finite sequence

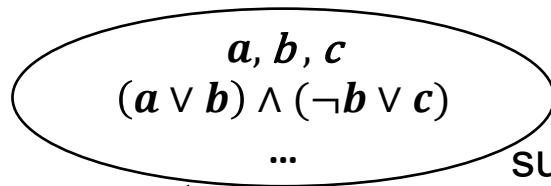
A string

$(a \vee b) \wedge (\neg b \vee c)$

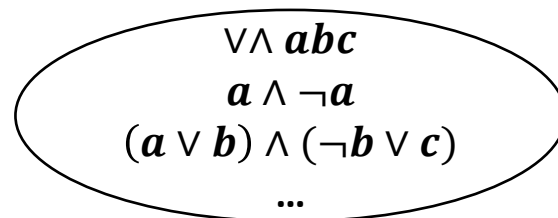
*Strings are
finite sequence
of characters*

*Characters are
individual
symbols*

All strings



subset of Σ^*



Set of all string :
 Σ^*

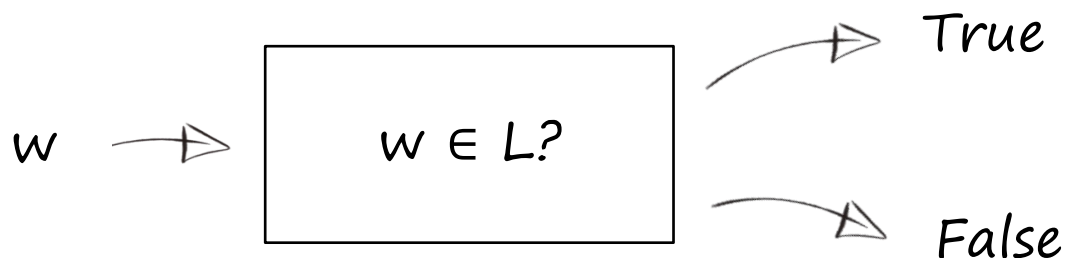
*Languages
are sets of
strings.*

A language

Recap: Language Recognizing

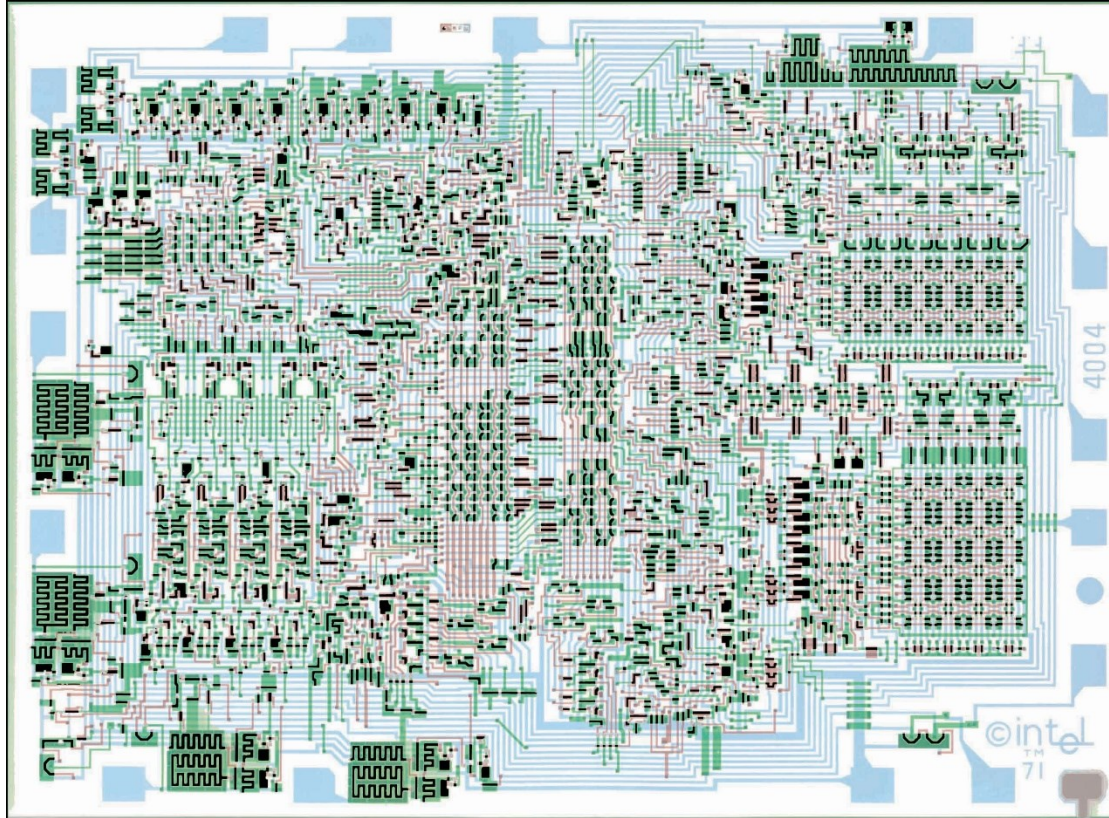
Throughout the whole course of computability, we'll focus on the membership problem of a language, or language recognizing, that is:

Given a language L and take w as input, determine whether $w \in L$.



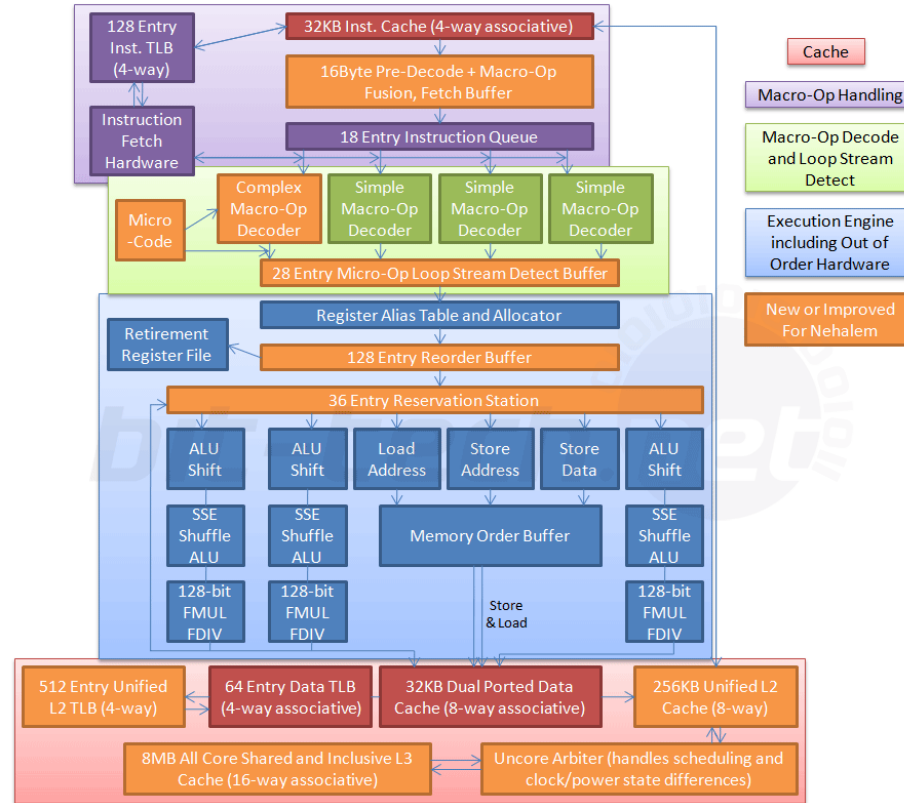
What languages can a *computer* recognize?

