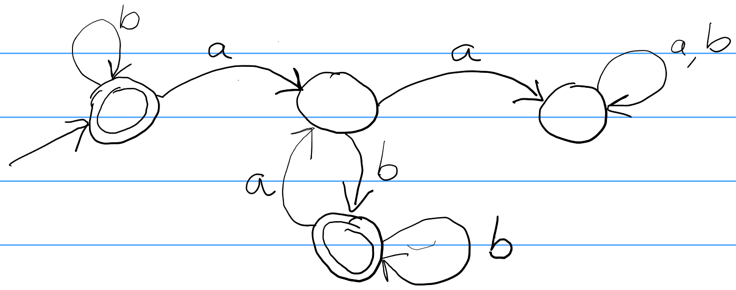
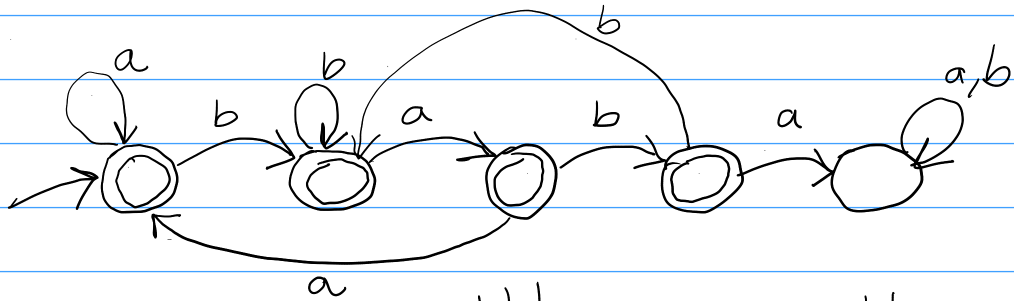
CS605
6 February 2025

W9 = $\{w : w \in \{a,b\}^*$, each a is followed by at least one $b\}$



W11 = $\{w : w \in \{a,b\}^*$, w does not contain the substring baba$\}$



abbbaa
baa

babb.....

DEFINITION 1.23

Let A and B be languages. We define the regular operations ***union***, ***concatenation***, and ***star*** as follows.

- **Union**: $A \cup B = \{x | \ x \in A \text{ or } x \in B\}$.
- **Concatenation**: $A \circ B = \{xy | \ x \in A \text{ and } y \in B\}$.
- **Star**: $A^* = \{x_1 x_2 \ldots x_k | \ k \geq 0 \text{ and each } x_i \in A\}$.

$L_1 = \{w : w \in \{0,1\}^*, w \text{ begins and ends with } 0\}$

$L_2 = \{w : w \in \{0,1\}^*, w \text{ begins and ends with } 1\}$

$L_1 \cup L_2$ is regular.

$L_1 = \{w : w \in \{0,1\}^* \; w \text{ begins with } 1\}$

$L_2 = \{e\}$

$L_1 \cup L_2$ is regular

THEOREM 1.25 ....

The class of regular languages is closed under the union operation.

In other words, if A_1 and A_2 are regular languages, so is $A_1 \cup A_2$.

We will not prove this theorem in CS605 this year, although you should know how to use it.

Example of its use in practice.

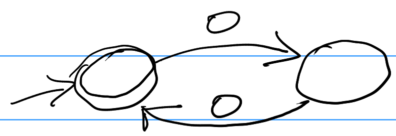
$\{w : w \in \{0\}^*, |w| \text{ is a multiple of 5 or 7}\}$

$= \{e, 0^5, 0^7, 0^{10}, 0^{14}, 0^{15}, 0^{20}, 0^{21}, \ldots$

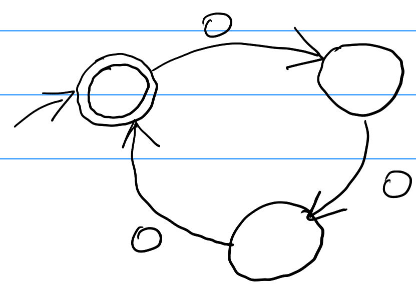
Creating two FA and quoting the closure thm for union would be the easiest way to prove this language is regular.

