

CS205 - LAB2

Write a program that takes a series of numbers on the command line, reads from stdin and displays on stdout the corresponding fields (fields in the input are supposed to be separated by a special separator character). The same character will be used in the output as in the input.

You are provided for testing with a big file called `unidata.csv` (coming from a US government website) with the list of higher-education institutes in the US, as well as a spreadsheet describing the data.

Your program will use Unix-style flags to specify some options. The function to use for processing options is called `getopt()` and `man 3 getopt` provides explanations as well as a complete example ("3" is required, there is another `getopt` in section 1 of the manual). You can copy, paste and adapt the example. Ask if something isn't clear for you.

You are asked to implement two flags:

- s followed by the separator used (default separator: one space)
- i followed by the number of lines to ignore (default: 0). You may have lines to ignore when the first line(s) contains headers, not data. It's the case in the sample file that is provided.

Example:

If you just want to extract from `unidata.csv` the state abbreviation (STABBR, 5th field) whether the institution is active (CYACTIVE, 39th field) and the institutional category (INSTCAT, 45th field), you should enter the command (I'm calling the program `extract`, you can give it any name you want):

```
extract -s ';' -i 1 5 39 45 < unidata.csv
```

-s ';' means that the separator is a semi-colon

-i 1 means to ignore (skip) the first line.

Simplifying assumptions:

- Field numbers are always provided in increasing sequence (typing "39 5" is forbidden, because 39 is greater than 5)
- You cannot repeat a field ("5 5" is forbidden)

Your program will check that assumptions are verified and display an error message on `stderr` if they aren't, and return something else than 0 to indicate the error.