

BÀI TẬP LỚN MÔN HỌC XLNNTN
LỚP CAO HỌC CNTT HK232

Đề bài: Xây dựng hệ thống hỏi đáp đơn giản về các môn học cao học ngành Khoa học máy tính.

1. Dữ liệu các môn học:

TT	Môn học	MSMH	NH1	NH2
01	An ninh mạng	055166	12	10
02	Bảo mật cơ sở dữ liệu	055203	10	02
03	Dữ liệu lớn	055153	02	12
04	Cơ sở toán học cho máy tính	055263	12	12
05	Học máy và ứng dụng	055241	12	12
06	Khai phá dữ liệu	055131	02	10
07	Học sâu và ứng dụng trong thị giác	055255	12	10
08	Kho dữ liệu và hệ hỗ trợ ra quyết định	055022	12	02
09	Kiến trúc phần mềm	055247	02	10
10	Kiểm thử phần mềm	055252	10	02
11	Mật mã học ứng dụng	055221	10	12
12	Phát triển ứng dụng IoT	055248	02	10
13	Phân tích chương trình	055146	12	12
14	Phân tích mã độc	055253	10	02
15	Quản lý dữ liệu doanh nghiệp	055234	12	00
16	Tính toán song song	055151	02	10
17	Xử lý ngôn ngữ tự nhiên	055143	12	10
18	Quản lý hệ thống thông tin	055102	12	02
19	Trí tuệ kinh doanh	055235	10	02
20	Quản trị hệ thống thông tin kinh doanh	055010	02	12

Giải thích về dữ liệu môn học: Cột 2 là tên các môn học; cột 3 là mã số môn học; cột 4 và 5 là năm học thứ 1 và thứ 2, có giá trị là 1 số trong 4 số nguyên không dấu là 10, 12, 02, 00, dùng để mô tả môn học sẽ được dạy trong học kỳ 1 (10), học kỳ 2 (02), cả hai học kỳ (12) hoặc không được dạy cả 2 học kỳ trong năm học tương ứng.

2. Yêu cầu bài toán:

2.1 Cho các câu truy vấn:

- 1) Có bao nhiêu môn học được dạy trong học kỳ 1, năm học thứ 1 ?.
- 2) Có bao nhiêu môn học được dạy trong học kỳ 2, năm học thứ 1 ?.
- 3) Có bao nhiêu môn học được dạy trong học kỳ 1 và 2, năm học thứ 1 ?.
- 4) Có bao nhiêu môn học được dạy trong học kỳ 1, năm học thứ 2 ?.
- 5) Có bao nhiêu môn học được dạy trong học kỳ 2, năm học thứ 2 ?.
- 6) Có bao nhiêu môn học được dạy trong học kỳ 1 và 2, năm học thứ 2 ?.
- 7) Có bao nhiêu môn học được dạy trong cả 2 học kỳ năm học thứ 1 ?.
- 8) Có bao nhiêu môn học được dạy trong cả 2 học kỳ năm học thứ 2 ?.
- 9) Có bao nhiêu môn học được dạy trong học kỳ 1, năm học thứ 1 và thứ 2 ?.
- 10) Có bao nhiêu môn học được dạy trong học kỳ 2, năm học thứ 1 và thứ 2 ?.
- 11) Có bao nhiêu môn học được dạy trong cả 2 học kỳ năm học thứ 1 và thứ 2 ?.
- 12) Cho biết tên môn học và mã môn học được dạy cả hai học kỳ năm học thứ 1 ?.
- 13) Cho biết tên môn học và mã môn học được dạy cả hai học kỳ năm học thứ 2 ?.
- 14) Cho biết tên môn học và mã môn học được dạy cả hai học kỳ năm học thứ 1 và thứ 2 ?.
- 15) Cho biết tên môn học có mã số: 055131 ?.
- 16) Cho biết mã số môn học “Quản lý dữ liệu doanh nghiệp” ?.
- 17) Cho biết môn học có mã số 055253, được dạy vào học kỳ nào, năm học nào ?.
- 18) Môn học có mã số **055151** và môn học có mã số **055235** có được dạy cùng học kỳ trong cùng năm học ?.
- 19) Môn học có mã số **055010** có được dạy cả hai học kỳ trong 1 năm học không ?.
- 20) Môn học nào chỉ dạy trong 1 năm học, cho biết tên, mã số môn học và học kỳ trong năm học ?.
- 21) Có môn học nào không được dạy trong 1 năm học, cho biết tên, mã số môn học và năm học ?.

2.2 Viết chương trình để hiện thực:

- a) Xây dựng **bộ phân tích cú pháp** của văn phạm phụ thuộc.
- b) Dùng bộ phân tích được xây dựng ở a) phân tích và xuất ra các **quan hệ cú pháp** của từng câu truy vấn ở mục 2.1.
- c) Từ kết quả ở b) tạo các **quan hệ văn phạm** cho từng câu truy vấn ở mục 2.1 với cơ sở dữ liệu đã cho ở mục 1.
- d) Tạo **dạng luận lý** từ các quan hệ văn phạm ở c).
- e) Tạo **ngữ nghĩa thủ tục** từ dạng luận lý ở d).

- f) *Truy xuất dữ liệu* theo mô tả ở bảng trong mục 1, để tìm và *xuất ra các thông tin trả lời* cho các câu truy vấn ở mục 2.1.

3. Yêu cầu khi thực thi:

3.1. *Ngôn ngữ sử dụng* là Python. Entry point cho chương trình là **main.py** nằm ở thư mục gốc. Bài sẽ được chấm trên Python 3.8 trở lên.

3.2. *Nộp bài*:

- Sinh viên nén toàn bộ các file/thư mục (bao gồm source code, data, ...) vào một file nén theo dạng **MSSV.zip** trong đó MSSV chính là mã số sinh viên của sinh viên. *Sinh viên không nén file theo các định dạng khác*. Chú ý: khi giải nén file thì sẽ xuất hiện các thư mục và 2 file dưới đây.
- **3 thư mục**:
 - Input: gồm 1 file: query.txt (chứa **21 câu hỏi**); các file dữ liệu, ...
 - Output: gồm 5 file kết quả thực thi từ *câu b đến câu f* trong phần 2.2, output của yêu cầu trước sẽ là input của yêu cầu sau.
 - Models: các lớp hoặc các module con thực thi bài toán.
- **2 file sau**:
 - **README.md** bao gồm thông tin cá nhân và các chú thích về các module/lớp có trong thư mục Models.
 - **main.py** (Chương trình khi thực thi KHÔNG có tham số)
