

程序代写代做 CS编程辅导

FIT2014 Theory of Computation



Lecture 10

Simplifying Finite Automata, and Lexical Analysis

WeChat: cstutorcs

Assignment Project Exam Help

slides by Graham Farr

based in part on previous slides by David Albrecht

QQ: 749389476

COMMONWEALTH OF AUSTRALIA
<https://tutorcs.com>

Copyright Regulations 1969

Warning

This material has been reproduced and communicated to you by or on behalf of Monash University in accordance with s113P of the Copyright Act 1968 (the Act).

The material in this communication may be subject to copyright under the Act.

Any further reproduction or communication of this material by you may be the subject of copyright protection under the Act.

Do not remove this notice.

Overview

程序代写代做 CS编程辅导



- ▶ Simplifying Finite Automata
- ▶ Implementing Finite Automata
- ▶ Lexical Analyzer
- ▶ Tokens and lexemes

WeChat: cstutorcs

Assignment Project Exam Help

Email: tutorcs@163.com

QQ: 749389476

<https://tutorcs.com>

Matching a Regular Expression

程序代写代做 CS编程辅导

Write a program which reads in a character string, over alphabet $\{a,b\}$, one character at a time and identifies whether the string matches the following regular expression.



$(a \cup bb \cup baa^*b)^*baa^*$
WeChat: cstutorcs

Assignment Project Exam Help

1. Convert regular expression to NFA.

2. Convert NFA to DFA

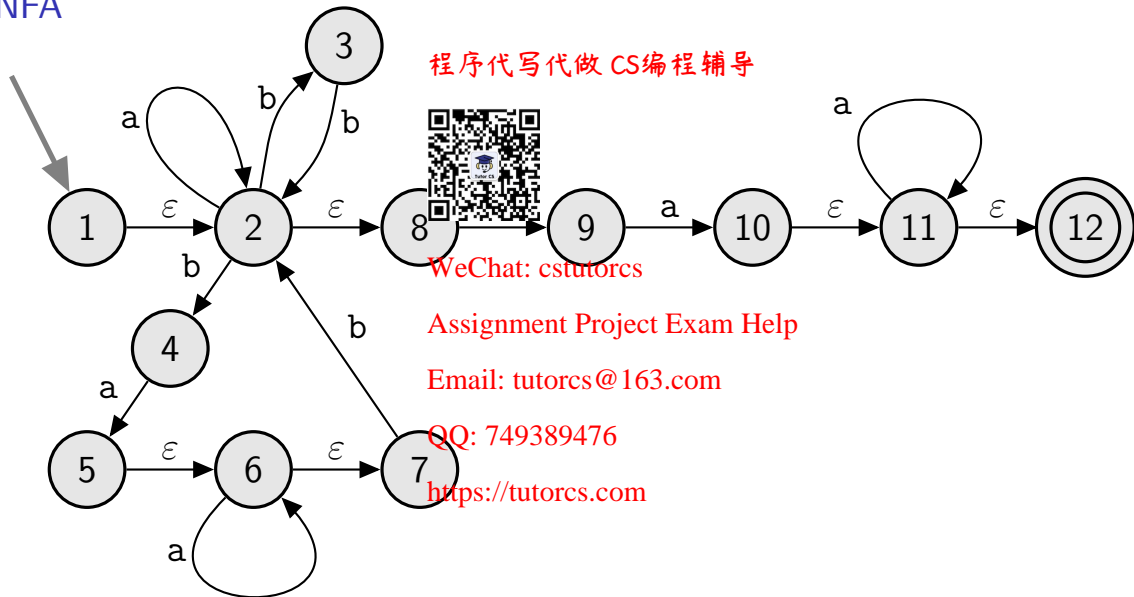
3. **Simplify DFA.**

Email: tutorcs@163.com

QQ: 749389476

<https://tutorcs.com>

NFA



程序代写代做 CS编程辅导

WeChat: cstutorcs

Assignment Project Exam Help

Email: tutorcs@163.com

QQ: 749389476

<https://tutorcs.com>

程序代写代做 CS编程辅导



			b
Start	{1,2,8}	{2,8}	{3,4,9}
	{2,8}	{2,8}	{3,4,9}
	{3,4,9}	{5,6,7,10,11,12}	{2,8}
Final	{5,6,7,10,11,12}	{6,7,11,12}	{2,8}
Final	{6,7,11,12}	{6,7,11,12}	{2,8}

QQ: 749389476

<https://tutorcs.com>

A **Final State** and a **non-Final State** are *fundamentally different*.

They *cannot be combined*.

程序代写代做 CS编程辅导

So:

- ▶ Give all **Final States** one
- ▶ Give all **non-Final States** a *different* colour.



WeChat: cstutorcs

Different colours \implies different states.

Assignment Project Exam Help

- ▶ They cannot be combined.

Email: tutorcs@163.com

Same colours \nRightarrow same states.

QQ: 749389476

- ▶ The states *may or may not* be combined.

