

BU CS 332 – Theory of Computation

Lecture 6: Assignment Project Exam Help

Reading:

Sipser Ch 1.3

“Myhill-Nerode” note

Sipser Ch 1.4 (optional)

- Regexes = NFAs

- Non-regular languages

<https://powcoder.com>
Add WeChat powcoder

Mark Bun

February 10, 2021

Regular Expressions – Syntax

A regular expression R is defined recursively using the following rules:

1. ϵ , \emptyset , and a are regular expressions for every $a \in \Sigma$
<https://powcoder.com>
2. If R_1 and R_2 are regular expressions, then so are
 $(R_1 \cup R_2)$, $(R_1 \circ R_2)$, and (R_1^*)
Add WeChat powcoder

Examples: (over $\Sigma = \{a, b, c\}$) (with simplified notation)
 ab $ab^*c \cup (a^*b)^*$ \emptyset

Regular Expressions – Semantics

$L(R)$ = the language a regular expression describes

1. $L(\emptyset) = \emptyset$
2. $L(\epsilon) = \{\epsilon\}$
3. $L(a) = \{a\}$ for every $a \in \Sigma$
4. $L((R_1 \cup R_2)) = L(R_1) \cup L(R_2)$
5. $L((R_1 \circ R_2)) = L(R_1) \circ L(R_2)$
6. $L((R_1^*)) = (L(R_1))^*$

Example: $L(a^*b^*) = \{a^m b^n \mid m, n \geq 0\}$

Regular Expressions Describe Regular Languages

Theorem: A language A is regular if and only if it is described by a regular expression

Theorem 1: Every regular expression has an equivalent NFA

<https://powcoder.com>

Theorem 2: Every NFA has an equivalent regular expression

Add WeChat powcoder

Regular expression \rightarrow NFA

Theorem 1: Every regex has an equivalent NFA

Proof: Induction on size of a regex

Base cases: **Assignment Project Exam Help**

$$R = \emptyset$$

<https://powcoder.com>

Add WeChat powcoder

$$R = \varepsilon$$

$$R = a$$

Regular expression \rightarrow NFA

Theorem 1: Every regex has an equivalent NFA

Proof: Induction on size of a regex

Inductive step: **Assignment Project Exam Help**

$$R = (R_1 \cup R_2)$$

Add WeChat powcoder

$$R = (R_1 R_2)$$

$$R = (R_1^*)$$

Example

Convert $(1(0 \cup 1))^*$ to an NFA

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

Regular Expressions Describe Regular Languages

Theorem: A language A is regular if and only if it is described by a regular expression

Theorem 1: Every regular expression has an equivalent NFA

<https://powcoder.com>

Theorem 2: Every NFA has an equivalent regular expression

Add WeChat powcoder

NFA \rightarrow Regular expression

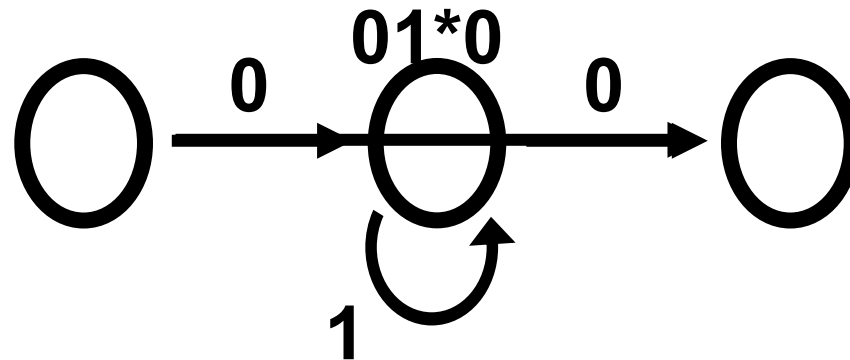
Theorem 2: Every NFA has an equivalent regex

Proof idea: Simplify NFA by “ripping out” states one at a time and replacing with regexes

