CS103 Problem Set 7 (Written)

Matthew Phonchay Vilaysack, Maria Shan Wang

TOTAL POINTS

24 / 27

QUESTION 1

1 Regexes and Accessible Design 3 / 3

✓ + **3 pts All correct**

+ **2 pts** Two correct
+ **1 pts** One correct
+ **0 pts** Click here to replace this description.

QUESTION 2

2 Finite Languages 1 / 2

+ **2 pts** Correct

✓ + **1 pts Partially Correct**

+ **0 pts** Incorrect
+ **0 pts** Check

1 If there are no strings in the language, $$\phi$$ would be the valid regular expression. Need to note this explicitly.

QUESTION 3

Embracing the Braces 7 pts

3.1 Part (i) 5 / 5

Intuition Scale

✓ + **2 pts Correct**

+ **1 pts** Partially correct intuition
+ **0 pts** Incorrect

Execution Scale

✓ + **3 pts Correct**

+ **2 pts** Minor execution error
+ **1 pts** Major execution error
+ **0 pts** Incorrect

3.2 Part (iii) 2 / 2

✓ + **2 pts Correct**

+ **1 pts** Partially correct
+ **0 pts** Incorrect

QUESTION 4

State Lower Bounds 15 pts

4.1 Part (i) 5 / 5

✓ + **5 pts Correct**

Intuition Scale

+ **2 pts** Correct
+ **1 pts** Partially correct intuition
+ **0 pts** Incorrect

Execution Scale

+ **3 pts** Correct
+ **2 pts** Minor execution error
+ **1 pts** Major execution error
+ **0 pts** Incorrect

+ **0 pts** Incorrect

4.2 Part (ii) 5 / 5

Intuition Scale

✓ + **2 pts Correct**

+ **1 pts** Partially correct intuition
+ **0 pts** Incorrect

Execution Scale

✓ + **3 pts Correct**

+ **2 pts** Minor execution error
+ **1 pts** Major execution error
+ **0 pts** Incorrect

4.3 Part (iii) 3 / 5

✓ + **5 pts Base score**

Intuition errors

- **1 pts** Minor intuition error
- **2 pts** Major intuition error

Execution errors

- 1 pts Minor execution error

✓ - 2 pts **Major execution error**

- 3 pts Fully incorrect execution

+ 0 pts No submission / fully incorrect.

2 You don't have $$q_0$$ defined in $$Q$$.

3 Incorrect transition function.

4 No explanation of how your DFA works.

QUESTION 5

5 Optional Fun Problem: Birget's Theorem

0 / 0

+ 0 pts Correct! Congratulations!!

✓ + 0 pts **Incorrect / Not attempted**

Problem Two: Regexes and Accessible Design

Marc Trevor Tessier-Lavigne
- The bracket is not allowed under the regex.
Timothée Hal Chalamet
- The accent on his first name is not allowed under the regex.
Abraham Lincoln
- He has no middle name.

1 Regexes and Accessible Design 3 / 3

- ✓ + **3 pts** **All correct**
- + **2 pts** Two correct
- + **1 pts** One correct
- + **0 pts** Click here to replace this description.

Problem Three: Finite Languages

A method that is guaranteed to work is by writing a union of every single string in the finite language. This can be its regular expression because it includes every string in the language. However this is not the best way to do it and can result in a huge (albeit finite) regex. The best way is to write an NFA for the language and reduce the NFA to a regex using the state-elimination algorithm.

2 Finite Languages 1 / 2

+ **2 pts** Correct

✓ + **1 pts** **Partially Correct**

+ **0 pts** Incorrect

+ **0 pts** Check

1 If there are no strings in the language, $$\phi$$ would be the valid regular expression. Need to note this explicitly.

Problem Five: Embracing the Braces

i.

Theorem: The language defined as $L_1 = \{w \in \{\{,\}\} | w$ is a string of balanced curly braces$\}$ is not a regular language.
Proof: Let $S = \{\{^n | n \in \mathbb{N}\}$. We will prove that S is infinite and that S is a distinguishing set for L_1.
To see that S is infinite, note that S contains one string for each natural number.
To see that S is a distinguishing set for L_1, consider any strings $\{^p, \{^q \in S$ where $p \neq q$. Note that $\{^p\}^p \in L_1$ but $\{^q\}^p \notin L_1$. Therefore, we see that $\{^p \not\equiv_{L_1} \{^q$, as required.
Since S is infinite and is a distinguishing set for L_1, by the Myhill-Nerode theorem, we see that L is not regular. ■

iii.

