

# Neoastra Systems | Internship | Sahil <> Assignment

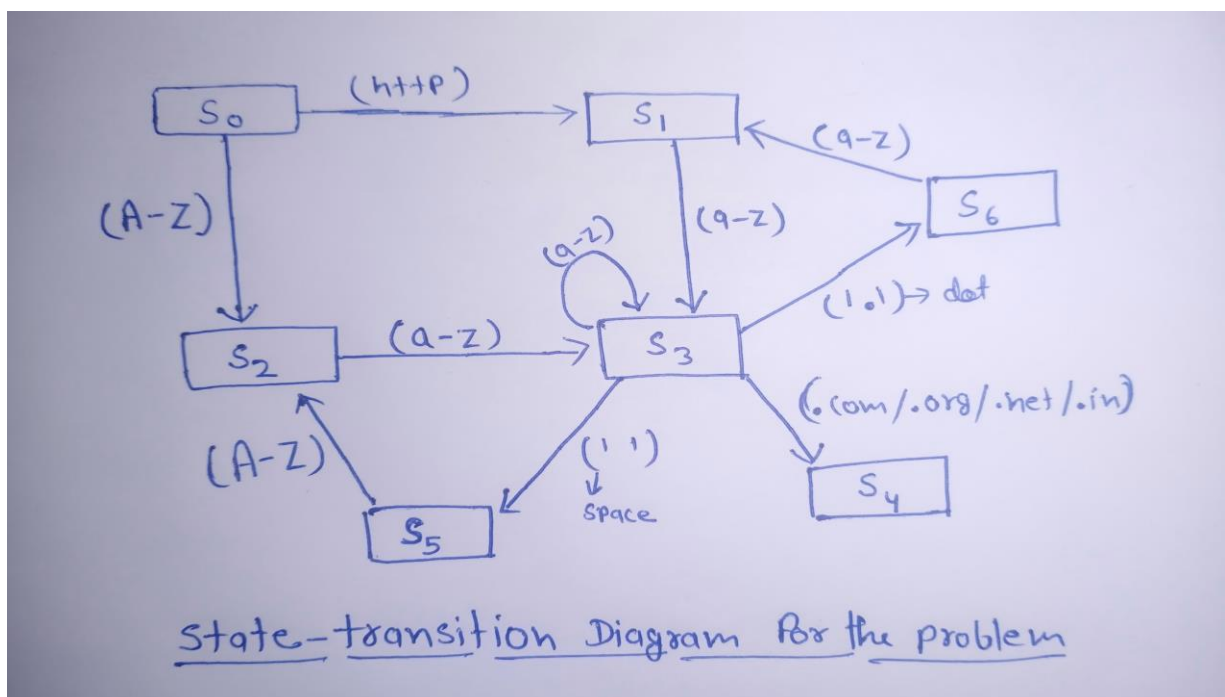
Summary of the assignment solution:-

The code is written in the C++ programming language and the code file is attached with this document.

This program takes an input string and uses finite-state parsing to output whether the given string is a name or web URL or neither of them.

There are some constraints for this solution which are given below:-

1. The name starts with a capital letter character and can have space in between them and again starting with a capital letter. For ex.- "Sahil Kumar" is a valid name but "sahil Kumar" and "sAhil kumar" and "Sahil kumar" are invalid.
2. The web URL starts with <https://> or <http://> only and can have small alphabetical characters, dots, and not any other characters. For ex.- <https://erp.iitkgp.ac.in> is valid but <https://www.NEOastra.com> is invalid
3. There are currently only 5 ending domains are valid (.com, .org, .net, .gov, .in)



states/Inputs	https	(A-Z)	(a-z)	.com	(' ')	('.')	Exception
state-0	S <sub>1</sub>	S <sub>2</sub>	-	-	-	-	-
state-1	-	-	S <sub>3</sub>	-	-	-	-
state-2	-	-	S <sub>3</sub>	-	-	-	-
state-3	-	-	S <sub>3</sub>	S <sub>4</sub>	S <sub>5</sub>	S <sub>6</sub>	-
state-4	-	-	-	-	-	-	-
state-5	-	S <sub>2</sub>	-	-	-	-	-
state-6	-	-	S <sub>1</sub>	-	-	-	-
state-7	-	-	-	-	-	-	-

# Neoastra Systems | Internship | Sahil <> Assignment

Here is the description of the vocabulary used in the state transition diagram

There are 8 states and 8 inputs values, in the program, it is implemented as a 2-D vector of integers. The corresponding integer values are given below :-

States:-

- 0 --> The initial state
- 1 --> This is the state just after getting http:// or https://
- 2 --> This is the state just after getting a capital letter character
- 3 --> This is the state just after getting a small letter character
- 4 --> This is the state after getting ".com" or ".in" or ".org"
- 5 --> This is the state just after getting a white space
- 6 --> This is the state just after getting a dot
- 7 --> This is for exception handling (Once we are in this state there is no way to come out)

Inputs:-

- 0 --> Initially input is 0
- 1 --> If http:// or https:// shows in the string
- 2 --> If there is a Capital letter character in the string
- 3 --> If a Small letter character shows
- 4 --> If .com or .in or .org shows in the string
- 5 --> If white space is there
- 6 --> If a '.' is there (there may be a dot in between the URLs for ex. <https://erp.iitkgp.ac.in>)
- 7 --> This is for exception handling, if any other character shows Input is 7.