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder



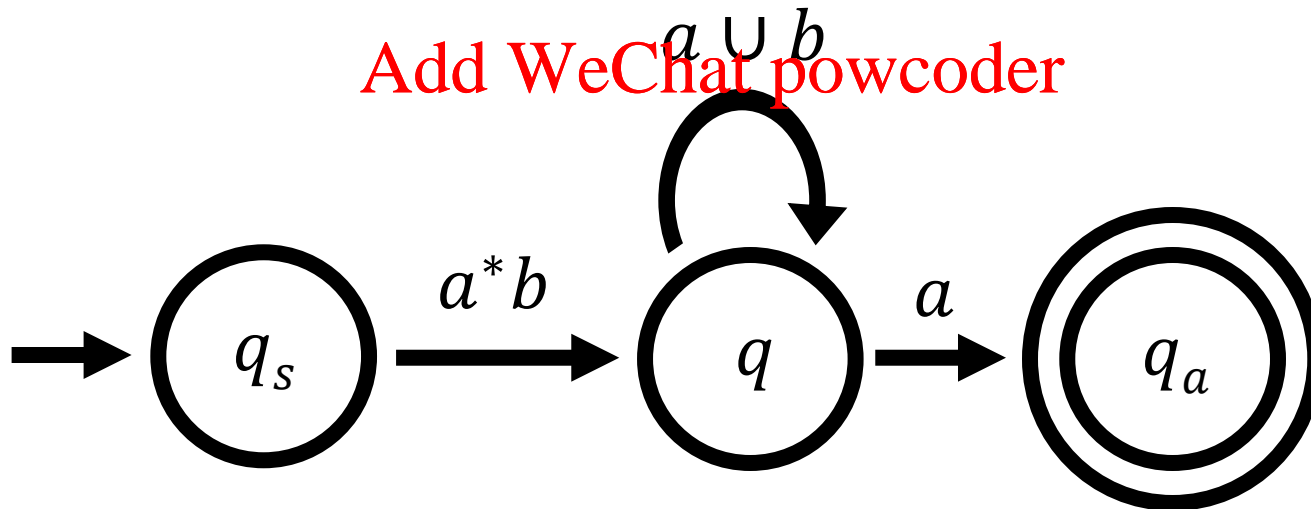
Generalized NFAs

- **Every transition is labeled by a regex**
- One start state with only outgoing transitions
- Only one accept state with only incoming transitions
- Start state and accept state are distinct