<https://tutorcs.com>

- ▶ We have not yet ruled out combining them.

程序代写代做 CS编程辅导



b

Start

{1,2,8}

{2,8}

{3,4,9}

Final

{5,6,7,10,11,12}

Final

{6,7,11,12}

We {2,8} tutorcs

{5,6,7,10,11,12}

Assignment Project

Email: tutorcs@163

QQ: 749389476

<https://tutorcs.com>

{3,4,9}

{3,4,9}

{2,8}

{2,8}

{2,8}

Help

程序代写代做 CS编程辅导



b

Start

{1,2,8}

{3,4,9}

{2,8}

We {2,8}

tutorcs

{3,4,9}

{3,4,9}

{5,6,7,10,11,12}

{2,8}

Final

{5,6,7,10,11,12}

Assignment Project

Help

{6,7,11,12}

{2,8}

Final

{6,7,11,12}

Email: tutorcs@163

{6,7,11,12}

{2,8}

QQ: 749389476

<https://tutorcs.com>

DFA

		a	b
Start	1	1	2
	2	3	1
Final	3	3	1

程序代写代做 CS编程辅导



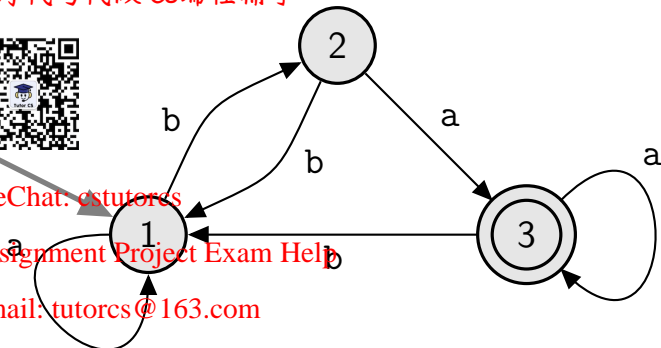
WeChat: cstutors

Assignment Project Exam Help

Email: tutorcs@163.com

QQ: 749389476

<https://tutorcs.com>



Minimum DFA

Algorithm: FA simplification

程序代写代做 CS编程辅导

Input: a FA

Colour all Final States with one colour, all non-Final States with a different colour.

repeat

for *each colour used so far* **do**

Consider all states with that colour.

if *their rows do not have the same pattern of colours* **then**

*/** States with different colour patterns along their rows must get different colours. So ... */*

Give each different row pattern a different colour, using new colours.

*/** Each set of states having the same row pattern gets the colour for that row pattern. */*

until *no new colour has been added in this iteration;*

Give each colour a unique number, and use these numbers to form the transition table.

Output: an equivalent FA with fewest states



WeChat: cstutorcs

Assignment Project Exam Help

Email: tutores@163.com

QQ: 749389476

https://tutres.com

Other Algorithms

程序代写代做 CS编程辅导



There are algorithms that can take a regular expression and produce a minimum state DFA without constructing a NFA.

WeChat: cstutorcs

There are algorithms that produce a fast and more compact representations of a DFA transition table than the straightforward two-dimensional table.

Email: tutorcs@163.com

QQ: 749389476

<https://tutorcs.com>

Programming Finite Automata

程序代写代做 CS编程辅导



Once we have a Finite Automaton for a regular expression,
we can write a program for it. WeChat: [cstutorcs](#)

Assignment Project Exam Help

Email: tutorcs@163.com

QQ: 749389476

<https://tutorcs.com>

Algorithm: detecting strings accepted by an FA given by a table.

Input: a string

currentState := 1;

table := table with rows (1,2) (3,1):

nextLetter := next character of

while nextLetter exists **do**

switch nextLetter **do**

case 'a' **do**

 currentState := table[currentState][0];

break;

case 'b' **do**

 currentState := table[currentState][1];

break;

 nextLetter := next character of input;

if currentState == 3 **then** print "Match"; **else** print "No Match";

程序代写代做 CS编程辅导



$\begin{pmatrix} 1, 2 \\ 3, 1 \\ 3, 1 \end{pmatrix}$

WeChat: cstutorcs

Assignment Project Exam Help

Email: tutorcs@163.com

QQ: 749889476

https://tutorcs.com

Application: Lexical Analysis

Many situations require text to be divided into strings that match various *patterns*.

Calculator: `exp(sqrt(-159265))+1`

Programming languages:



`read(n); sum := 0; i := 1; while(i <= n) {sum += 1.0/i; i++} write(sum).`

Assignment Project Exam Help

Personnel records:

`//Employer: Harvard College Observatory.// Annie Jump Cannon,
11/12/1863. Williamina Paton Stevens Fleming, 15/5/1857. Henrietta
Swan Leavitt 4/7/1868. Edward Charles Pickering, 19/7/1846.`

WeChat: estutorcs

Email: tutorcs@163.com

QQ: 749389476

<https://tutorcs.com>

Application: Lexical Analysis

Many situations require text to be divided into substrings that match various *patterns*.