$L \cup \{\} = L$

Aside: $L = \{w : w \in \{0\}^*, |w| \bmod 2 = 0\} = \{e, 00, 0000, 000000, \ldots\}$



$L = \{w : w \in \{0\}^*, |w| \bmod 3 = 0\} = \{e, 000, 000000, 000000000, \ldots\}$



$e \circ 01 = 0 \circ 01 = 01 \circ e = 01$

Concatenation:

$\{cat, dog\} \circ \{cat, bird\}$

$= \{catcat, catbird, dogcat, dogbird\}$

$e \circ e = e$

$e \circ 1 = 1$

$\cancel{0 \circ 1 = 0}$

$e \circ 10 = 10$

$e \circ 11 = 11$

$00 \circ 1 = 001$

$\{w \in \{0\}^*, |w| \bmod 2 = 0\} \circ \{w \in \{0,1\}^*, w \text{ begins with } 1\}$

$= \{1, 10, 11, 001, \ldots\ldots$

THEOREM 1.26

The class of regular languages is closed under the concatenation operation.

In other words, if A_1 and A_2 are regular languages then so is $A_1 \circ A_2$.

We will not prove this theorem in CS605 this year, although you should know how to use it.

$\{a, bb\} \circ \{\} \circ \{0, 00, 000, \ldots\ldots\} = \{\}$

Star operation (Kleene star operation)

$\{00\}^* = \{e, 00, 0000, 000000, \ldots\}^*$

The star operation on a language L is the lang of words that can be formed by taking zero or more words from L and concatenating them together.

$\{1, 001\}^* = \{e, 1, 11, 001, 111, 0011, 1001, 1111, \ldots\}$

$\{\}^* = \{e\}$

THEOREM 1.49

The class of regular languages is closed under the star operation.

We will ~~not~~ prove this theorem in CS605 this year, although you should know how to use it.

Complement

$$L = \{w : w \in \{0,1\}^*, w \text{ begins with } 0\}$$
$$= \{0, 00, 01, 000, 001, 010, \ldots\}$$

$$\overline{L} = \{w : w \in \{0,1\}^*, w \text{ does not begin with } 0\}$$

$$= \{e, 1, 10, 11, 100, 101, 110, \ldots\}$$

Thm

The regular languages are closed under complement. In other words, if A, is a regular language, then $\overline{A}$ is regular.

We will not prove this theorem in CS605 this year, although you should know how to use it.

Intersection:

$\{w : w \in \{0\}^*, |w| \text{ is a multiple of } 5 \text{ and } 7\}$

$= \{e, 0^{35}, 0^{70}, 0^{105}, \ldots\} = L_1 \cap L_2$

where

$L_1 = \{w : w \in \{0\}^*, |w| \text{ is a multiple of } 5\}$

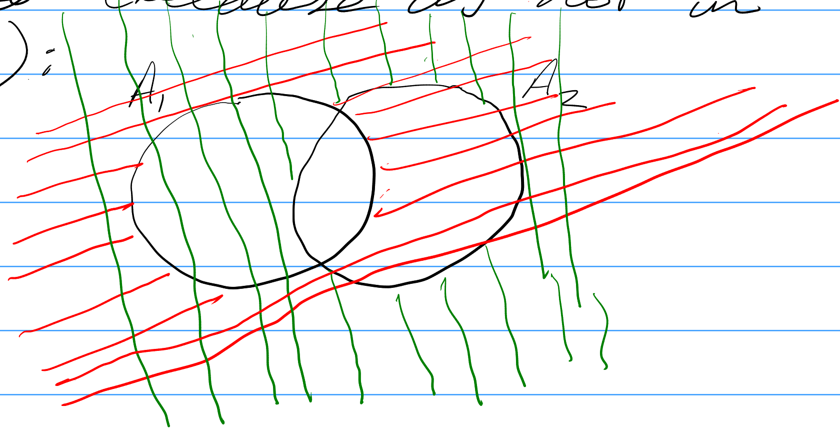
$L_2 = \{w : w \in \{0\}^*, |w| \text{ is a multiple of } 7\}$

Thm

The regular languages are closed under intersection. In other words, if A_1 and A_2 are regular languages then $A_1 \cap A_2$ is regular.

We will ~~not~~ prove this theorem in CS605 this year, although you should know how to use it.