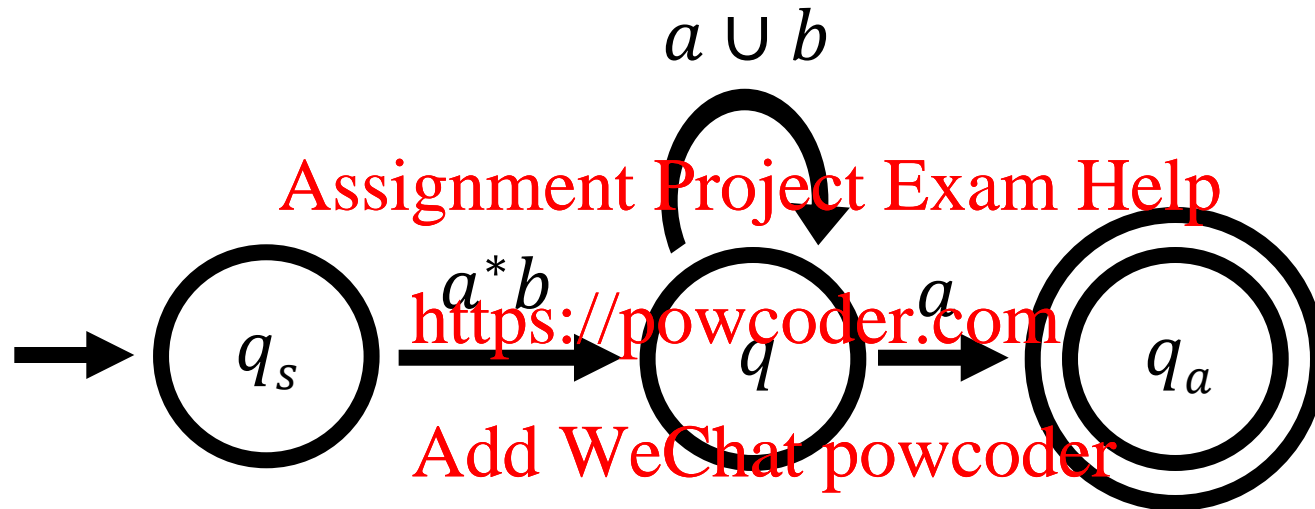
Assignment Project Exam Help

<https://powcoder.com>

Add WeChat $a \cup b$ powcoder



Generalized NFA Example

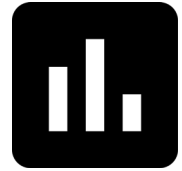


$$R(q_s, q) =$$

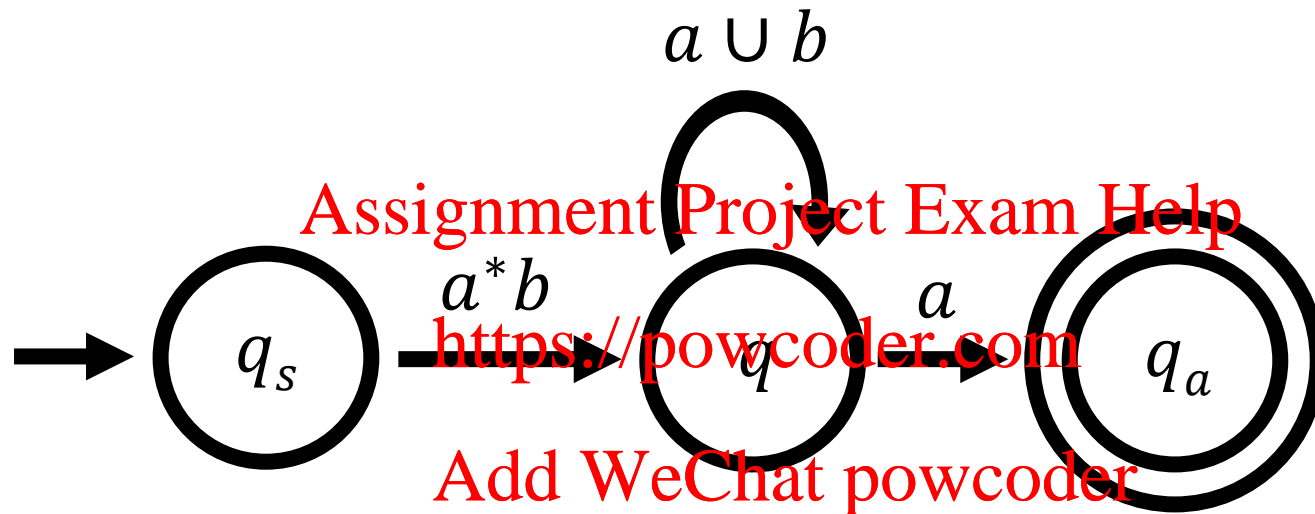
$$R(q_a, q) =$$

$$R(q, q_s) =$$

Which of these strings is accepted?

