

## BÀI THỰC HÀNH SỐ 4

# TỔNG HỢP CÁC BÀI TẬP

### **Buổi 1**

1. *Viết hàm hoán đổi 2 ký tự đầu tiên của 2 chuỗi đơn và sinh ra chuỗi mới từ 2 chuỗi kết quả (cách nhau bằng khoảng trắng).*

*Ví dụ:*

*Input: 'abc' 'efg'*

*Output: 'efc abg'*

2. *Viết hàm loại bỏ các ký tự tại vị trí index chẵn của một chuỗi cho trước Ví dụ:*

*Input: 'python'*

*Output: 'yhn'*

3. *Viết hàm đếm số lần xuất hiện của các từ trong câu Ví dụ:*

*Input: 'no pain no gain'*

*Output: {'no': 2, 'pain': 1, 'gain': 1}*

4. *Viết hàm mã hóa một chuỗi bằng cách dịch chuyển các ký tự sang trái 3 bước. Ví dụ: d->a, e->b, f->c*

*Input: 'def'*

*Output: 'abc'*

5. *Viết hàm kiểm tra một chuỗi hợp lệ hay không? (có được sinh từ bộ chữ cái nhập) Ví dụ: bộ chữ cái nhập  $S = \{0,1,2\}$*

*Các chuỗi hợp lệ: 0011, 12, 012, 122...*

*Các chuỗi không hợp lệ: 123, 24,...*

6. *Viết hàm chuyển một chuỗi thành danh sách Ví dụ:*

*Input: 'This is a list'*

*Output: ['This', 'is', 'a', 'list']*

7. *Viết hàm in ra ký tự không lặp lại đầu tiên trong chuỗi Ví dụ:*

*Input: 'abcdef'*

*Output: 'a'*

*Input: 'abcabcdef'*

*Output: 'd'*

*Input: 'aabbcc'*

*Output: 'None'*

8. *Viết hàm loại bỏ khoảng trắng trong chuỗi cho trước Ví dụ:*

*Input: 'a b c'*

*Output: 'abc'*

9. *Viết hàm in ra từ được lặp lại đầu tiên trong chuỗi Ví dụ:*

*Input: 'ab ca bc ca ab bc'*

*Output: 'ca'*

*Input: 'ab ca bc ab'*

*Output: 'ab'*

10. *Viết hàm tìm độ dài tối đa của chuỗi con gồm các số 0 liên tiếp trong chuỗi nhị phân cho trước.*

*Ví dụ:*

*Input: '1110001110000'*

*Output: 4*

## BÀI TẬP NÂNG CAO

### Bài tập 1

*Kiểm tra 2 danh sách khác nhau có thể sinh ra cùng một chuỗi hay không?*

Input danh sách  $A = (s_1, s_2, \dots, s_n)$ , danh sách  $B = (t_1, t_2, \dots, t_m)$

**Output:** Có, nếu tồn tại một chuỗi được sinh ra từ danh sách A và danh sách B

*Ví dụ:*

Input:  $A = (110, 0011, 0110)$  và  $B = (110110, 00, 110)$ .

Output: có vì  $s_2s_3s_1 = t_2t_3t_1$

*Kiểm tra các trường hợp sau:*

1.  $A = (0011, 11, 1101)$ ,  $B = (101, 011, 110)$

2.  $A = (100, 0, 1)$ ,  $B = (1, 100, 0)$

*(Gợi ý: sử dụng itertools (import itertools))*

### Bài tập 2:

*Chuỗi DNA được tạo thành từ sự kết hợp của các codon, là 3 ký tự khác nhau bất kỳ thuộc bộ ký hiệu  $\{A, C, G, T\}$ , ví dụ ACT, ACG, TCG... Một đoạn gen là tập hợp của ít nhất 3 codon bắt đầu là ATG và kết thúc bằng TAA, TAG hoặc TGA.*

*Hãy viết chương trình kiểm tra 1 đoạn gen có hợp lệ hay không?*

*Ví dụ:*

Input: ATGCCCTAG

Output: không hợp lệ

Input: ATGCGTTGA

Output: hợp lệ

## **Buổi 2**

1. *Xây dựng DFA và kiểm tra chuỗi nhập từ bàn phím có thuộc DFA đã cho hay?*
2. *Dựa trên DFA đã xây dựng, mở rộng xây dựng lớp NFA và kiểm tra 1 chuỗi có thuộc ngôn ngữ sinh bởi NFA đã cho hay không*

## **Buổi 3**

### **A. *Đọc file txt, và xây dựng btcq để tìm và in:***

- 1) Các dòng bắt đầu bằng ‘t’ hoặc ‘h’, và có chứa ‘re’ (sử dụng phương thức re.match())
- 2) Các dòng có chiều dài tối thiểu 20 ký tự
- 3) Các dòng kết thúc bởi cặp dấu ‘?!’
- 4) Các dòng chứa những ký tự: a, r, s, m, l (không cần liên tục)
- 5) Nội dung file ko chứa các dấu ‘,’ và ‘.’
- 6) Các dòng có chứa chữ “mouse”
- 7) Các từ có số chữ ‘a’ bất kỳ và theo sau bởi ‘b’
- 8) Domain địa chỉ mail (ví dụ: abc@gmail.com, in gmail.com )
- 9) Nội dung giữa cặp tag <head> </head> (gợi ý: sử dụng open và read để mở và đọc file)

### **B. *Viết chương trình kiểm tra một câu có đúng cú pháp hay không?***

- 1) Câu được bắt đầu bằng ký tự in hoa, theo sau bởi ký tự thường
- 2) Kết thúc bằng dấu chấm hoặc sau một ký tự in hoa
- 3) Các từ cách nhau bằng một khoảng trắng, không chấp nhận nhiều hơn 1 khoảng trắng liên tiếp
- 4) Không tồn tại hai ký tự liên tiếp viết hoa

### **C. *Viết chương trình in ra các từ thỏa một trong các trường hợp sau:***

- 1) Các từ có chứa các ký tự thường ‘a-z’ và số từ ‘0-9’
- 2) Các từ có chứa ký tự ‘a’ theo sau bởi b (b xuất hiện ít nhất 0 lần)
- 3) Các từ bắt đầu bằng ‘a’, theo sau là ký tự bất kỳ và kết thúc bằng ‘b’
- 4) Các từ chỉ chứa ký tự thường ‘a-z’ và ‘\_’
- 5) Các từ có chiều dài là 5
- 6) Các từ có chứa ký tự ‘h’
- 7) Các từ bắt đầu là số từ ‘0-9’
- 8) Các từ có chứa dấu ‘\_’ và thay bằng khoảng trắng
- 9) Có chứa định dạng mm-dd-yy, và chuyển thành định dạng dd-mm-yy

### **D. *Sử dụng biểu thức chính quy để giải quyết bài tập ở bài thực hành số 1***

## **Buổi 4**

1. *Kiểm tra văn phạm tuyến tính trái/ tuyến tính phải (Gợi ý: đọc file lưu cấu trúc của văn phạm đã ch, và kiểm tra)*

2. *Kiểm tra văn phạm chính quy.*
3. *Kiểm tra chuỗi nhập có thuộc văn phạm đã cho hay không?*