Programming languages (TC-2006)

Homework 03

In this homework, you will use the strategies seen in class to design a scanner that detects valid words of simple toy languages. This homework may be solved in teams of at most three members. If you want, you can submit this homework individually.

To simplify this homework, assume that you can use DIGITS or LETTERS to refer to the digits (characters 0 to 9) or the upper case letters in the English alphabet (characters A to Z), respectively. If you need it, you can even use natural language to express a particular condition.

1 Designing a DFA - Case I (30%)

Design a DFA that recognizes the lexical elements defined by the following information.

NUMBER. At least one digit followed by zero or more digits.

VARIABLE. The symbol \$ followed by any combination of letters and digits, as long as they start with a letter.

OPERATOR. Any of the following characters: +, -, *, or /.

For simplicity, assume that you will analyze only one word at the time and that such a word contains no spaces. Then, you can safely assume that inputs such as 10+\$X will produce an error (it includes three valid words: 10, +, and \$X). However, be warned: this is not the case when designing a real scanner. Also, consider that any transition not indicated by the previous information must produce an error (and you are not requested to show it in your DFA explicitly).

Some examples to consider when designing your DFA are listed below.

Input	Output	Justification
123	Accepted	It is a valid NUMBER.
-10	Rejected	It contains two valid words: - and 10.
\$XY25	Accepted	It is a valid VARIABLE.
+	Accepted	It is a valid OPERATOR.
\$25XY	Rejected	A VARIABLE cannot contain a digit immediately after the symbol \$.

2 Designing the transition matrix - Case I (25%)

Given the DFA defined for case I, propose a feasible transition matrix for such a DFA. Remember to use indexes larger or equal to 100 to indicate goal states and the constant `ERROR` for transitions leading to the error state.

3 Designing a DFA - Case II (30%)

Design a DFA that recognizes the lexical elements defined by the following information.

NUMBER. The symbol '#' followed by any combination of digits or the letters 'a' to 'f' (hexadecimal numbers). Please note that the characters 'a' to 'f' are different from the ones defined in LETTERS, which are upper case characters.

VARIABLE. The symbol '_' followed by any combination of upper case letters and digits, as long as they start with a letter and finish with the symbol '_'.

RESERVED-WORD. Any of the following strings: "START" or "END".

As in the previous case, assume that you will analyze only one word at the time and that such a word contains no spaces. Then, you can safely assume that inputs such as START#24 will produce an error (it includes two valid words: START and #24). However, be warned: this is not the case when designing a real scanner. Also, consider that any transition not indicated by the previous information must produce an error (and you are not requested to show it in your DFA explicitly).

Some examples to consider when designing your DFA are listed below.

Input	Output	Justification
#aa5	Accepted	It is a valid NUMBER.
53	Rejected	There is no definition for anything that starts with a digit.
#b5C	Rejected	The letters in NUMBER can only be the characters a to f.
_XY25_	Accepted	It is a valid VARIABLE.
_25XZ_	Rejected	In a VARIABLE, a letter must appear immediately after the symbol _.
STAR	Rejected	Although similar to a reserved word, STAR is not defined in the vocabulary.

4   Designing the transition matrix - Case II (25%)

Given the DFA defined for case II, propose a feasible transition matrix for such a DFA. Remember to use indexes larger or equal to 100 to indicate goal states and the constant ERROR for transitions leading to the error state.

Deliverables

Prepare a brief PDF document (at most two pages) that contains the information requested and submit it to Canvas. **Please, do not submit other formats but PDF**.

I promise to apply my knowledge, strive for its development, and not use unauthorized or illegal means to complete this activity, following the Tecnológico de Monterrey Student Code of Honor.