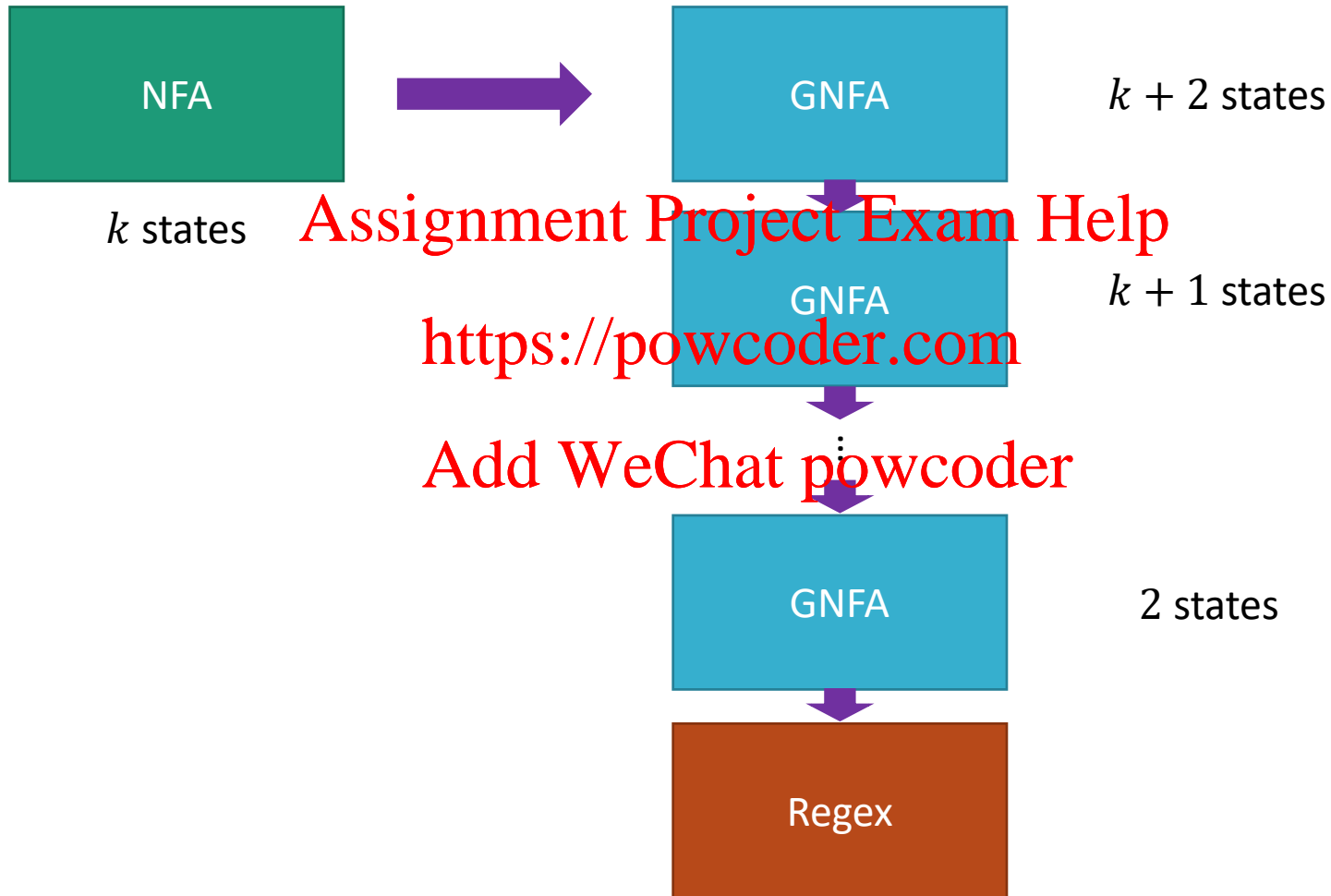


Which of the following strings is accepted by this GNFA?