$\{^p\}^p \notin L_2$ if $p > 4$, so you could not pick any string $\{^p \in S$ and therefore could not conclude that S is a distinguishing set.

3.1 Part (i) 5 / 5

Intuition Scale

✓ + **2 pts** **Correct**

+ **1 pts** Partially correct intuition
+ **0 pts** Incorrect

Execution Scale

✓ + **3 pts** **Correct**

+ **2 pts** Minor execution error
+ **1 pts** Major execution error
+ **0 pts** Incorrect

Problem Five: Embracing the Braces

i.

Theorem: The language defined as $L_1 = \{w \in \{\{,\}\} | w$ is a string of balanced curly braces$\}$ is not a regular language.
Proof: Let $S = \{\{^n | n \in \mathbb{N}\}$. We will prove that S is infinite and that S is a distinguishing set for L_1.
To see that S is infinite, note that S contains one string for each natural number.
To see that S is a distinguishing set for L_1, consider any strings $\{^p, \{^q \in S$ where $p \neq q$. Note that $\{^p\}^p \in L_1$ but $\{^q\}^p \notin L_1$. Therefore, we see that $\{^p \not\equiv_{L_1} \{^q$, as required.
Since S is infinite and is a distinguishing set for L_1, by the Myhill-Nerode theorem, we see that L is not regular. ■

iii.

$\{^p\}^p \notin L_2$ if $p > 4$, so you could not pick any string $\{^p \in S$ and therefore could not conclude that S is a distinguishing set.

3.2 Part (iii) 2 / 2

✓ + **2 pts** Correct

\+ **1 pts** Partially correct

\+ **0 pts** Incorrect

Problem Six: State Lower Bounds

i.

Theorem: Let L be a language over Σ and let S be a distinguishing set for L. If S is finite, then any DFA for L must have at least $|S|$ states.
Proof: Assume that S is a finite distinguishing set for L. Let $S = \{x_1, \ldots, x_k\}$. We will prove that any DFA for L must have at least $|S| = k$ states. Consider two cases:
Case 1: Assume that L is not regular. There does not exist a DFA for L and therefore there are no DFAs for L with less than k states.
Case 2: Assume that L is regular. Let D be a DFA for L. Pick arbitrary strings $x_m, x_n \in S$ where $m \neq n$. Because S is a distinguishing set and $m \neq n$, then there exists a string $w \in \Sigma*$ such that $x_m w \in L$ and $x_n w \notin L$. Therefore one of $x_m w$ and $x_n w$ leads to an accepting state q_m and one leads to a rejecting state q_n, without loss of generality. Therefore each of x_m and x_n lead to different states. Therefore there must be at least k distinct states $q_1, \ldots, q_k$ in D.
In either case, any DFA for L must have at least $|S| = k$ states. ■

ii.

Theorem: Any DFA for $TWEETS = \{w \in \Sigma * ||w| \leq 280\}$ must have at least 282 states.
Proof: Pick an arbitrary DFA D for $TWEETS$. We will prove that D has at least 282 states. To do so, we will show that there exists a distinguishing set S such that $|S| = 282$.
Let $S = \{a^n | n \in \mathbb{N} \text{ and } n \leq 281\}$. We see that S contains one string for every $0 \leq n \leq 281$, therefore S is finite and $|S| = 282$.
Pick any $a^m, a^n \in S$ such that $m < n$ without loss of generality. Let the string $w_{mn} \in \Sigma*$ be the string a^{280-m}. We see that $a^m w_{mn} = a^m a^{280-m} = a^{280}$ and that $a^{280} \in L$. We also see that $a^n w^{mn} = a^n a^{280-m} = a^{280-m+n}$. Because $m < n$, then $(280 - m + n) > 280$ and that $a^{280-m+n} \notin L$. Therefore there exists a $w_{mn} \in \Sigma*$ such that $a^m w_{mn} \in \Sigma*$ and $a^n w_{mn} \notin \Sigma*$ and that $a^m \not\equiv_L a_n$. Therefore S is a distinguishing set.
Because S is a finite distinguishing set of where $|S| = 282$, by the theorem in Problem 6, Part I, D must have at least 282 states. ■

iii.

