

# BÁO CÁO CUỐI KỲ

## MÔN PHƯƠNG PHÁP LẬP TRÌNH

### (HK 1/ 2021-2022)

---

## Thông tin chung

### Một số lưu ý trước khi làm bài

- Ngôn ngữ sử dụng để làm bài là C, không phải C++.
- Bài tập lớn là bài làm cá nhân, không làm theo nhóm, mỗi sinh viên tự làm bài của mình.
- Sinh viên phải nộp bài trên Classroom, giảng viên không nhận bài qua email hay bất kỳ hình thức nào khác.
- Classroom sẽ bị ẩn khi hết hạn nộp bài do vậy sinh viên cần nộp bài sớm nhất có thể.
- Khi nộp bài sinh viên phải bấm nút Submit và chờ đến khi báo nộp thành công, kiểm tra cẩn thận để đảm bảo là bài làm đã được nộp.
- Sinh viên cần thường xuyên cập nhật thông báo trên Classroom.
- Bài làm sẽ được biên dịch trên C chuẩn (C99) bằng câu lệnh `gcc`. Sinh viên cần thận tuyệt đối không dùng thêm thư viện của C++. Những bài vi phạm quy định này sẽ bị cho 0 điểm.
- Bài làm sẽ được chấm tự động bằng máy do vậy sinh viên chỉ nhập và xuất dữ liệu theo đúng yêu cầu của đề không nhập xuất thừa, không cần in ra các câu thông báo nhập hay các câu thông báo kết quả.
- Không dùng hàm `getch()` hay `system("pause")` để dừng màn hình vì như vậy khi chấm tự động sẽ bị xem là lỗi.

### Xử lý đạo văn (sao chép bài của nhau hoặc sao chép trên mạng internet)

- Sinh viên phải tự làm bài của mình, bài làm sẽ bị xem là đạo văn nếu có sự giống nhau giữa các mã nguồn các bài nộp.
- Mã nguồn bài làm của sinh viên sẽ được kiểm tra bằng đạo văn bằng phần mềm chuyên dụng để đối chiếu với các bài làm của sinh viên khác cũng như đối chiếu với các nguồn trên internet.
- Những bài bị kết luận là đạo văn sẽ bị 0 điểm và sinh viên đạo văn sẽ bị cấm thi cuối kỳ.

# Nội dung

Mục tiêu bài tập: giúp sinh viên hiểu và vận dụng được cấu trúc cơ bản của ngôn ngữ C, các kiểu dữ liệu, các phép toán số học, cách thức đọc/ghi và xử lý trên tập tin, cách thức tổ chức, dịch và thực thi chương trình khi lập trình trên máy tính.

Trong bài tập này, sinh viên được cung cấp 03 tập tin \*.csv chứa các dữ liệu về danh sách sinh viên, danh sách môn học và kết quả học tập. Thông tin chi tiết của từng tập tin như sau:

STT	Tên tập tin	Nội dung	Ghi chú
1	dssv.csv	Chứa danh sách thông tin sinh viên	
2	dsmh.csv	Chứa danh sách thông tin môn học	
3	diem.csv	Chứa kết quả học tập của các sinh viên	Số lượng sinh viên ở các môn học không nhất thiết bằng nhau. <b>Điểm của một sinh viên có thể nằm rải rác trong tập tin, không nhất thiết phải nằm cạnh nhau.</b>

*Tập tin \*.csv là tập tin lưu trữ dữ liệu theo từng dòng, mỗi dòng có các thuộc tính được ngăn cách với nhau bằng dấu phẩy (comma separated values). Dòng đầu tiên trong 3 tập tin mà đề bài cung cấp là dòng chứa tên cột.*

Hãy viết một chương trình C đọc nội dung từ các tập tin csv đã cho, sau đó chương trình sẽ chờ người dùng nhập vào một câu lệnh (trong bảng bên dưới), thực hiện lệnh đã nhập rồi kết thúc. Mỗi lần chạy, chương trình chỉ tiếp nhận và thực thi một câu lệnh duy nhất. Một câu lệnh hoàn chỉnh gồm có hai từ được ngăn cách với nhau bởi duy nhất một khoảng cách, ví dụ lệnh “**top 10**” trong đó **top** có thể hiểu là phần lệnh điều khiển còn **10** là tham số cho lệnh điều khiển.

Nếu muốn thực hiện câu lệnh khác thì người dùng cần phải chạy lại chương trình từ đầu. Khi chờ người dùng nhập lệnh, chương trình không cần in bất cứ thông báo nào ra màn hình console.