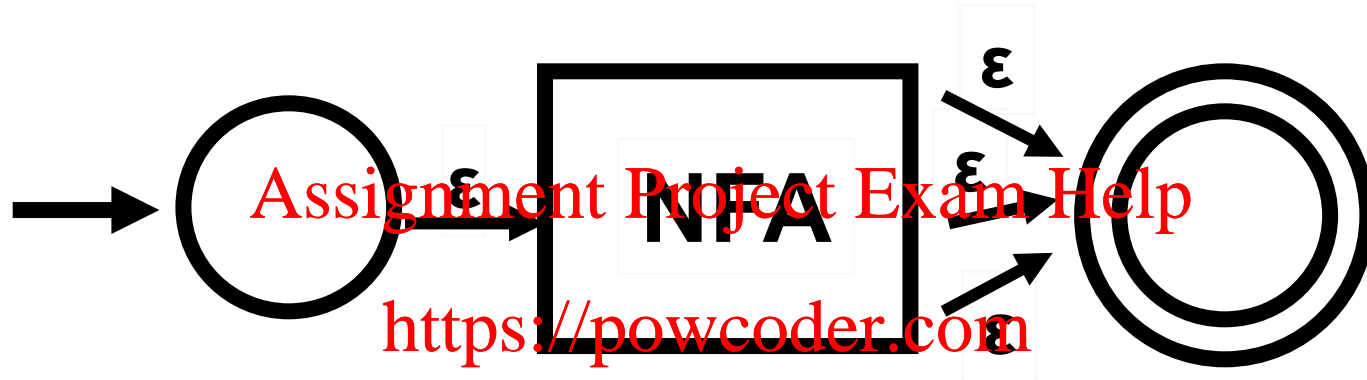


- a) aaa
- b) $aabb$
- c) bbb
- d) bba

NFA \rightarrow Regular expression



NFA \rightarrow GNFA

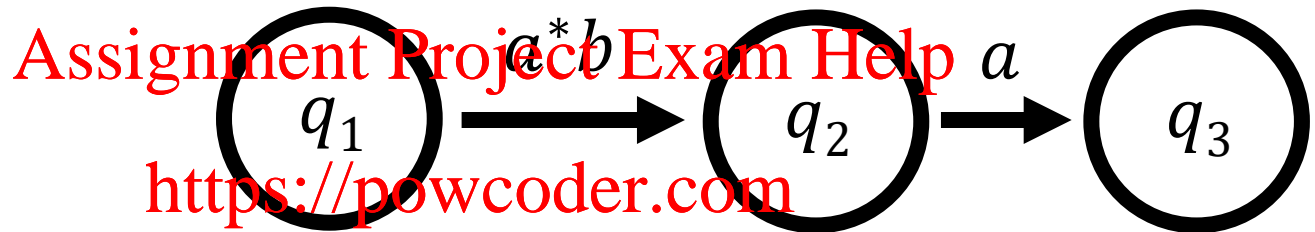


Add WeChat powcoder

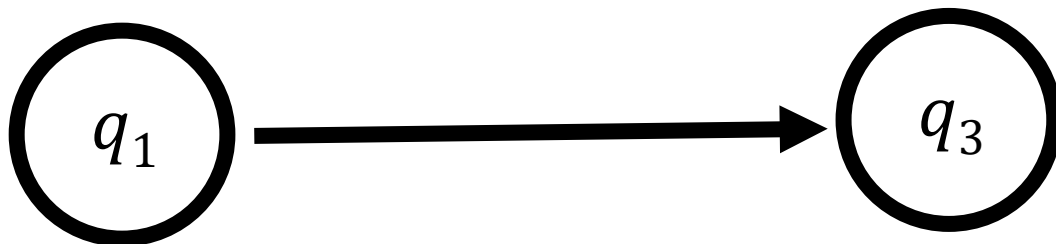
- Add a new start state with no incoming arrows.
- Make a unique accept state with no outgoing arrows.

GNFA \rightarrow Regular expression

Idea: While the machine has more than 2 states, rip one out and relabel the arrows with regexes to account for the missing state