Calculator:

exp (sqrt(3.14159265) + 1



Programming languages:

```
read ( n ); sum := 0; i := 1;
while ( i <= n ) { sum += 1.0/i; i++; } write ( sum );
```

Email: tutors@163.com

Personnel records:

QQ: 749389476

```
//Employer: Harvard College Observatory.// Annie Jump Cannon ,
11/12/1863 . Williamina Paton Stevens Fleming , 15/5/1857 .
Henrietta Swan Leavitt , 4/7/1868 . Edward Charles Pickering ,
19/7/1846 .
```

<https://tutors.com>

Terminology

程序代写代做 CS编程辅导

A **pattern** is specified by a regular expression.



A **token** is a name of a pattern.

- ▶ It may also have an attribute value associated with it.

WeChat: tutorcs

A **lexeme** is a sequence of characters that matches the pattern corresponding to a token.

Assignment Project Exam Help

Email: tutorcs@163.com

So a pattern describes the form that the lexemes of a token may take.

QQ: 749389476

<https://tutorcs.com>

Lexical Analyzer

程序代写代做 CS编程辅导



A **lexical analyzer**:

- ▶ reads the input one character at a time, and
- ▶ splits the input up into lexemes with their associated tokens,
- ▶ where each token corresponds to a specific regular language.
- ▶ It outputs a sequence of tokens (along with any attribute values that any tokens have).
- ▶ It is implemented using a Finite Automaton.

WeChat: cstutorcs

Assignment Project Exam Help

Email: tutorcs@163.com

QQ: 749389476

<https://tutorcs.com>

Matching Regular Expressions

程序代写代做 CS编程辅导



Write a program which reads in a character string, over alphabet $\{a,b\}$, one character at a time and identifies whether or not the string matches one the following regular expressions, and which one.

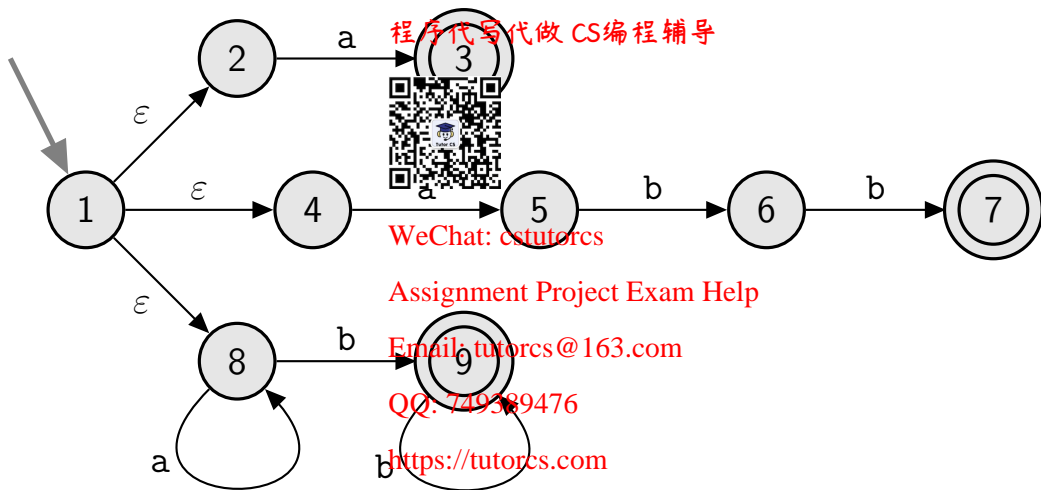
WeChat: cstutorcs

Assignment Project Exam Help

Email: tutorcs@163.com

QQ: 749389476

<https://tutorcs.com>



程序代写代做 CS编程辅导

Convert to FA, giving:



	a	b
Start	{2,4,8}	{3,5,8}
Final a	{3,5,8}	{6,9}
Final a*b	{8}	{9}
Final a*bb	{8}	{9}
Final abb	{7,9}	{9}
	\emptyset	\emptyset

QQ: 749389476

<https://tutorcs.com>

Conventions

程序代写代做 CS编程辅导

Often it is possible to split a sequence of characters up into tokens in more than one way.



- ▶ Consider `abbbb`
- ▶ **Convention:** Match the largest possible lexeme at each stage.

WeChat: cstutorcs

Often a sequence of characters can match more than one token.

Assignment Project Exam Help

- ▶ Consider `abb`
- ▶ **Convention:** If the lexemes are the same length, choose the first token that is listed.

Email: tutorcs@163.com

QQ: 749589476

<https://tutorcs.com>

Revision

程序代写代做 CS编程辅导



- ▶ Know how to find the DFA with the minimum number of states
- ▶ Know how to implement a finite automaton.
- ▶ Understand what a lexical analyzer does.

WeChat: cstutorcs

Assignment Project Exam Help

Email: tutorcs@163.com

QQ: 749389476

<https://tutorcs.com>