

## HW#4 RE Programming (COSE215, 2023 Spring)

### Problem 1: Bacon's Cipher

Your task is to create a Regular Expression program with **re(python)** to **decode a Bacon Cipher code** into a text message. You will receive a lengthy Bacon ciphertext, which translates to the English alphabet. After proper conversion, you will obtain a URL leading to Problems 2 and 3. The separator character is a space, and any character other than capital A or B between A and B should be removed. If you find a **URL** format, eliminate spaces and convert it to lowercase.

Input::

```
"ABstartBBB roAA4AAB leAB3BA2A 2'B2AfindA3BA ABABB iABBAB k AABBB tBA://ABB
ps://wBABBA BA.noBBA .noBtAeABB ioABBAB .so/cAAABA s-AAAAB inaBAAAB y/AABBB
BABBA -4-6AABAB 0AAA.AA 449A-ABAA 2836487AA.BAA 89297AAfindAAB AAlinkABB
AAreallyABA 5f67AAAAB 014
ABBBB BAA.AB obAB-ABB em3'BAaaaABA ABABB iABBAB k AAhBBB BAttpABB BAA-BB
ABBBB s://BA5B1BA BA-BBA BA22BBA .ABB-AB ABBBA BAABB ABAAA ABBBA
ABBAB .sABBBA /cAAABA s-AAgoodAAB inAA-AAA B-AAAB y/AA-BBB BABBA -4-AAABB
354496AABAA AAAAB AAAAB 4AA3BAB 49AAABA 4AA-AAA AAfightingAAA AAyouABA
AAcanAAA 53AAtheoryAofAcomputationB AABdoAB 2e2AAitABB AA.ABB e07"
```

Bacon's Cipher table

letter	code	letter	code
A	<u>AAAAA</u>	N	<u>ABBAB</u>
B	<u>AAAAB</u>	O	<u>ABBBA</u>
C	<u>AAABA</u>	P	<u>ABBBB</u>
D	<u>AAABB</u>	Q	<u>BAAAA</u>
E	<u>AABAA</u>	R	<u>BAAAB</u>
F	<u>AABAB</u>	S	<u>BAABA</u>
G	<u>AABBA</u>	T	<u>BAABB</u>
H	<u>AABBB</u>	U	<u>BABAA</u>
I	<u>ABAAA</u>	V	<u>BABAB</u>
J	<u>ABAAB</u>	W	<u>BABBA</u>
K	<u>ABABA</u>	X	<u>BABBB</u>
L	<u>ABABB</u>	Y	<u>BBAAA</u>
M	<u>ABBAA</u>	Z	<u>BBAAB</u>

Ex) "BABBA AABthisAA ABisABB AfakeAhttp:ABA AB//BBA ABexampleBAA A.ABAA BAACoBB  
AmBBBA AAABA AAABA S hBAABB BAABB AB BBB BAABA ://AAABA AAABA BAABA .kABBBA  
reAAAAA .aAAABA .ABABA BAAAB

" → "W E L C O M E T O C C s h T T P S ://C C S .kO reA .aC .K R

**Output:** return all part of decoded text

You don't need to find link from decoded text with regex.

**Remember to turn all Capital(upper-case) English alphabet to small(lower-case) alphabets**

**The URL only take lower-case alphabets**

Once your program runs correctly, you will get the **URL** to the next problems(problem2 and problem3)!

### Constraints:

Use only *Python re*, *NLTK* in *Python3*

### Submission

1. Submit only a .py file.
2. Include a report containing the environment, algorithm, answers to each problem, regular expressions, and an explanation of your code.
3. Late submissions will have a one-day 10% deduction, a one-week 20% deduction, and no late submissions will be allowed after one week.