However, just to give you a flavour of what's involved (because its not in Sipser's book):

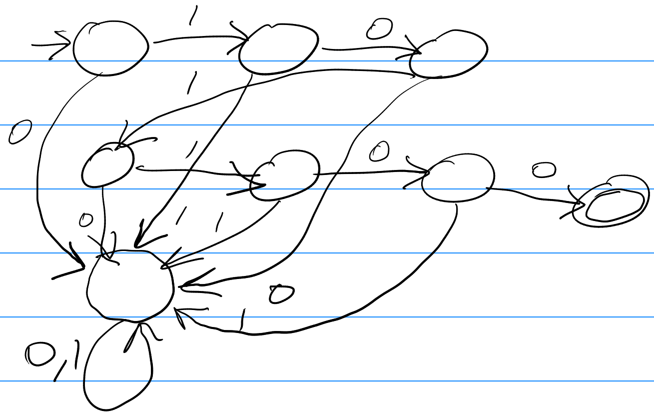


$$\overline{\overline{A_1} \cup \overline{A_2}} = A_1 \cap A_2$$

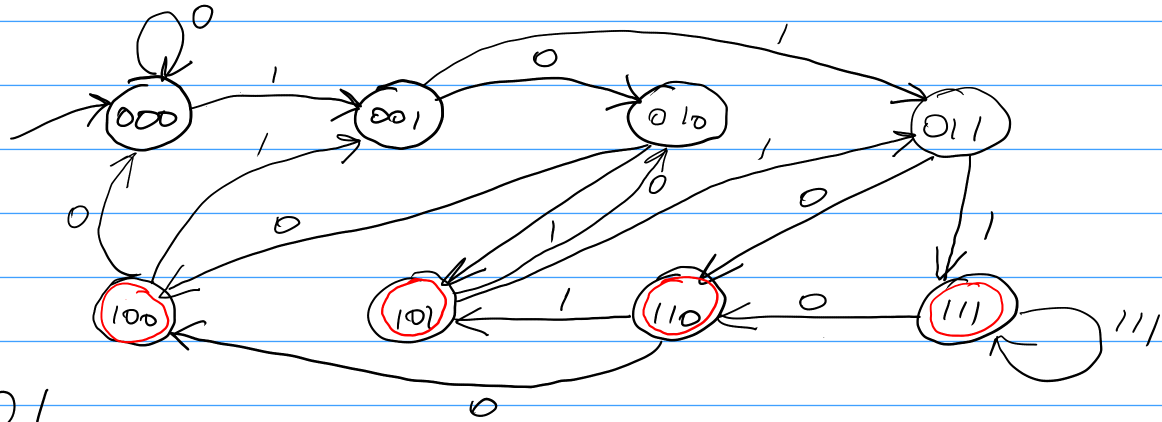
All finite languages are regular.

The proof involves the union of a finite number of languages of one word each.

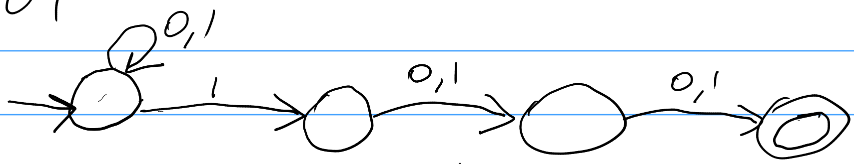
$L_1 \{101100\}$ $M_1 =$



$A = \{ w : w \in \{0,1\}^*, \text{ contains a } 1 \text{ in its third last position} \}$



10110101



↳ This is a nondeterministic FA

Nondeterminism

Useful. Great impact on CS.

THEOREM 1.39

Every nondeterministic finite automaton has an equivalent deterministic finite automaton.

We'll look at nondeterminism again in the context of PDA and the class NP.

THE PUMPING LEMMA FOR REGULAR LANGUAGES

Our technique for proving nonregularity stems from a theorem about regular languages, traditionally called the ***pumping lemma***. This theorem states that all regular languages have a special property. If we can show that a language does not have this property, we are guaranteed that it is not regular. The property states that all strings in the language can be "pumped" if they are at least as long as a certain special value, called the ***pumping length***. That means each such string contains a section that can be repeated any number of times with the resulting string remaining in the language.

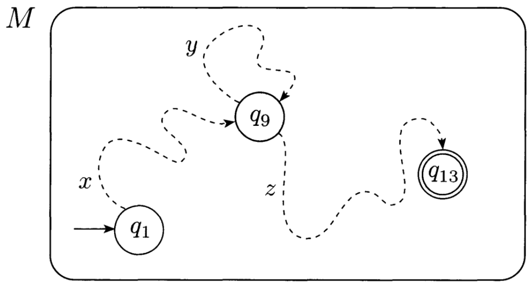


FIGURE 1.72
Example showing how the strings x, y, and z affect M

THEOREM 1.70

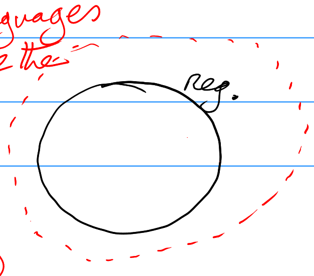
Pumping lemma If A is a regular language, then there is a number p (the pumping length) where, if s is any string in A of length at least p, then s may be divided into three pieces, $s = xyz$, satisfying the following conditions:

1. for each $i \geq 0$, $xy^iz \in A$,
2. $|y| > 0$, and
3. $|xy| \leq p$.

