CPT 411 Assignment

2023/2024

Write a well-structured, well-documented recognizer Deterministic Finite Automata (DFA) for the assigned language. The program must be based on a complete DFA for the language, but you can terminate the program on entering a trap state. You MUST process one character at a time from left to right simulating a finite state machine. No other strategy for your program is allowed.

Each student will be given a different language to search for. For program development, a text file will be given as example for implementation and testing. During the demonstration day, you may use the given sample text or new text file to demo your machine.

For demonstration purpose, the output from the run must show the following:

 The pattern (input string).

The text used for demo.

The status (whether accept/reject).

Additional information (the position of the pattern found, occurrences of patterns, visualization using boldface of the pattern occurred in the text etc.).

You also hand in a typed technical report on your project. Recommended outline:

I. Introduction – state your language, define your scope and give the complete DFA (sample DFA if the complete one is too huge).

II. Implementation Information

        a. how your read and processed the strings

        b. overview of programming constructs used for your program

III. Conclusion – Summary

IV. Appendix – Sample/Full programs.

Σ = { a,..z, A,..Z, 0,…9, and other symbols found the sample text}

Example languages

L = {w ∈ Σ * | w contain substring "Malaysia", "Kuala Lumpur", "Penang" …}

L = {w ∈ Σ * | w contain substring "2 litres", "1kg", "100%" …}

Assignment Problems

L1. Place Finder (E.g., Country, Organization, Shops, States etc.)

Example: Malaysia, Australia, Penang, Pizza Hut, Intel etc.

L2. Number Data Finder

Example: 3rd, 100%, 3 million, 2017, 16 September 2016, 5 litres, 2 cups etc.

L3. People Finder

Example: John, Mr. Lim, Ahmad, Dr Tan etc.

L4. Food Finder

Example: roti canai, laksa, chapati etc.

L5. Conference/Workshop Finder

Example: CITA 2021, WI 2021 etc.

L6. English Conjunctions/Adverb/Adjectives Finder

Example: and, most, good, bad, pretty, dirty, blue etc.

L7. English Stop Words Finder

Example: the, or, it etc.

Sample Text I (You may use other texts to demo)

Malaysia

From Wikipedia, the free encyclopedia

Malaysia is a federal constitutional monarchy located in Southeast Asia. It consists of thirteen states and three federal territories and has a total landmass of 329,847 square kilometres (127,350 sq mi) separated by the South China Sea into two similarly sized regions, Peninsular Malaysia and East Malaysia (Malaysian Borneo). Peninsular Malaysia shares a land and maritime border with Thailand and maritime borders with Singapore, Vietnam, and Indonesia. East Malaysia shares land and maritime borders with Brunei and Indonesia and a maritime border with the Philippines. The capital city is Kuala Lumpur, while Putrajaya is the seat of the federal government. By 2015, with a population of over 30 million, Malaysia became 43rd most populous country in the world. The southernmost point of continental Eurasia, Tanjung Piai, is in Malaysia, located in the tropics. It is one of 17 megadiverse countries on earth, with large numbers of endemic species.

Malaysia has its origins in the Malay kingdoms present in the area which, from the 18th century, became subject to the British Empire. The first British territories were known as the Straits Settlements, whose establishment was followed by the Malay kingdoms becoming British protectorates. The territories on Peninsular Malaysia were first unified as the Malayan Union in 1946. Malaya was restructured as the Federation of Malaya in 1948, and achieved independence on 31 August 1957. Malaya united with North Borneo, Sarawak, and Singapore on 16 September 1963, with is being added to give the new country the name Malaysia. Less than two years later in 1965, Singapore was expelled from the federation.

Since its independence, Malaysia has had one of the best economic records in Asia, with its GDP growing at an average of 6.5% per annum for almost 50 years. The economy has traditionally been fuelled by its natural resources, but is expanding in the sectors of science, tourism, commerce and medical tourism. Today, Malaysia has a newly industrialised market economy, ranked third largest in Southeast Asia and 29th largest in the world. It is a founding member of the Association of Southeast Asian Nations, the East Asia Summit and the Organisation of Islamic Cooperation, and a member of Asia-Pacific Economic Cooperation, the Commonwealth of Nations, and the Non-Aligned Movement.

Sample Text II

The 3rd International Workshop on Machine Learning and Knowledge Graphs (MLKgraphs2021)
September 27 - 30, 2021 - Linz, Austria

http://www.dexa.org/mlkgraphs2021
email: dexa@iiwas.org
Papers submission: https://easychair.org/conferences/?conf=mlkgraphs2021

**** IMPORTANT DATES ****
Paper submission: April 23, 2021 (SHARP)
Notification of acceptance: June 1, 2021

Camera-ready copies due: June 30, 2021

*** PUBLICATION ***
All accepted papers will be published by Springer in "Communications in Computer and Information Science".

*** SCOPE ***
Knowledge Graphs are becoming a key technology for large-scale information processing systems containing massive collections of interrelated facts. Specifically, Knowledge Graphs provide the means for development of the newest data methods for data management, data fusion, data merging, and graph optimization and modeling, serving as a source of high quality data and a base for web-scale information integration.

The 3rd International Workshop on Machine Learning and Knowledge Graphs aims to be a meeting point for researchers and practitioners working on the latest advances in the intersection of machine learning technologies and knowledge graphs. Therefore, we welcome submissions of novel research that brings together the two topics of Machine Learning (ML) and Knowledge Graphs (KGs) either applying ML models for semantic data management structures (like KGs or ontologies), or by presenting newly assembled Knowledge Graphs that support the task of Machine Learning for certain application domains. Examples areas are Business Analytics, Customer Relationship Management, Fault Detection, Industry 4.0, or Social Networking.