Add WeChat powcoder



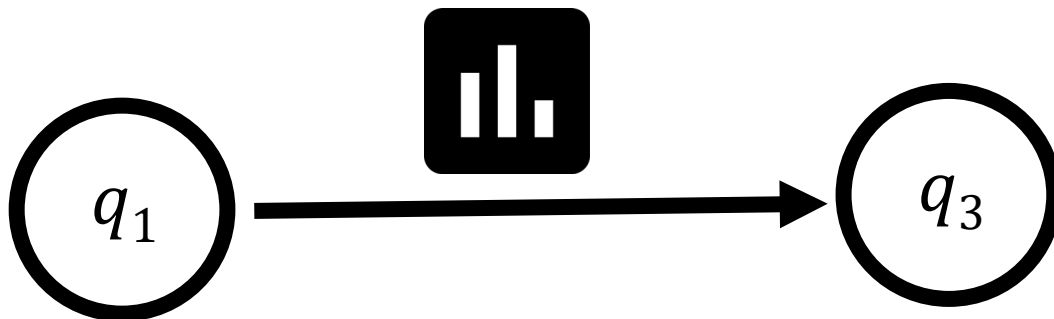
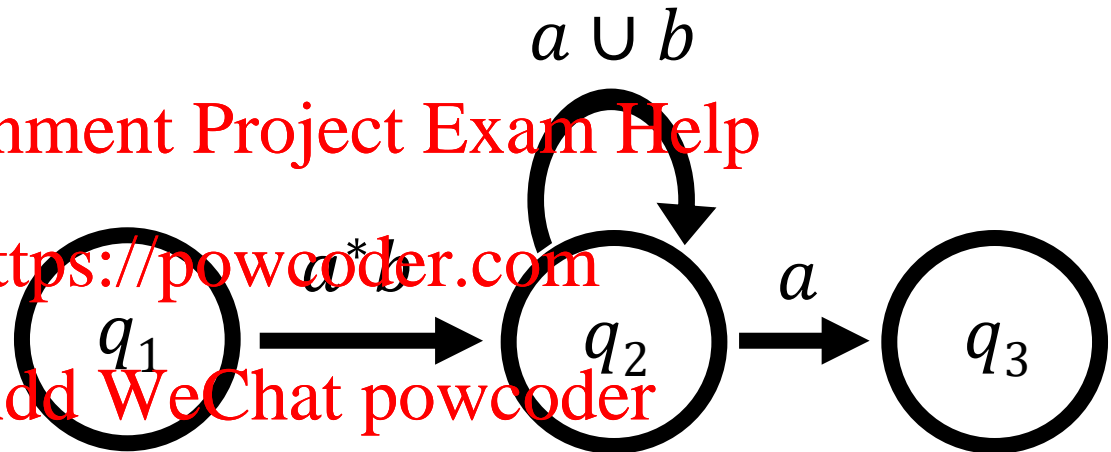
GNFA \rightarrow Regular expression

Idea: While the machine has more than 2 states, rip one out and relabel the arrows with regexes to account for the missing state

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder



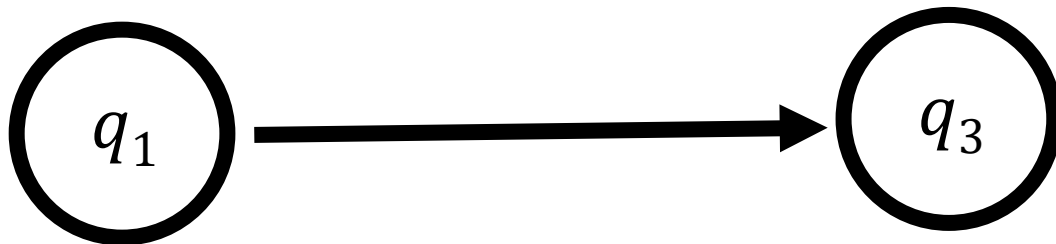
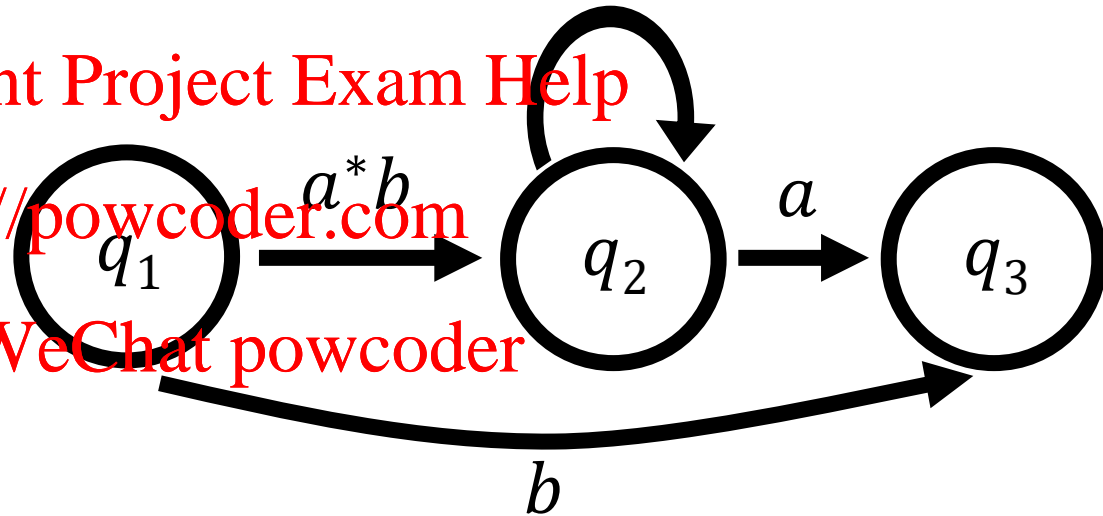
GNFA \rightarrow Regular expression