Let the set $Q = \{q_1, \ldots, q_{282}\}$.
Let Σ be the set of all unicode characters allowed in a tweet.
Let the set $\delta = \{q_m \times \Sigma \rightarrow q_{m+1} | m \in \mathbb{N} \text{ where } 0 \leq m \leq 281\} \cup q_{282} \times \Sigma \rightarrow q_{282}$
Let the set $F = \{q_1, \ldots, q_{281}\}$.
The formal definition of a DFA for $TWEETS$ is $(Q, \Sigma, \delta, q_1, F)$.

4.1 Part (i) **5 / 5**

✓ **+ 5 pts** **Correct**

Intuition Scale

+ **2 pts** Correct
+ **1 pts** Partially correct intuition
+ **0 pts** Incorrect

Execution Scale

+ **3 pts** Correct
+ **2 pts** Minor execution error
+ **1 pts** Major execution error
+ **0 pts** Incorrect

+ **0 pts** Incorrect

Problem Six: State Lower Bounds

i.

Theorem: Let L be a language over Σ and let S be a distinguishing set for L. If S is finite, then any DFA for L must have at least $|S|$ states.
Proof: Assume that S is a finite distinguishing set for L. Let $S = \{x_1, \ldots, x_k\}$. We will prove that any DFA for L must have at least $|S| = k$ states. Consider two cases:
Case 1: Assume that L is not regular. There does not exist a DFA for L and therefore there are no DFAs for L with less than k states.
Case 2: Assume that L is regular. Let D be a DFA for L. Pick arbitrary strings $x_m, x_n \in S$ where $m \neq n$. Because S is a distinguishing set and $m \neq n$, then there exists a string $w \in \Sigma*$ such that $x_m w \in L$ and $x_n w \notin L$. Therefore one of $x_m w$ and $x_n w$ leads to an accepting state q_m and one leads to a rejecting state q_n, without loss of generality. Therefore each of x_m and x_n lead to different states. Therefore there must be at least k distinct states $q_1, \ldots, q_k$ in D.
In either case, any DFA for L must have at least $|S| = k$ states. ■

ii.

Theorem: Any DFA for $TWEETS = \{w \in \Sigma * \,||w| \leq 280\}$ must have at least 282 states.
Proof: Pick an arbitrary DFA D for $TWEETS$. We will prove that D has at least 282 states. To do so, we will show that there exists a distinguishing set S such that $|S| = 282$.
Let $S = \{a^n | n \in \mathbb{N} \text{ and } n \leq 281\}$. We see that S contains one string for every $0 \leq n \leq 281$, therefore S is finite and $|S| = 282$.
Pick any $a^m, a^n \in S$ such that $m < n$ without loss of generality. Let the string $w_{mn} \in \Sigma*$ be the string a^{280-m}. We see that $a^m w_{mn} = a^m a^{280-m} = a^{280}$ and that $a^{280} \in L$. We also see that $a^n w^{mn} = a^n a^{280-m} = a^{280-m+n}$. Because $m < n$, then $(280 - m + n) > 280$ and that $a^{280-m+n} \notin L$. Therefore there exists a $w_{mn} \in \Sigma*$ such that $a^m w_{mn} \in \Sigma*$ and $a^n w_{mn} \notin \Sigma*$ and that $a^m \not\equiv_L a_n$. Therefore S is a distinguishing set.
Because S is a finite distinguishing set of where $|S| = 282$, by the theorem in Problem 6, Part I, D must have at least 282 states. ■

iii.

Let the set $Q = \{q_1, \ldots, q_{282}\}$.
Let Σ be the set of all unicode characters allowed in a tweet.
Let the set $\delta = \{q_m \times \Sigma \to q_{m+1} | m \in \mathbb{N} \text{ where } 0 \leq m \leq 281\} \cup q_{282} \times \Sigma \to q_{282}$
Let the set $F = \{q_1, \ldots, q_{281}\}$.
The formal definition of a DFA for $TWEETS$ is $(Q, \Sigma, \delta, q_1, F)$.

4.2 Part (ii) 5 / 5

Intuition Scale

✓ + **2 pts** **Correct**

+ **1 pts** Partially correct intuition

+ **0 pts** Incorrect

Execution Scale

✓ + **3 pts** **Correct**

+ **2 pts** Minor execution error

+ **1 pts** Major execution error

+ **0 pts** Incorrect

Problem Six: State Lower Bounds

i.