Mỗi lần chương trình chạy, chương trình sẽ tạo ra tập tin log.txt nằm cùng thư mục. Tập tin này chứa 03 số nguyên trên cùng một dòng. Ba số nguyên này đại diện cho số dòng dữ liệu (**không bao gồm tiêu đề**) có trong các tập tin dssv.csv, dsmh.csv, diem.csv. Nếu tập tin log.txt đã tồn tại thì chương trình sẽ ghi đè lên nội dung mới. Ví dụ tập tin log.txt chứa nội dung “10 5 30” có nghĩa là trong tập tin dssv.csv có 10 dòng, trong tập tin dsmh.csv có 5 dòng và trong tập tin diem.csv có 30 dòng. **Tập tin log.txt cần được ghi trước khi tiếp nhận lệnh từ bàn phím, vì vậy dù lệnh nhập vào hợp lệ hay không hợp lệ thì tập tin log.txt vẫn được tạo ra.**

Chương trình chỉ được viết trong một tập tin duy nhất đặt tên là *mssv.c*, trong đó *mssv* chính là mã số của sinh viên thực hiện (ví dụ *520H1234.c*). Các tập tin .csv và .c phải được đặt trong cùng một thư mục. Sinh viên **không** được tự ý đổi tên các tập tin .csv. Ngoài ra, sinh viên **cần** sử dụng đường dẫn tương đối khi viết code đọc file, **không** được sử dụng đường dẫn tuyệt đối (ví dụ không được sử dụng đường dẫn theo kiểu như sau: *C:/Users/admin/Desktop/dssv.csv*) vì khi chạy chương trình trên máy tính người chấm, đường dẫn có thể khác, dẫn đến việc đọc tập tin sẽ bị lỗi.

Chương trình chỉ tiếp nhận các lệnh được mô tả trong bảng bên dưới (không phân biệt viết hoa và viết thường). Nếu người dùng nhập vào bất cứ lệnh nào khác với các lệnh trong bảng bên dưới hoặc lệnh giống nhưng với tham số không hợp lệ (ví dụ lệnh *top n* nhưng với *n* là số khác số nguyên dương) thì đều được xem là lệnh không hợp lệ. Khi đó, chương trình cần tạo ra một tập tin tên là **error.txt** nằm cùng thư mục với chương trình, tập tin này chứa nội dung là: *invalid command*. Sau đó, chương trình sẽ kết thúc mà không in ra bất cứ nội dung nào trên màn hình console. Nếu tập tin error.txt đã tồn tại thì chương trình sẽ tạo tập tin mới và ghi đè (override) lên tập tin đã có.

Danh sách các câu lệnh được hỗ trợ

STT	Câu lệnh	Mô tả
1	<b>list</b> <i>classname</i>	<ul style="list-style-type: none"> <li>- Tìm tất cả các sinh viên thuộc lớp <i>classname</i> và ghi kết quả vào tập tin <i>result.csv</i>, mỗi sinh viên được ghi trên một dòng. Trong đó <i>classname</i> chính là tên lớp cần tìm.</li> <li>- Ví dụ: <i>list 5202058</i> sẽ ghi ra danh sách sinh viên thuộc lớp 5202058.</li> <li>- Nếu người dùng nhập vào một mã lớp không tồn tại thì tập tin <i>result.csv</i> vẫn được tạo ra nhưng sẽ có nội dung rỗng.</li> </ul>
2	<b>count male</b> <b>count female</b>	Đếm số lượng sinh viên nam/nữ có trong danh sách sinh viên và ghi kết quả đếm được vào tập tin <i>result.csv</i> (kết quả trong tập tin <i>result.csv</i> là một số nguyên)
3	<b>top</b> <i>n</i>	<ul style="list-style-type: none"> <li>- Thực hiện xuất ra file <i>result.csv</i> danh sách <i>n</i> sinh viên có điểm trung bình từ cao xuống thấp, mỗi sinh viên trên một dòng khác nhau. <b>Lưu ý: câu lệnh này in ra thông tin sinh viên giống với định dạng có trong tập tin <i>dssv.csv</i> (trong tập tin <i>result.csv</i> không có thông tin điểm trung bình, điểm trung bình chỉ dùng để lọc danh sách trước khi ghi)</b></li> <li>- Nếu <math>n \leq 0</math> thì lệnh được xem là không hợp lệ.</li> <li>- <b>Điểm trung bình được tính bằng công thức <math>S = (d_1 + d_2 + \dots + d_n) / n</math>. Trong đó <math>d_i</math> là điểm của môn thứ <math>i</math>. Trong công thức của điểm trung bình không quan tâm đến hệ số tín chỉ.</b></li> </ul>

4	<b>sort asc sort desc</b>	Sắp xếp và xuất ra file <b>result.csv</b> danh sách sinh viên theo thứ tự tăng dần/giảm dần trên thuộc tính <i>first_name</i> , mỗi sinh viên trên một dòng khác nhau. Nếu 2 sinh viên có <i>first_name</i> giống nhau thì sắp xếp theo ngày tháng năm sinh. <b>Nếu vẫn trùng ngày tháng năm sinh thì sắp xếp theo last_name. Nếu vẫn trùng last_name thì sắp xếp theo country.</b>
5	<b>country <u>name</u></b>	<ul style="list-style-type: none"> <li>- Tìm tất cả các sinh viên đến từ quốc gia <i>name</i> và ghi kết quả vào tập tin <b>result.csv</b>, mỗi sinh viên được ghi trên một dòng. Trong đó <i>name</i> chính là tên quốc gia cần tìm.</li> <li>- Ví dụ: <code>country singapore</code> sẽ ghi ra danh sách sinh viên đến từ nước singapore.</li> <li>- Nếu người dùng nhập vào một mã tên nước không tồn tại thì tập tin <b>result.csv</b> vẫn được tạo ra nhưng sẽ có nội dung rỗng.</li> </ul>

Ngoài mô tả bên trên cần lưu ý thêm một số điều sau về các câu lệnh:

- Các câu lệnh chỉ có thể chứa các ký tự alphanumeric (ký tự viết hoa, viết thường và số 0-9) và duy nhất một khoảng trắng. Nếu lệnh chứa bất kỳ ký tự nào khác thì chắc chắn là một lệnh KHÔNG hợp lệ.
- Để cho thuận tiện, câu lệnh *list classname* (lệnh số 1) và câu lệnh *country name* (lệnh số 5) thì trong tên *classname* và tên quốc gia được phép chứa chữ cái và số. Nếu người dùng nhập lệnh “*classname 12a112*” với 12a112 giả sử là một mã lớp không tồn tại thì cách giải quyết chỉ đơn giản là tạo ra tập tin **result.csv** với nội dung rỗng.
- Dữ liệu ở cột giới tính trong testcase luôn là “Male” hoặc “Female”, dữ liệu tên quốc gia luôn bắt đầu bởi một ký tự viết hoa. Tuy nhiên câu lệnh 5 (*country name*) vẫn phải chạy đúng nếu nhập tên quốc gia viết thường.

#### Lưu ý:

- Khi người dùng nhập một lệnh hợp lệ thì tập tin **error.txt** đương nhiên sẽ không được tạo ra. Nếu trước đó đã có sẵn một tập tin **error.txt** thì chương trình KHÔNG cần xóa hay làm gì với tập tin này cả.
- Dữ liệu trong các testcase lúc chấm sẽ có sự khác biệt về số dòng cũng như giá trị dữ liệu nhưng vẫn đảm bảo theo định dạng và thứ tự ban đầu. Số dòng trong các tập tin csv có thể lên đến 1500 dòng.
- Trong tập tin **diem.csv**, có thể có sinh viên A xuất hiện 10 lần, sinh viên B xuất hiện 2 lần và sinh viên C không xuất hiện lần nào. Sinh viên A xuất hiện 10 lần nhưng không nhất thiết có dữ liệu là 10 dòng liền kề với nhau.
- Mỗi lần chạy chương trình thì tập tin **result.csv** sẽ ghi đè lên nội dung cũ (nếu có).
- Danh sách sinh viên xuất ra trong tập tin **result.csv** không bao gồm tên cột, **không in dư dấu xuống dòng ở dòng cuối cùng.**

- Thứ tự các thuộc tính của mỗi sinh viên trong tập tin result.csv phải đúng thứ tự với các thuộc tính trong tập tin dssv.csv.
- Các câu lệnh hợp lệ là lệnh thỏa mãn đồng thời hai yếu tố:
  - o Có từ khóa giống như trong bảng mô tả
  - o Có đúng số lượng và kiểu dữ liệu của tham số như trong bảng mô tả
- Ví dụ về các lệnh hợp lệ:
  - o `sort asc`
  - o `list abcdef` (mặc dù không có lớp nào tên `abcdef` nhưng lệnh vẫn hợp lệ về mặt cú pháp)
  - o `count male`
- Ví dụ về lệnh không hợp lệ (ghi 'invalid command' ra tập tin error.txt)
  - o `sort xyz` (do `sort xyz` không được mô tả trong bảng, chỉ có `sort asc` và `sort desc`)
  - o `list 52101234 abcd` (do lệnh `list` chỉ nhận 1 tham số, ở đây truyền tới 2 tham số)
  - o `list 52101234` (do có hai khoảng trắng trong câu lệnh)
  - o `list 52101234` (do dư một khoảng trắng phía cuối câu lệnh)
  - o `shutdown` (do lệnh `shutdown` không được mô tả trong bảng)
  - o `top` (do thiếu tham số `n`)
  - o `top abc` (do tham số cần truyền là số nguyên nhưng ở đây lại truyền chuỗi `abc`)
  - o `top 10x` (do `10x` không phải là một số hợp lệ)
  - o `top -10` (do tham số `n`  $\leq 0$  nên không hợp lệ)
  - o `top +10` (do có dấu `+` là ký tự không hợp lệ)
  - o `top 10` (do dư một khoảng trống trước từ khóa `top`).
  - o `top 5` (do có nhiều hơn một dấu cách giữa từ khóa `top` và tham số `5`).

## Thời hạn và cách nộp bài

- Sinh viên nộp bài trên phần Assignment của Google classroom của từng lớp học.
- Sinh viên đăng nhập vào Classroom và nộp bài trong mục Assignments theo thời gian quy định trong Assignments.
- Sinh viên chỉ nộp 1 file duy nhất đặt tên là **MSSV.c** không nén (Trong đó **MSSV** là mã số sinh viên của người nộp bài). Bài nộp đặt tên sai sẽ bị trừ  $\frac{1}{2}$  số điểm.

Mọi thắc mắc sinh viên gửi email: [levanvang@tdtu.edu.vn](mailto:levanvang@tdtu.edu.vn) để được giải đáp.