Recall the notation where $|s|$ represents the length of string s, y^i means that i copies of y are concatenated together, and y^0 equals ε.

When s is divided into xyz, either x or z may be ε, but condition 2 says that $y \neq \varepsilon$. Observe that without condition 2 the theorem would be trivially true. Condition 3 states that the pieces x and y together have length at most p. It is an extra technical condition that we occasionally find useful when proving certain languages to be nonregular. See Example 1.74 for an application of condition 3.

set of languages that have the pumping property (can be pumped)



Non-Regular languages

• $a^n b^n$

• $0^n 1^n$

Prove that $B = \{0^n 1^n : n \geq 0\}$ is not regular.

Proof: We will prove this by contradiction using the pumping lemma.

Assume B is regular. Let p be the pumping length given by the pumping lemma. Let $w = 0^p 1^p$.

According to the pumping lemma, w can be divided into three pieces $w = xyz$ satisfying three conditions.

These are all the ways that w can be divided:

$0_1 0_2 \ldots\ldots 0_p 1_1 1_2 \ldots\ldots 1_p$

y only zeros: $\underbrace{\quad}_{x}\underbrace{\quad\quad}_{y}\underbrace{\quad\quad\quad\quad\quad\quad}_{z}$ X 1 more 0s than 1s

y has 0s and 1s: $\underbrace{\quad\quad\quad}_{x}\underbrace{\quad\quad\quad}_{y}\underbrace{\quad\quad}_{z}$ X 1 3 out of order symbols

y only has 1s: $\underbrace{\quad\quad\quad\quad\quad}_{x}\underbrace{\quad\quad}_{y}\underbrace{\quad\quad}_{z}$ X 1 3 more 1s than 0s

There is no way to pump this word. A contradiction because the pumping lemma states it should be possible. Therefore this proves B is not regular.

Why the second way of dividing w does not work.

$p = 4$

$x = 0^{p-1}$ $\quad xyz = w$

$y = 01$ $\quad xy^2z = 0^p 1 0 1^p$

$z = 1^{p-1}$

$\underbrace{000}_{x}\underbrace{01}_{y}\underbrace{111}_{z}$ (0000 1111)

$xy^2z = xyyz = 000010 1111$

$C = \{w : w \in \{0,1\}^*, w \text{ has an equal number of } 0\text{s and } 1\text{s}\}$

Prove C is not regular.

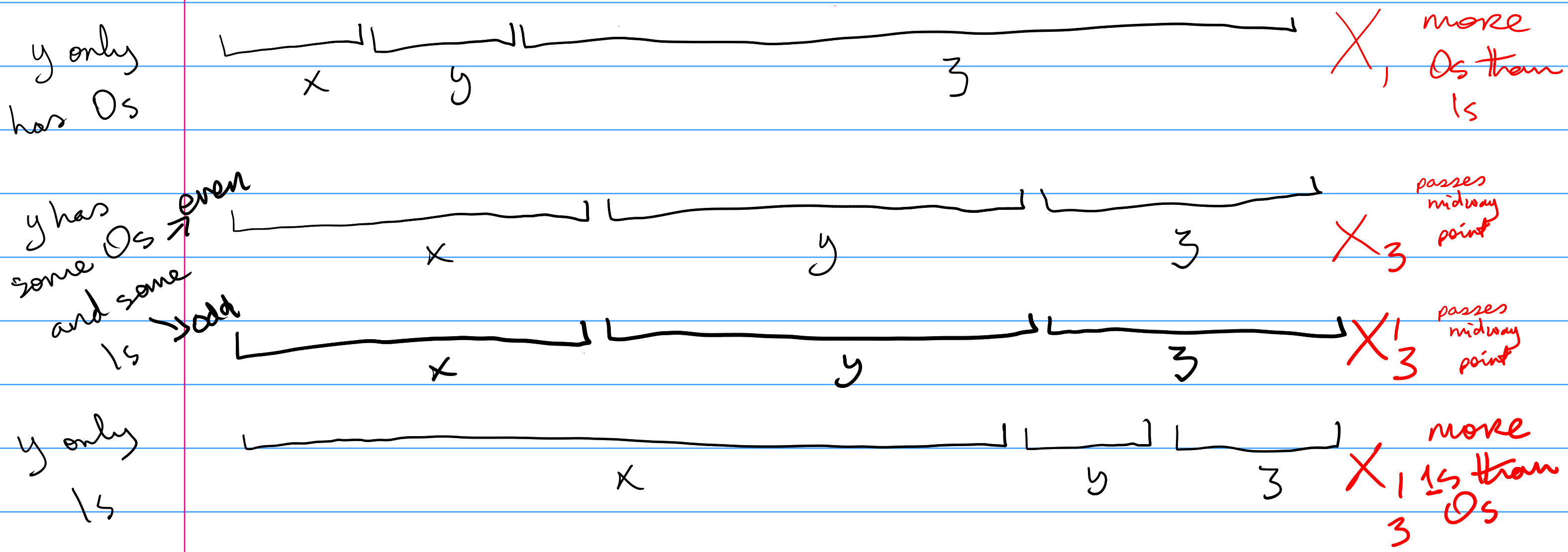
Proof: We will prove this by contradiction using the pumping lemma.

