AUTOMATA THEORY AND FORMAL LANGUAGES

-

CODING ASSIGNMENT

Yasemin İzel Çangal - 041701021

Ertuğrul Yılmaz - 041701030

Introduction

In this assignment, we are expected to write a decryption program. Program takes an encrypted .txt file and gives the decrypted version of it. Program consists of two main algorithms. First is algorithm implements more than one regex function to every line of the original.txt file and write it to a .txt file named revert.txt . Second part is implementing of context free grammar to revert.txt file to avoid fake lines from our file. Algorithm applies the context free grammar and decides which line is fake then writes the non-fake lines to a .text file named decrypted.txt . After running the program the decrypted content can seen in the decrypted.txt .

Regex Functions

To make alterations, we wrote functions for every different alteration methods by using regex library. Explanation of the functions:

- date(): Function takes the date if it is in right form. Then changes the digit location to right locations.
  Ex: 17-01-2021. -> 10-12-2017.
- extension(): It changes encrypted file extension to right file extensions.
  Ex: abc.cat -> abc.txt or abc.mouse -> abc.docx or abc.dog -> abc.pdf
- phoneNumber(): In the encrypted file the phone number are in the form of IBAN TR...
  It extracts the phone number from the proper form.
  Ex: TR00 4560 1230 7890 0000 0000 00. -> Phone number : 123-456-0789.
- verb(): This function includes some alterations about verbs. In encrypted file verbs can be in different forms. Function them to proper verb form.
  Ex: I forwarded -> I won't be able to forward
- time(): It decrypts time to proper form by changing digit location to right places.
  Ex: at 16:09. -> at 19:06.
- passives_to_pural(): This function makes a complex sentences alteration. In encrypted files, sentences can be in very different forms to detect. Function detects that form and changes it to right form.
  Ex: Cat is looked by human. -> Cat and human are looking.

Context Free Grammar

After applying alteration by using regex then the algorithm can eliminate the fake lines from our file. To achieve that, we created a context free grammar according to our decryption method. Context free grammar can be seen below:

S->”I”J | “We”P | “She”O | “He”O | “It”O | “They”P | “You”P |abcH | abcdH | CatO

A -> won’t be able toB | couldB |won’tB | mightB | mayB | andL

J -> amG | A

O -> isG | A

P -> areG | A

L -> dog areK | human areK

K -> looking.D

G -> wanted C | looked C | forwarded C |

waited C

B -> wantC | lookC | forwardC | waitC

C -> to Phone number : 531-357-8664. | at 10/12/2017.| at 19:06. | abc.txt. |abc.docx. |abc.pdf. | abcd.txt. |abcd.docx. |abcd.pdf. | .D

D -> S | ε

With this context free grammar, we are capable of detecting the lines which are against our language.

```
1 I forwarded to IBAN: TR00 4560 1230 7890 0000 0000 00. I couldn't wait at 16:09.
2 I may not look abc.cat. I will forward abcd.mouse.
3 I might look abbbc.cat. I have forwarded at 17-01-2021.
4 He is bored.
5 Cat is looked by human.
6 We are waited.
7 Dog is looked by human.
8 She waited at 17-01-2021.
9 You cleaned at 16:09.
```

Figure 1: Input file

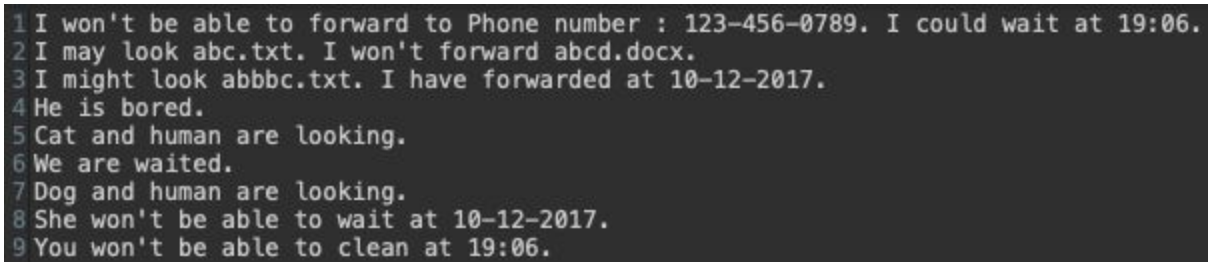

Figure 2: Intermediate file

Figure 3: Output file

Conclusion

In conclusion, the program decrypts the text from the given original.txt file. It uses some alterations to see the right form of the sentences then inserts the text file to context free to detect the fake line in the file. Finally the program gives the fully decrypted file. Decrypted file can reached from decrypted.txt