Theorem: Let L be a language over Σ and let S be a distinguishing set for L. If S is finite, then any DFA for L must have at least $|S|$ states.
Proof: Assume that S is a finite distinguishing set for L. Let $S = \{x_1, \ldots, x_k\}$. We will prove that any DFA for L must have at least $|S| = k$ states. Consider two cases:
Case 1: Assume that L is not regular. There does not exist a DFA for L and therefore there are no DFAs for L with less than k states.
Case 2: Assume that L is regular. Let D be a DFA for L. Pick arbitrary strings $x_m, x_n \in S$ where $m \neq n$. Because S is a distinguishing set and $m \neq n$, then there exists a string $w \in \Sigma*$ such that $x_m w \in L$ and $x_n w \notin L$. Therefore one of $x_m w$ and $x_n w$ leads to an accepting state q_m and one leads to a rejecting state q_n, without loss of generality. Therefore each of x_m and x_n lead to different states. Therefore there must be at least k distinct states $q_1, \ldots, q_k$ in D.
In either case, any DFA for L must have at least $|S| = k$ states. ■

ii.

Theorem: Any DFA for $TWEETS = \{w \in \Sigma * ||w| \leq 280\}$ must have at least 282 states.
Proof: Pick an arbitrary DFA D for $TWEETS$. We will prove that D has at least 282 states. To do so, we will show that there exists a distinguishing set S such that $|S| = 282$.
Let $S = \{a^n | n \in \mathbb{N} \text{ and } n \leq 281\}$. We see that S contains one string for every $0 \leq n \leq 281$, therefore S is finite and $|S| = 282$.
Pick any $a^m, a^n \in S$ such that $m < n$ without loss of generality. Let the string $w_{mn} \in \Sigma*$ be the string a^{280-m}. We see that $a^m w_{mn} = a^m a^{280-m} = a^{280}$ and that $a^{280} \in L$. We also see that $a^n w^{mn} = a^n a^{280-m} = a^{280-m+n}$. Because $m < n$, then $(280 - m + n) > 280$ and that $a^{280-m+n} \notin L$. Therefore there exists a $w_{mn} \in \Sigma*$ such that $a^m w_{mn} \in \Sigma*$ and $a^n w_{mn} \notin \Sigma*$ and that $a^m \not\equiv_L a_n$. Therefore S is a distinguishing set.
Because S is a finite distinguishing set of where $|S| = 282$, by the theorem in Problem 6, Part I, D must have at least 282 states. ■

iii.

Let the set $Q = \{q_1, \ldots, q_{282}\}$.
Let Σ be the set of all unicode characters allowed in a tweet.
Let the set $\delta = \{q_m \times \Sigma \to q_{m+1} | m \in \mathbb{N} \text{ where } 0 \leq m \leq 281\} \cup q_{282} \times \Sigma \to q_{282}$
Let the set $F = \{q_1, \ldots, q_{281}\}$.
The formal definition of a DFA for $TWEETS$ is $(Q, \Sigma, \delta, q_1, F)$.

4.3 Part (iii) 3 / 5

✓ **+ 5 pts** **Base score**

Intuition errors

- **- 1 pts** Minor intuition error
- **- 2 pts** Major intuition error

Execution errors

- **- 1 pts** Minor execution error
- ✓ **- 2 pts** **Major execution error**
- **- 3 pts** Fully incorrect execution

+ 0 pts No submission / fully incorrect.

2 You don't have $$q_0$$ defined in $$Q$$.

3 Incorrect transition function.

4 No explanation of how your DFA works.

5 Optional Fun Problem: Birget's Theorem **0 / 0**

+ **0 pts** Correct! Congratulations!!

✓ + **0 pts** **Incorrect / Not attempted**

CS 103: Mathematical Foundations of Computing
Problem Set #7

[TODO: Replace this with your name(s)]

May 20, 2022

Due Friday, November 12 at 2:30 pm Pacific

Do not put your answers to Problems 1, 4, and 5.ii. in this file.
Some symbols you may want to use here:

- The empty string is denoted ε.
- Alphabets are denoted Σ.
- The language of an automaton is denoted $\mathscr{L}(D)$.
- The Greek letter delta is denoted δ.
- You can say two strings are distinguishable relative to L by writing $x \not\equiv_L y$.
- The union symbol for regexes is denoted $\cup$.
- Kleene stars are denoted $L^\star$.
- You can type a brace character by writing { or }.
- The Optional Fun Problem uses the notation $\mathscr{F}$.

Optional Fun Problem: Birgetâ€™s Theorem

Write your answer to the Optional Fun Problem here.