Text

Description automatically generated

SEMESTER 1, 2021/2022

CSCI 4342 NLP Section 01

NATURAL LANGUAGE PROCESSING

**Exercise 2**

**PREPARED BY:**

|  |  |
| --- | --- |
| **NAME** | **MATRIC NO.** |
| MOHAMED MOUBARAK MOHAMED MISBAHOU MKOUBOI | 1820705 |

LECTURER:

Dr. SURIANI BT. SULAIMAN

1. Find the shortest string that is NOT in the language represented by the regular expression w\*(wa)\*a\*.

It is clear that β, w, and a are strings in the language with lengths of one or less. The language has the strings ww, aa, and ab with lengths of two. aw, on the other hand, isn't in it. As a result, the answer is aw.

1. Find a regular expression corresponding to the language of all strings over the alphabet { w, a } that contains exactly two a’s (refer slide #5).

In this language, a string must include at least two w's. Because any string of a's may be placed in front of the first w, behind the second w, or between the two w's, and any string of a's can be represented by the regular expression a\*, a\*w a\*w a\* is a regular expression for this language.

1. Find a regular expression corresponding to the language of all strings over the alphabet { w, a } that do not end with wa (refer slide #5).

Any string in a language that is longer than w, a must finish with an or a. As a result, if a string does not conclude with wa, it must end with w, and if it ends with a, the last a must be preceded by the symbol a. (w + a)\*(w + aa) is a regular expression for the language because it can have any string in front of the last an or aa.