Assume C is regular. Let p be the pumping length given by the pumping lemma. Let $w = 0^p 1^p$.

According to the pumping lemma, w can be divided into three pieces $w = xyz$ satisfying three conditions.

These are all the ways that w can be divided:

$0_1 0_2 0_3 \ldots\ldots\ldots 0_p 1_1 1_2 1_3 \ldots\ldots 1_p$



There is no way to pump this word. A contradiction because the pumping lemma states it should be possible. Therefore this proves C is not regular.

$F = \{ww : w \in \{0,1\}^*\} = \{e, 00, 11, 0000, 0101, \ldots$

Proof We will prove this by contradiction using the pumping lemma.

Assume F is regular. Let p be the pumping length given by the pumping lemma. Let $w = 0^p 1 0^p 1$.

According to the pumping lemma, w can be divided into three pieces $w = xyz$ satisfying three conditions.

These are all the w's that w can be divided:

$0_1 0_2 \ldots\ldots 0_p 1 \; 0_1 0_2 \ldots\ldots 0_p 1$

y only zeros: $x \quad y \quad z$ — X, more 0s in first w compared to 2nd w OR the second w has two 1s.

y has 1st 1: $x \quad y \quad z$ — X, 3 could have an odd no. of 1s.

y only second group of 0s: $x \quad y \quad z$ — X, 3 first w not end with 1

y has 2nd 1: $x \quad y \quad \underset{z}{e}$ — X, 3 odd number of 1s

y has both 1s: $x \quad y \quad \underset{z}{e}$ — X_3

There is no way to pump this word. A contradiction because the pumping lemma states it should be possible. Therefore this proves F is not regular.

Friday morning lab work (7 February 2025)
Exercises to do yourself.
Prove these languages (L1 - L3) are not regular using the text and approach
I have demonstrated (30 mins), and then get into groups of two (with one group of 3) and discuss your approaches (20 mins).
Then construct PDAs for the below L4, L5 languages to prove they are regular (40 mins).

L1.

EXAMPLE 1.77

Sometimes "pumping down" is useful when we apply the pumping lemma. We use the pumping lemma to show that $E = \{0^i 1^j | \; i > j\}$ is not regular. The proof is by contradiction.

Assume that E is regular. Let p be the pumping length for E given by the pumping lemma. Let $s = 0^{p+1}1^p$. Then s can be split into xyz, satisfying the conditions of the pumping lemma. By condition 3, y consists only of 0s. Let's examine the string $xyyz$ to see whether it can be in E. Adding an extra copy of y increases the number of 0s. But, E contains all strings in 0^*1^* that have more 0s than 1s, so increasing the number of 0s will still give a string in E. No contradiction occurs. We need to try something else.

The pumping lemma states that $xy^iz \in E$ even when $i = 0$, so let's consider the string $xy^0z = xz$. Removing string y decreases the number of 0s in s. Recall that s has just one more 0 than 1. Therefore xz cannot have more 0s than 1s, so it cannot be a member of E. Thus we obtain a contradiction.

L2 $\{ww^R : w \in \{0,1\}^*\}$

L3 $\{wxw^R : w \in \{0,1\}^*\}$

PDAs

L4 $\{xwyw^R z : x, y, z, w \in \{a,b\}^*, \; |w| > 0\}$

L5 $\{w : w \in \{0,1\}^*, \; w \text{ has twice as many 0s as 1s}\}$