Idea: While the machine has more than 2 states, rip one out and relabel the arrows with regexes to account for the missing state

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder



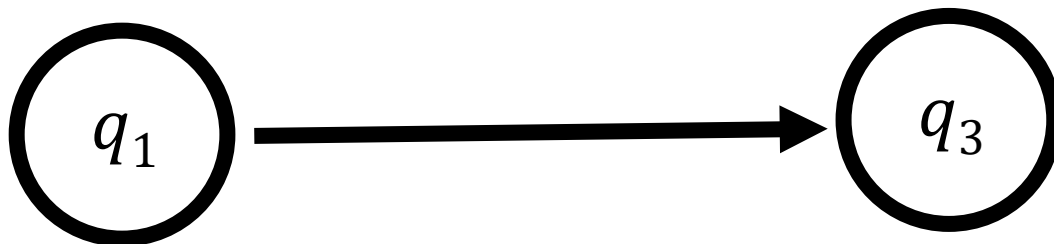
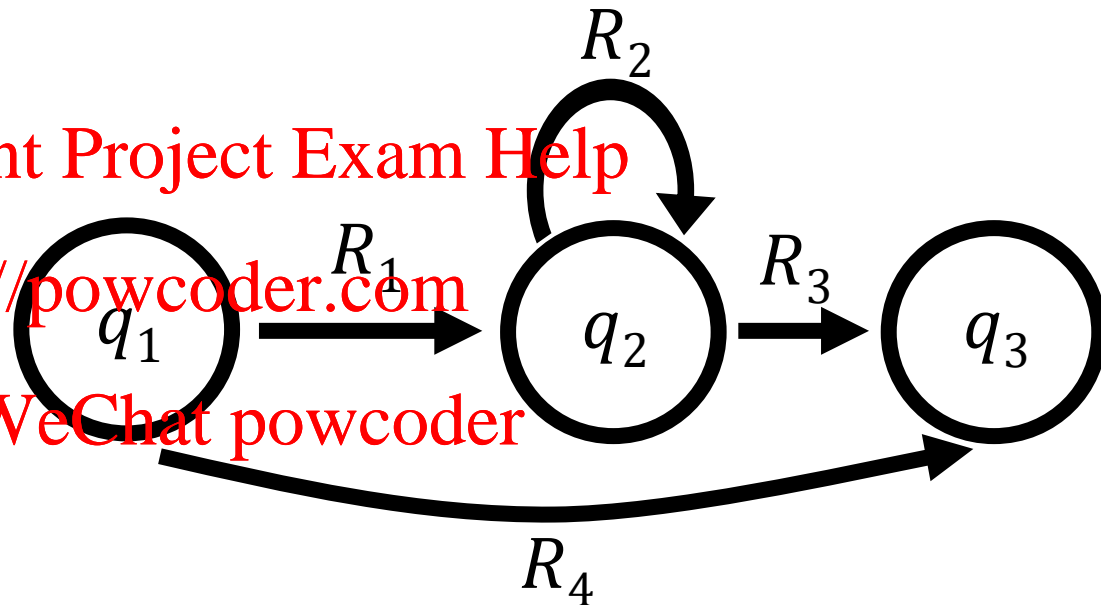
GNFA \rightarrow Regular expression