Computers are complicated.



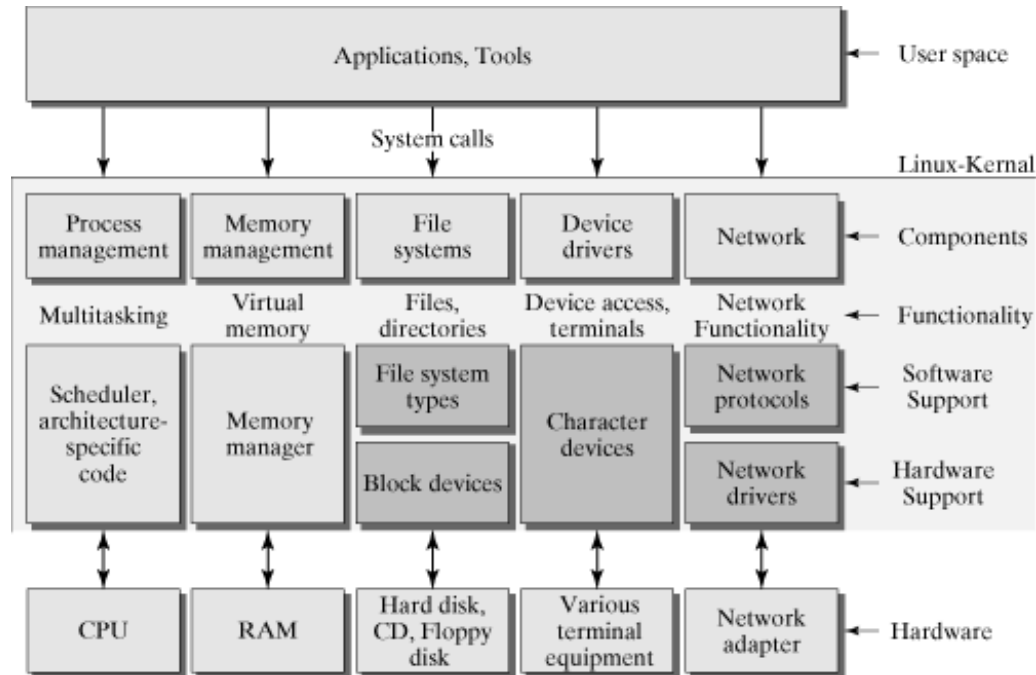
Intel 4004, first commercial CPU

Computers are complicated.



Core i7 Architecture.

Computers are complicated.



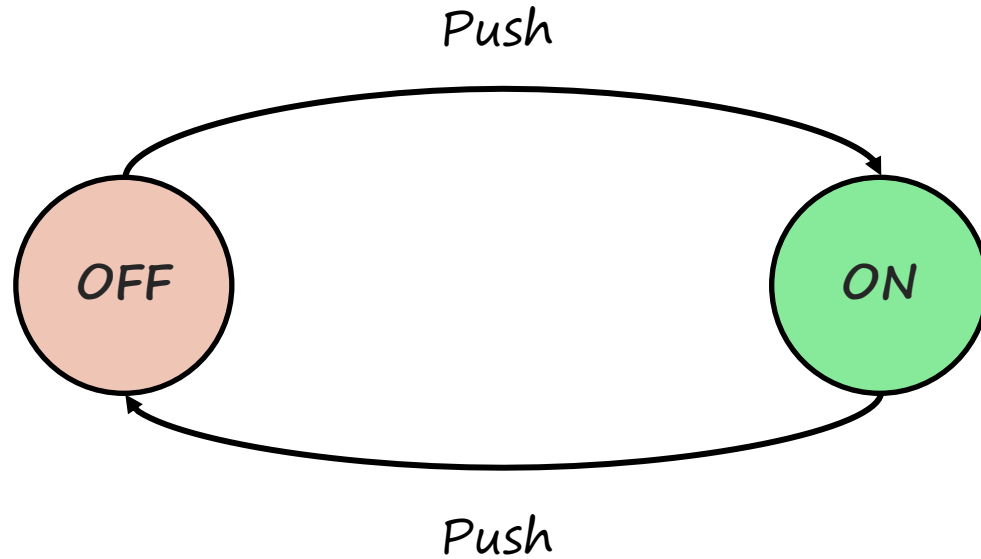
Architecture of Operating System

*We need to model computer in a simple
and clean manner!!*

Automata(自动机) Theory

- An automaton is a mathematical model of a computing device
 - It is significantly easier to manipulate our abstract models of computers than it is to manipulate actual computers
 - If we pick our models correctly, we can make broad, sweeping claims about huge classes of real computers by arguing that they're just special cases of our more general models.

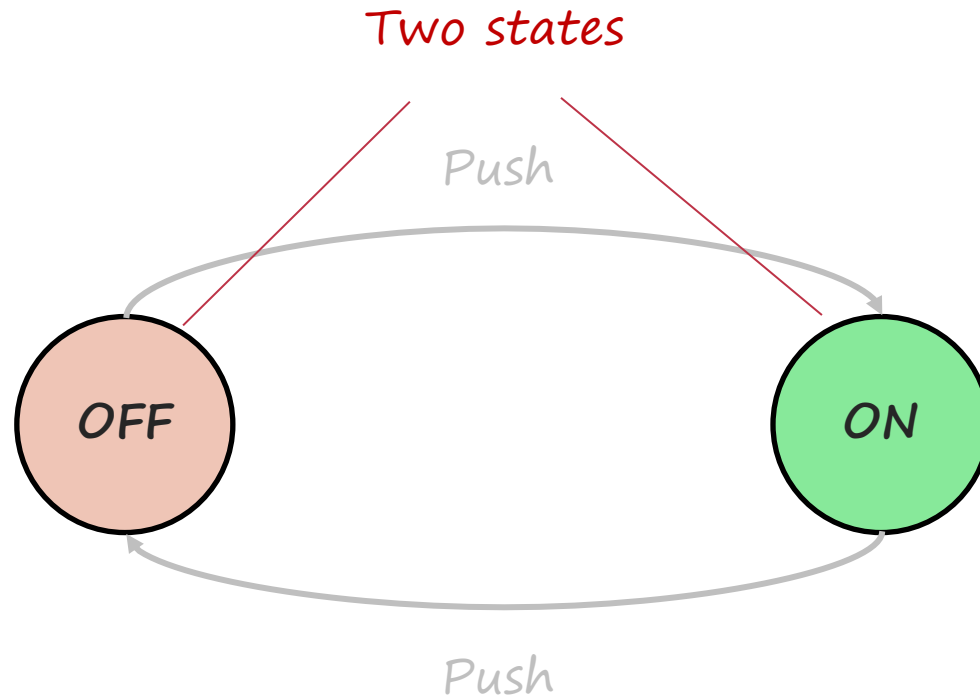
What is an Automaton



An “automata” for an on-off switch

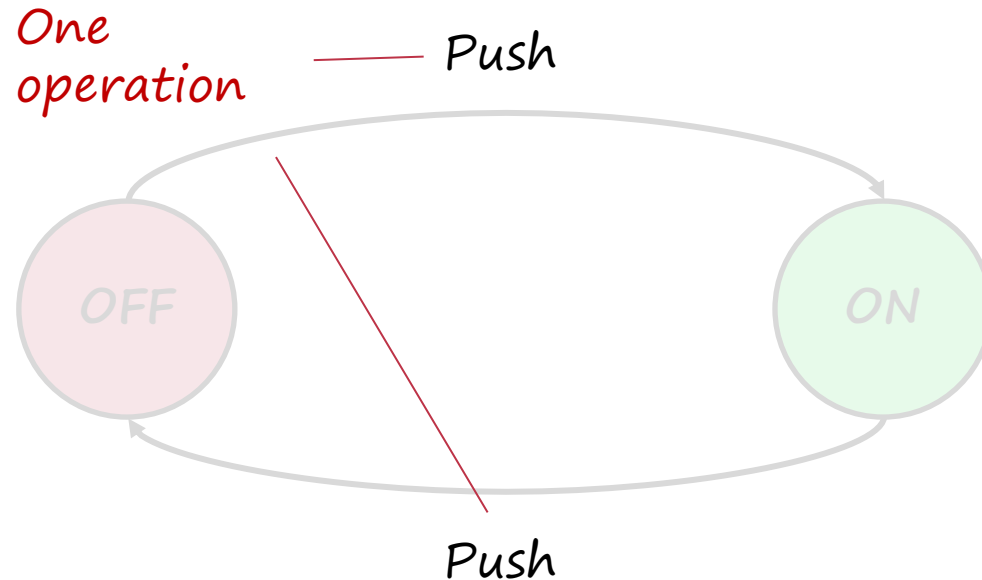
What is an Automaton

- $Q: \{\text{"off"}, \text{"on"}\}$



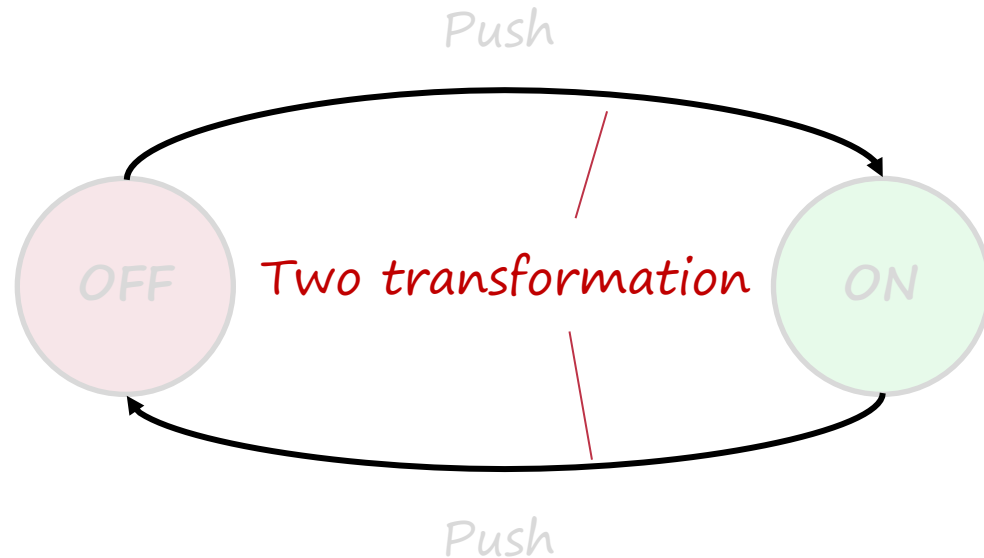
What is an Automaton

- $Q: \{\text{"off"}, \text{"on"}\}$
- $\Sigma : \{\text{"push"}\}$

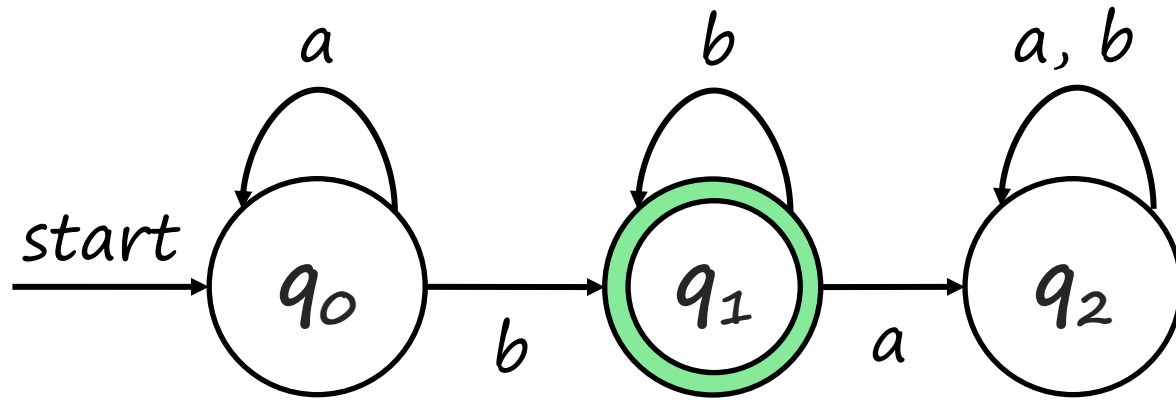


What is an Automaton

- $Q: \{\text{"off"}, \text{"on"}\}$
- $\Sigma : \{\text{"push"}\}$
- $\delta: \{(\text{"on"}, \text{"push"}, \text{"off"}), (\text{"off"}, \text{"push"}, \text{"on"})\}$

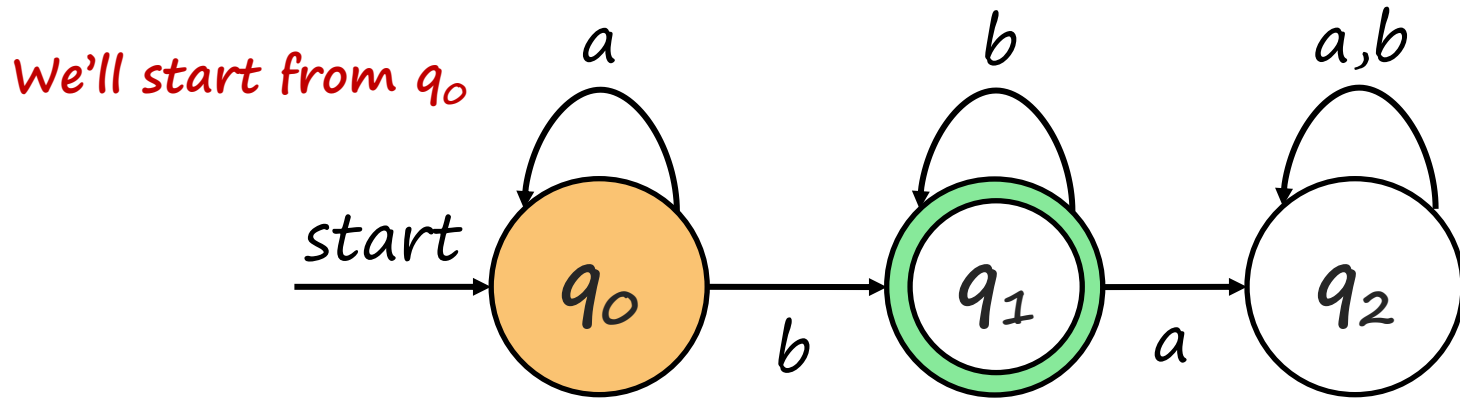


What is an Automaton



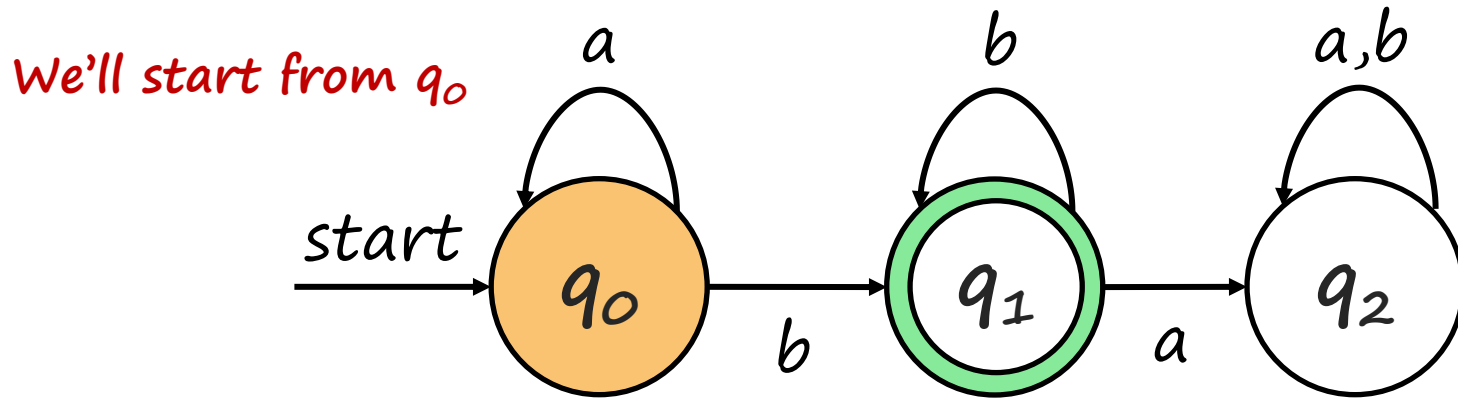
A automata for “aa...bb...”

What is an Automaton



Input: a a b b

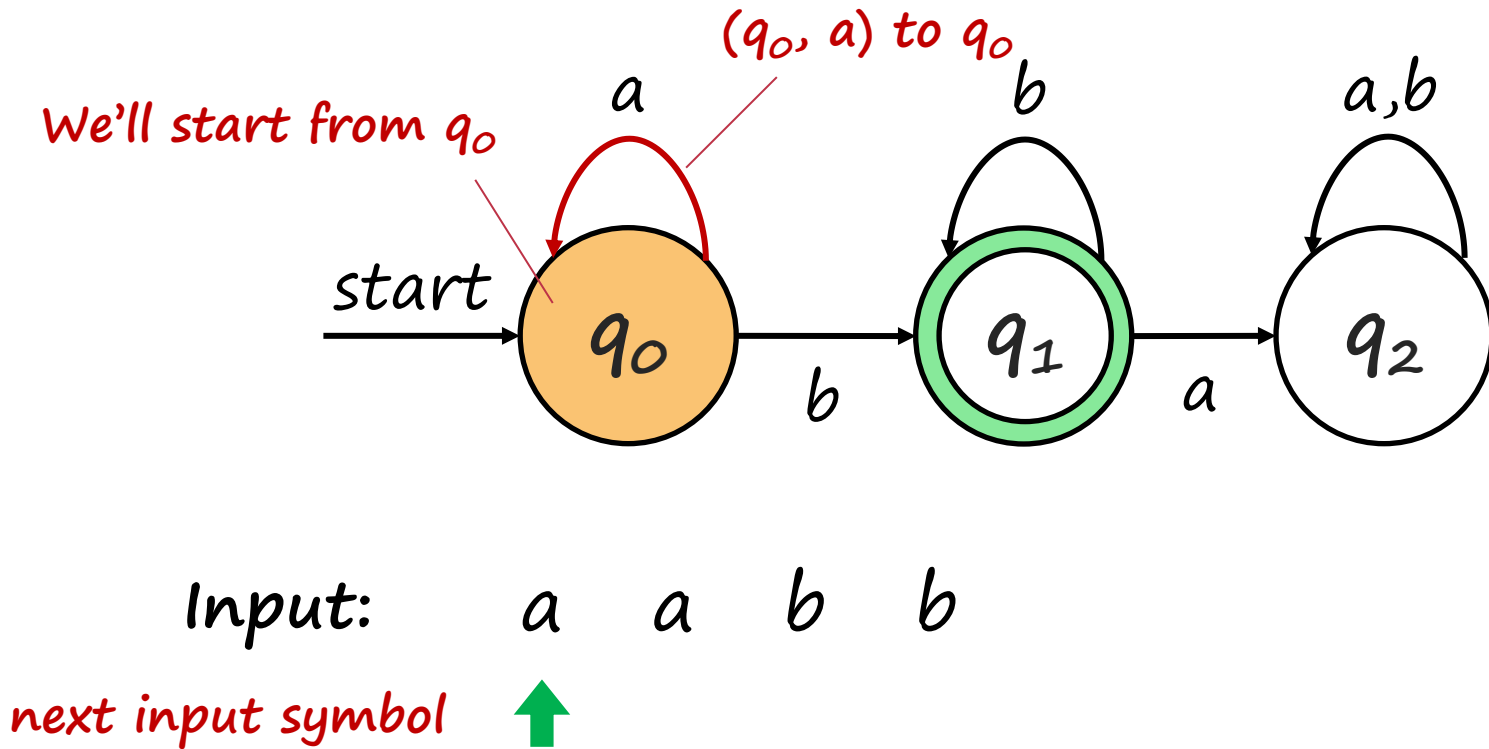
What is an Automaton



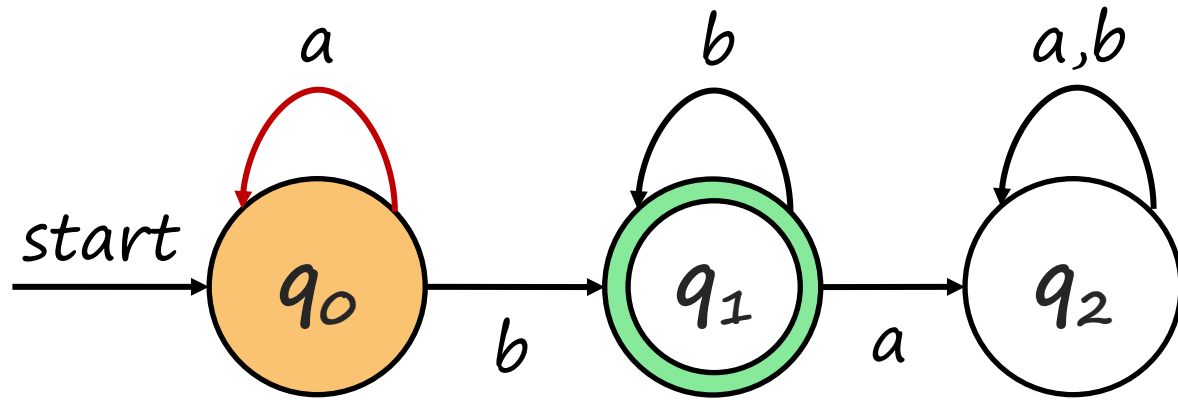
Input: a a b b

next input symbol ↑

What is an Automaton



What is an Automaton

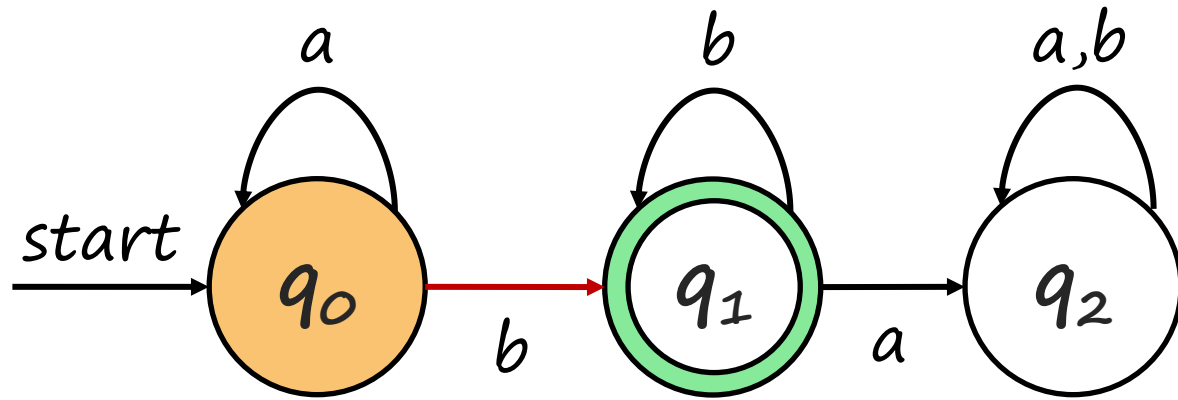


Input: a a b b

next input symbol



What is an Automaton

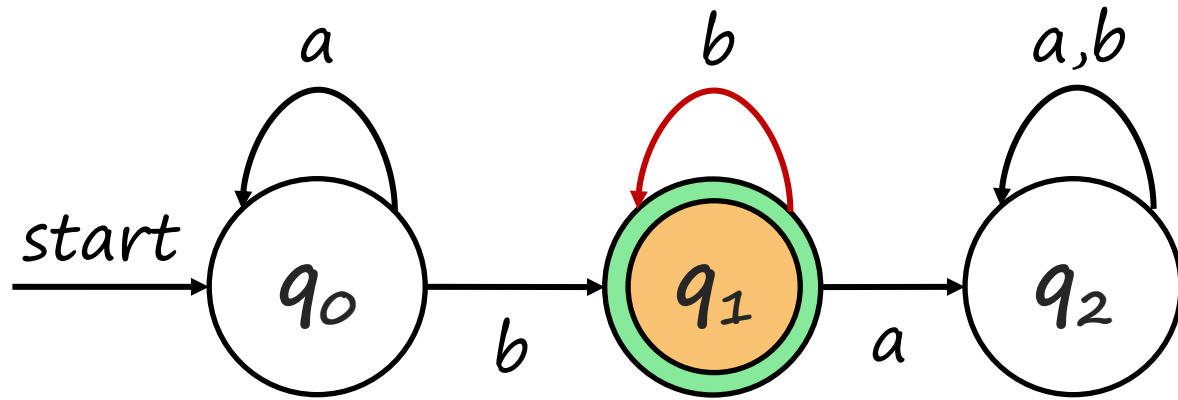


Input: a a b b

next input symbol



What is an Automaton



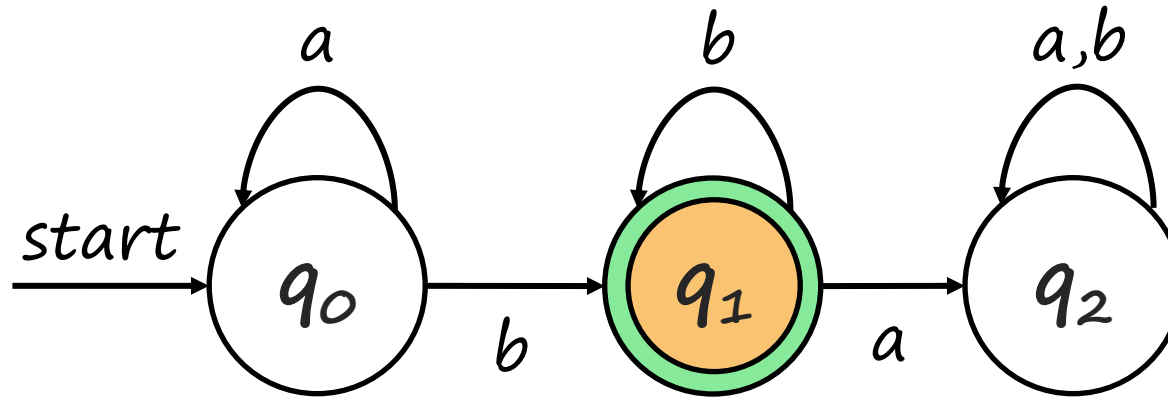
Input: a a b b

next input symbol



What is an Automaton

q_1 is the final state.

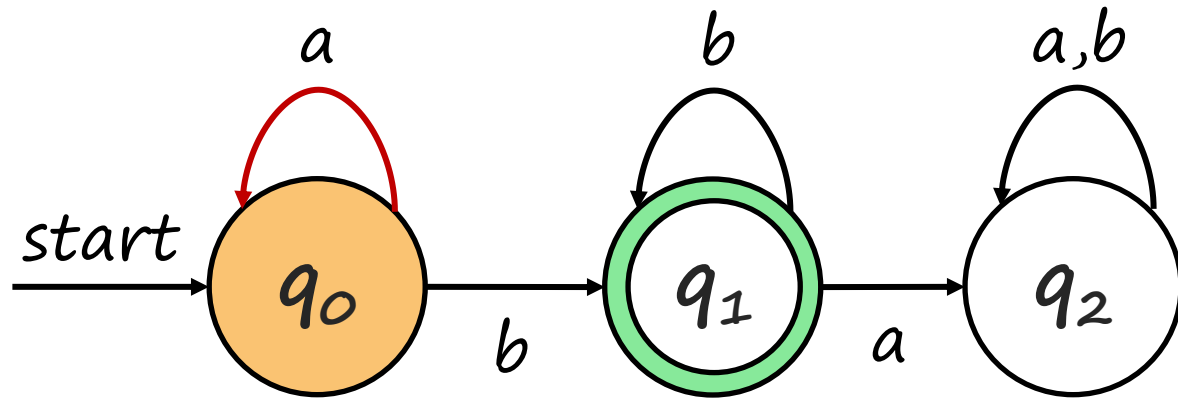


Input: a a b b

next input symbol



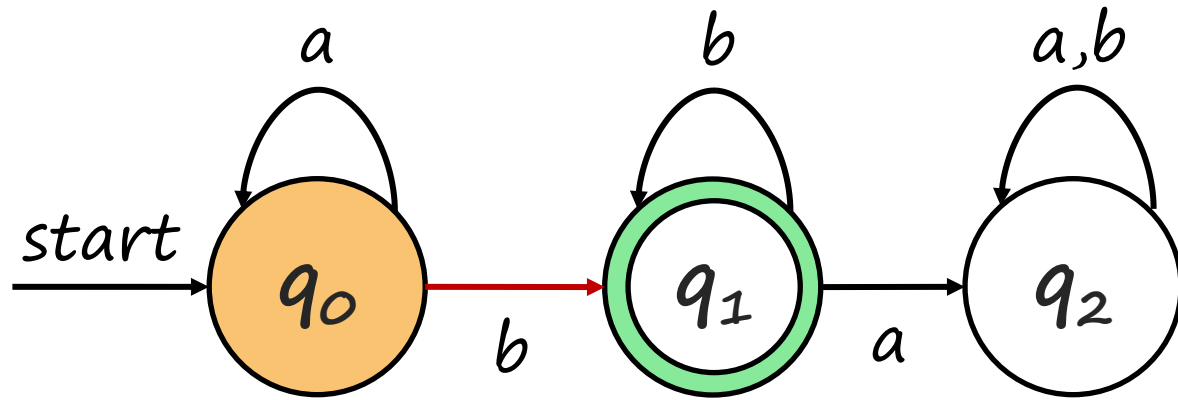
What is an Automaton



Input: a b b a

next input symbol ↑

What is an Automaton

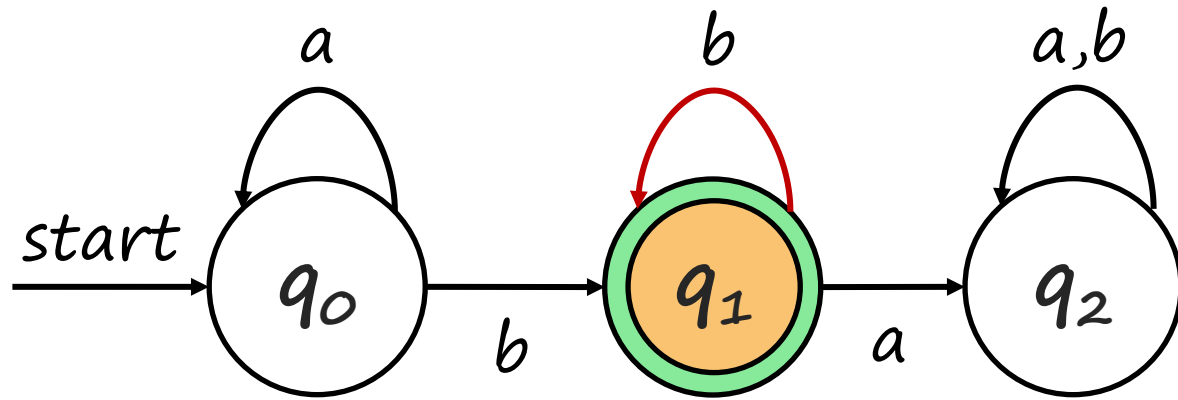


Input: a b b a

next input symbol



What is an Automaton

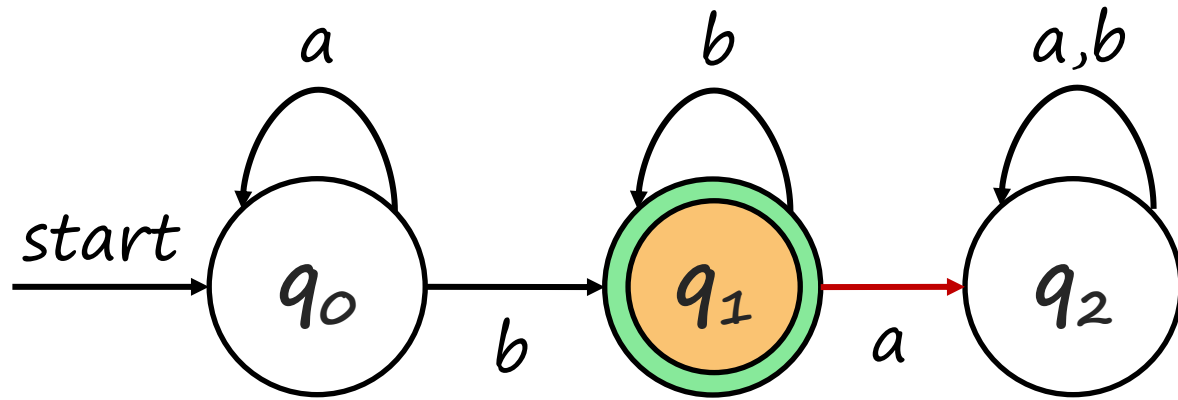


Input: a b b a

next input symbol



What is an Automaton



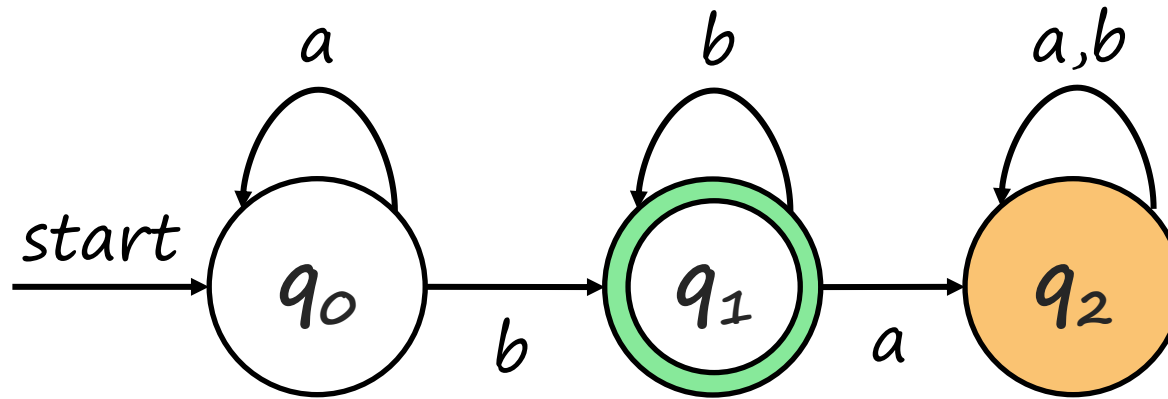
Input: a b b a

next input symbol



What is an Automaton

q_2 is not the final state.



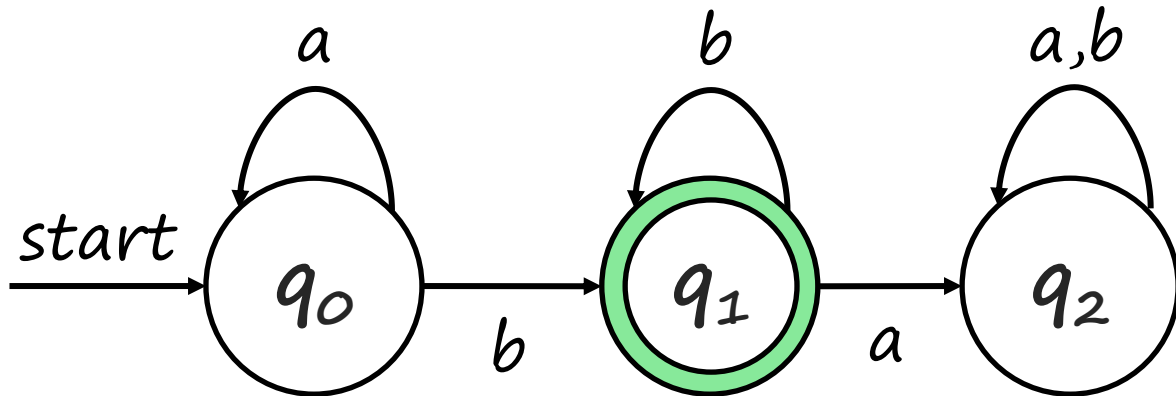
Input: a b b a

next input symbol



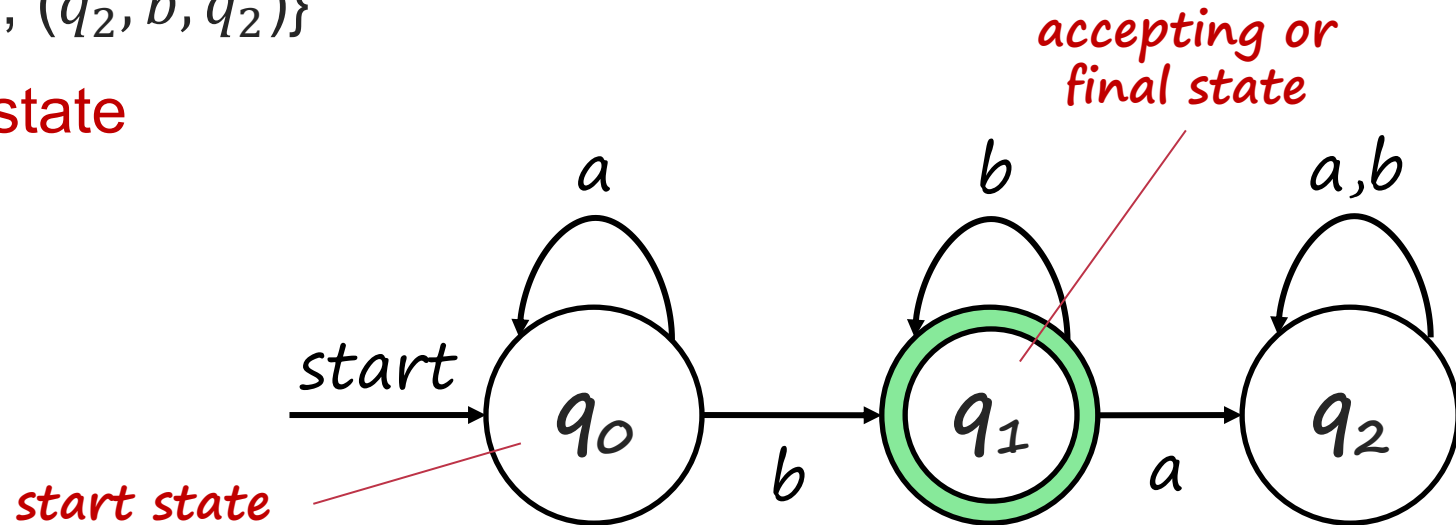
What is an Automaton

- $Q: \{q_0, q_1, q_2\}$
- $\Sigma : \{a, b\}$
- $\delta: \{(q_0, a, q_0), (q_0, b, q_1), (q_1, b, q_1), (q_1, a, q_2), (q_2, a, q_2), (q_2, b, q_2)\}$



What is an Automaton

- $Q: \{q_0, q_1, q_2\}$
- $\Sigma : \{a, b\}$
- $\delta: \{(q_0, a, q_0), (q_0, b, q_1), (q_1, b, q_1), (q_1, a, q_2), (q_2, a, q_2), (q_2, b, q_2)\}$
- q_0 : start state
- $F: \{q_1\}$



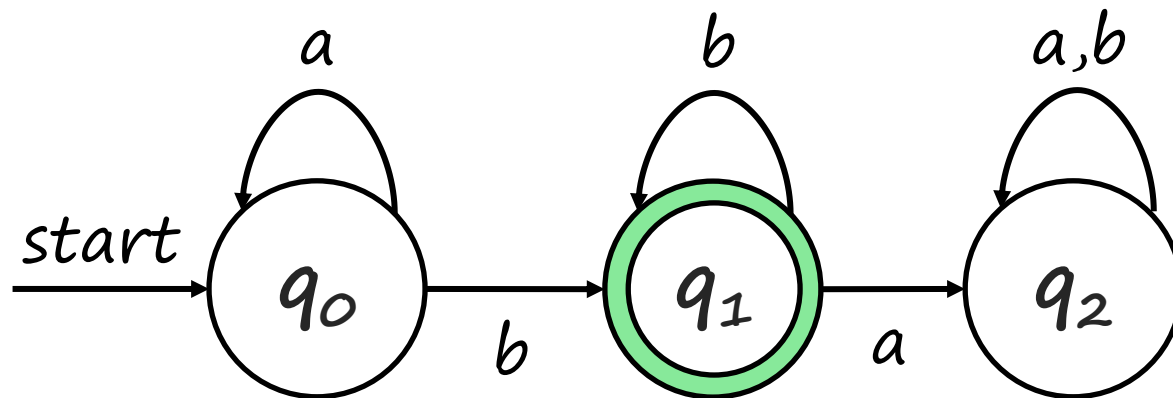
Deterministic Finite Automaton(DFA, 确定有限自动机)

- A **deterministic finite automaton(DFA)** consists of:
 - Q : A finite set of states
 - Σ : A finite set of input symbols
 - δ : A transition function. $Q \times \Sigma \rightarrow Q$
 - q_0 : The start state
 - F : A set of final or accepting states. $F \subseteq Q$
- We can denote a DFA as a “five tuple”
$$A = (Q, \Sigma, \delta, q_0, F)$$

Deterministic Finite Automaton(DFA)

- A DFA takes a finite string(sequence of input symbols) as input, and process the input symbols of the string one by one, and change its state according to the transition function δ .
- DFA will always stop, if the state is in the set of accepting states(F) when a DFA stops, we say that this DFA **accepts** the input. Otherwise it **rejects** the input.

Simpler Notations for DFA



Transition
Diagram

		a	b
start state →	q ₀	q ₀	q ₁
	* q ₁	q ₂	q ₁
	q ₂	q ₂	q ₂

Transition
Table

Extending the Transition Function

For a DFA $(Q, \Sigma, \delta, q_0, F)$, we can define an **extended transition function**

$$\hat{\delta}: Q \times \Sigma^* \rightarrow Q$$

$\hat{\delta}$ takes a state q and a string w as input, and returns the state after processing string w from state q .

- $\hat{\delta}(q, \epsilon) = q$
- If $w = xa$, where $a \in \Sigma$. Then $\hat{\delta}(q, w) = \delta(\hat{\delta}(q, x), a)$

Extending the Transition Function

- $\hat{\delta}(q_0, \epsilon) = q_0$
- $\hat{\delta}(q_0, 1) = \delta(\hat{\delta}(q_0, \epsilon), 1) = \delta(q_0, 1) = q_1$
- $\hat{\delta}(q_0, 11) = \delta(\hat{\delta}(q_0, 1), 1) = \delta(q_1, 1) = q_0$
- $\hat{\delta}(q_0, 110) = \delta(\hat{\delta}(q_0, 11), 0) = \delta(q_0, 0) = q_2$
- $\hat{\delta}(q_0, 1101) = \delta(\hat{\delta}(q_0, 110), 1) = \delta(q_2, 1) = q_3$
- $\hat{\delta}(q_0, 11010) = \delta(\hat{\delta}(q_0, 1101), 0) = \delta(q_3, 0) = q_1$
- $\hat{\delta}(q_0, 110101) = \delta(\hat{\delta}(q_0, 11010), 1) = \delta(q_1, 1) = q_0$

	0	1
* $\rightarrow q_0$	q_2	q_1
q_1	q_3	q_0
q_2	q_0	q_3
q_3	q_1	q_2

The Language of DFA

- For a DFA $A = (Q, \Sigma, \delta, q_0, F)$, the language of A , denoted as $L(A)$, is the set of all strings that A accepts.

$$L(A) = \{w \mid \hat{\delta}(q_0, w) \in F\}$$

- For a certain language L , if there exists a DFA A such that $L = L(A)$, then we call L is a **regular language(正则语言)**.

We'll not take much time in DFA and regular language. Feel free to read chapter 2,3,4 in IALC if you are interested in it.

Regular Expression(正则表达式)

- Regular expressions are used to denote regular languages.
- For an alphabet Σ
 - $\emptyset, \epsilon, a$ are regular expressions corresponding to languages $\emptyset, \{\epsilon\}, \{a\}$, where $a \in \Sigma$
 - If r, s are regular expression corresponding to language L_r, L_s .
 - $r + s$ (or $r|s$): $L_r \cup L_s$
 - rs : $\{wu | w \in L_r \wedge u \in L_s\}$
 - r^k : $\{w_1 w_2 \cdots w_{k-1} w_k | w_1 \in L_r \wedge w_2 \in L_r \wedge \cdots w_k \in L_r\}$
 - r^+ : $L_{r^1} \cup L_{r^2} \cup L_{r^3} \dots \cup L_{r^n} \dots$
 - r^* : $L_{r^+} \cup \{\epsilon\}$



Next

- **Turing Machine Basics.**