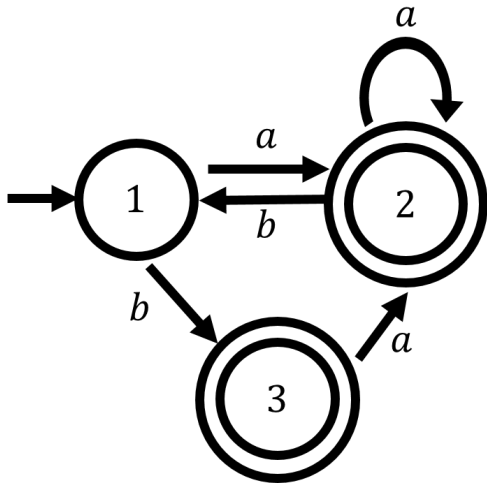
Idea: While the machine has more than 2 states, rip one out and relabel the arrows with regexes to account for the missing state

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder





Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

Non-Regular Languages

Assignment Project Exam Help
<https://powcoder.com>

Add WeChat powcoder

Motivating Questions

- We've seen techniques for showing that languages are regular

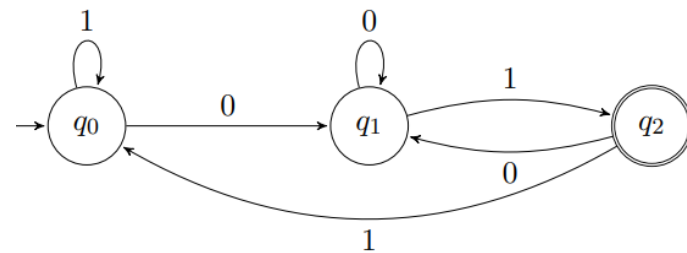
Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

- How can we tell if we've found the smallest DFA recognizing a language?
- Are all languages regular? How can we prove that a language is not regular?

An Example



$$A = \{ w \in \{0, 1\}^* \mid w \text{ ends with } 01 \}$$

Claim: Every DFA recognizing A needs at least 3 states

Proof: Let M be any DFA recognizing A . Consider running M on each of $x = \varepsilon, y = 0, w = 01$

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

A General Technique

$$A = \{w \in \{0, 1\}^* \mid w \text{ ends with } 01\}$$

Definition: Strings x and y are **distinguishable** by L if there exists z such that exactly one of xz or yz is in L .

Ex. $x = \varepsilon$, $y = 0$ **Assignment Project Exam Help**

<https://powcoder.com>

Definition: A set of strings S is **pairwise distinguishable** by L if every pair of distinct strings $x, y \in S$ is distinguishable by L .

Ex. $S = \{\varepsilon, 0, 01\}$

A General Technique

Theorem: If S is pairwise distinguishable by L , then every DFA recognizing L needs at least $|S|$ states

Proof: Let M be a DFA with $< |S|$ states. By the pigeonhole principle, there are $x, y \in S$ such that M ends up in same state on x and y

<https://powcoder.com>

Add WeChat powcoder

Back to Our Example

$$A = \{ w \in \{0, 1\}^* \mid w \text{ ends with } 01 \}$$

Theorem: If S is pairwise distinguishable by L , then every DFA recognizing L needs at least $|S|$ states

Assignment Project Exam Help

$$S = \{\varepsilon, 0, 01\} \quad \text{https://powcoder.com}$$

Add WeChat powcoder

Another Example

$$B = \{ w \in \{0, 1\}^* \mid |w| = 2 \}$$

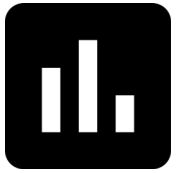
Theorem: If S is pairwise distinguishable by L , then every DFA recognizing L needs at least $|S|$ states

Assignment Project Exam Help

$S = \{$ <https://powcoder.com>

Add WeChat powcoder

Distinguishing Extension



Which of the following is a distinguishing extension for $x = 0$ and $y = 00$ for language $B = \{w \in \{0, 1\}^* \mid |w| = 2\}$?

- a) $z = \varepsilon$
- b) $z = 0$
- c) $z = 1$
- d) $z = 00$

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

Non-Regularity

Theorem: If S is pairwise distinguishable by L , then every DFA recognizing L needs at least $|S|$ states

Corollary: If S is an infinite set that is pairwise distinguishable by L , then no DFA recognizes L

Add WeChat powcoder

The Classic Example

Theorem: $A = \{0^n 1^n \mid n \geq 0\}$ is not regular

Proof: Construct an infinite pairwise distinguishable set

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder