BOOK SCANNING

GROUP MEMBERS:

- ANUSHKA AGRAWAL(2019140002)
- ADITI BHOSALE(2019140009)
- SHWETA CHAVAN(2019140011)

Table of Contents

01

INTRODUCTION

- To bring world’s best books from various languages online and accessible to everyone
- Increase the speed and efficiency for those looking to specific information
- Digital copy of all possible books
- Ancient books,manuscripts have a high book score
- Books easily available have slightly lower score

02 PROBLEM STATEMENT

- There are **B** different books(0 to B-1), **L** libraries(0 to L-1) and **D** days(0 to D-1)
- Copies of same book can be available in different libraries
- Each book has a book score
- A book can be scanned multiple times but its score is calculated once
- Books can be scanned in parallel from multiple libraries
- Each library has 3 additional parameters
  - Set of books in library
  - Time (days)for signing up library for scanning
  - Number of books that can be scanned per day
- First library starts its sign up at D=0
- D-1 is the deadline for the scanning process

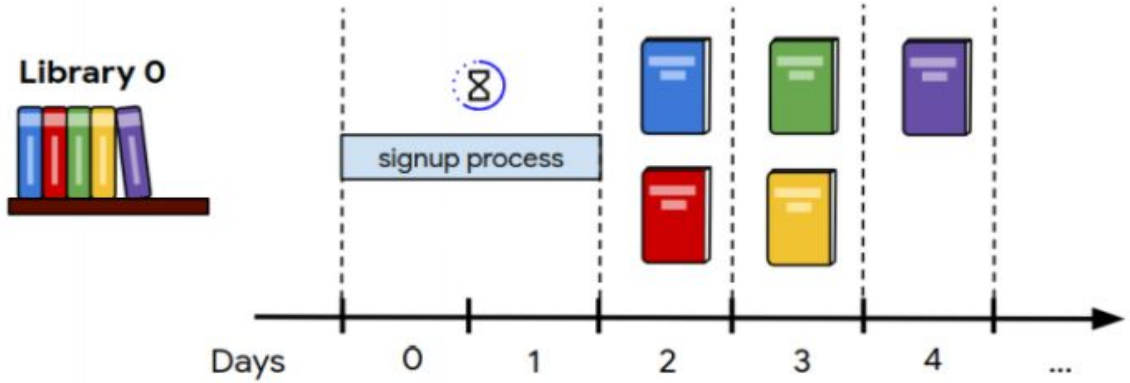
PROBLEM STATEMENT

SIGN UP PROCESS:

- For each and every library
- Any order
- Scanning process followed by sign up
- One library at a time

For **example**, if library 0 has 5 books, can ship 2 books per day, and completes the signup process on day 1, total books are 5 and total days are 5

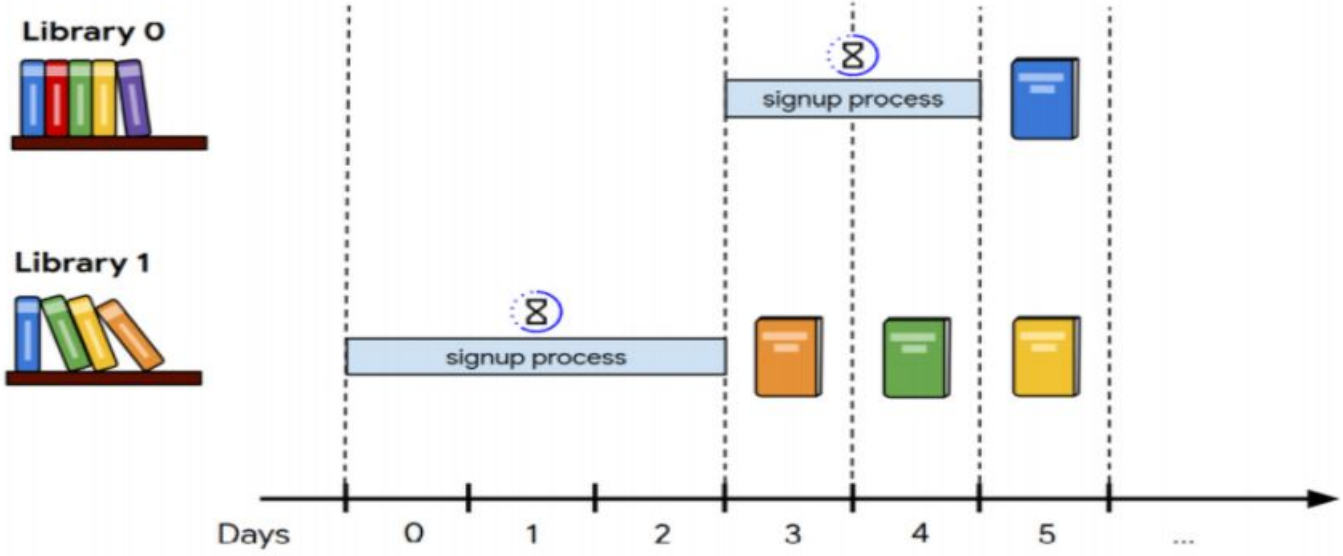
- 2 books can be scanned on day 2
- 2 books can be scanned on day 3
- the one remaining book can be scanned on day 4



For **example**, if

- the signup process for library 0 (that is, the library with ID 0) takes 2 days, and
- the signup process for library 1 takes 3 days, and
- library 1 is signed up before library 0

total days are 6 and total books in library 0 and 1 are 5 and 6



PROBLEM STATEMENT

Given the description of books ,libraries and days assigned and sign up process ,we are supposed to find the **maximum score** that can be obtained after scanning the books from various libraries

FORMULAS AND FUNCTIONS

- score for each library:
  library_score = library_books_score / library_signup_days
- library_available_days = all_days - library_signup_days
- total_books(k) = available days * max books per day.
- After finding total_books, find the k most top scoring books for the library
- Sum up the scores of all these k books for each library = library_book_score

FORMULAS AND FUNCTIONS

- Find library_score which is the total_books_score we calculated divided by signup days.
- Then sort according to library_score
- Take one at a time, find library_available_days = all_days - prev_lib_days-current_lib_signup_days
- scan top total_books of the library, which haven't been scanned previously, until days are available

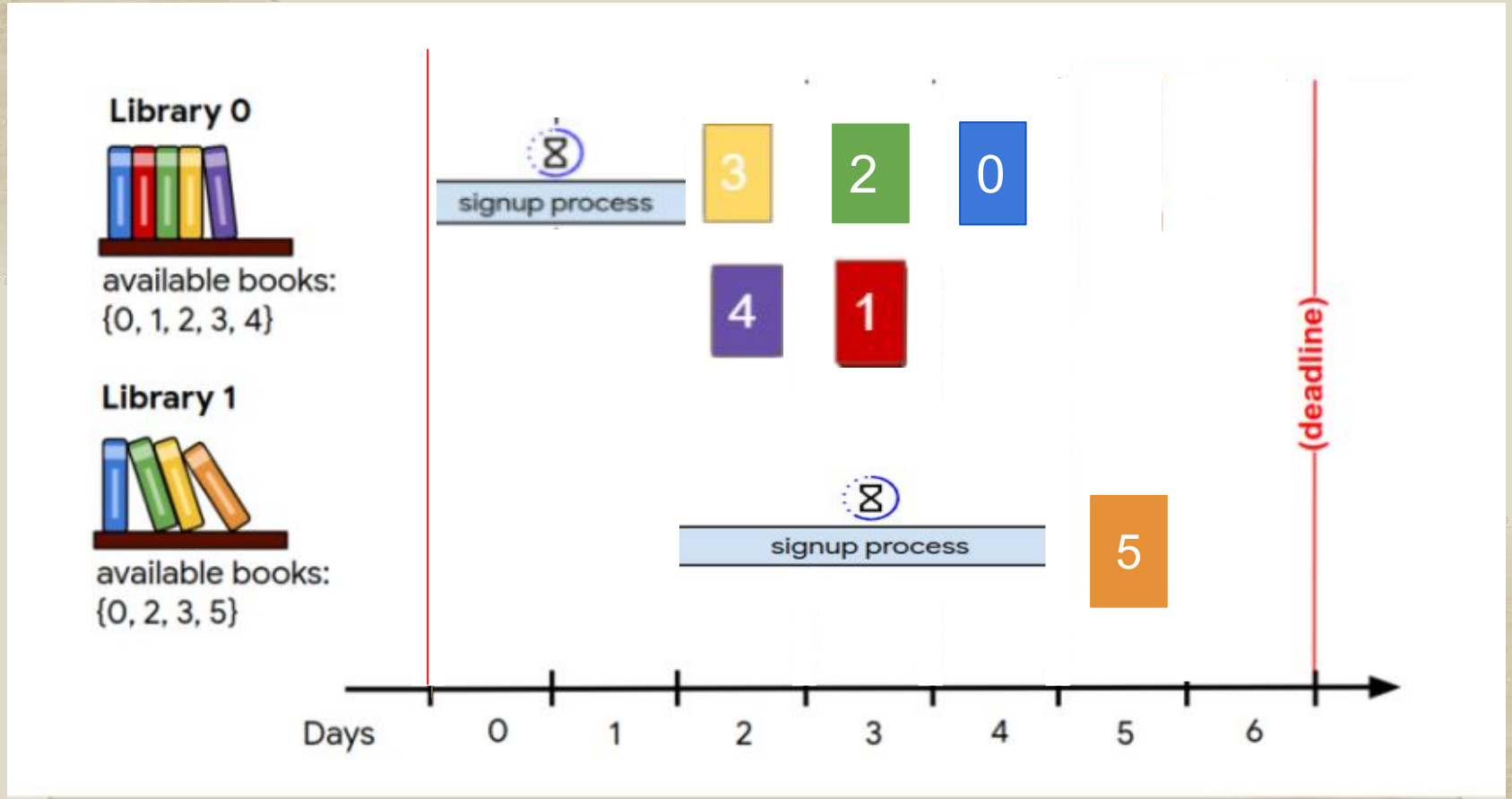
Input file	Description
`6 2 7`	There are 6 books, 2 libraries, and 7 days for scanning.
`1 2 3 6 5 4`	The scores of the books are 1, 2, 3, 6, 5, 4 (in order).
`5 2 2`	Library 0 has 5 books, the signup process takes 2 days, and the library can ship 2 books per day.
`0 1 2 3 4`	The books in library 0 are: book 0, book 1, book 2, book 3, and book 4.
`4 3 1`	Library 1 has 4 books, the signup process takes 3 days, and the library can ship 1 book per day.
`3 2 5 0`	The books in library 1 are: book 3, book 2, book 5 and book 0.

*daaproj - Notepad

File Edit Format View Help

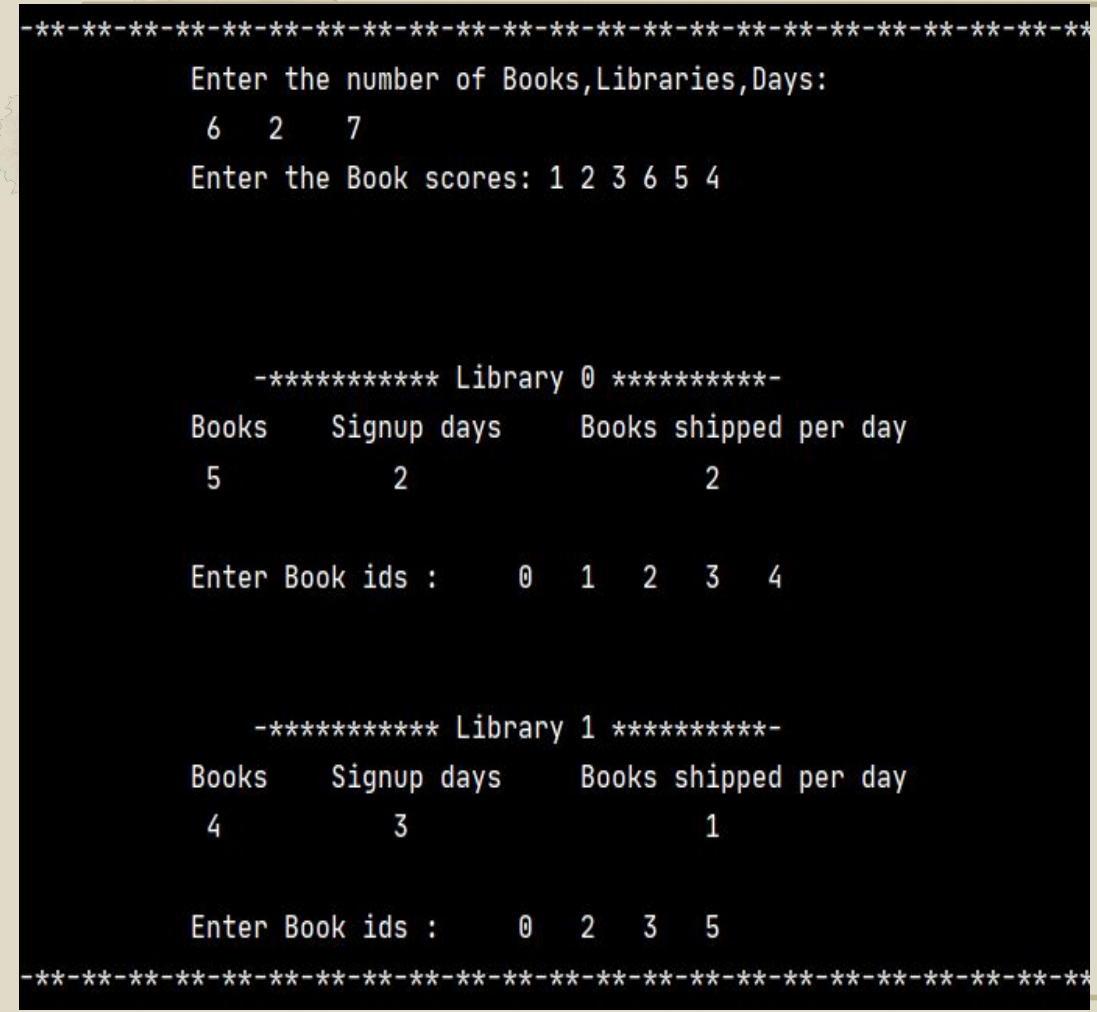
```
L0-5 avaialble days after signup
L1-4 available days after signup

L0: 5*2=10
L1: 4*1=4

L0: 3 4 2 1 0 (6 5 3 2 1)

L1: 3 5 2 0   (6 4 3 1)

L0 book score = 16
L1 book score = 14

L0 score: 16/2=8
L1 score: 14/3=4.6


Therefore first library 0 then Library 1

L0 available days = 5(D2-6)
scanned books: 3 and 4 (D2) 2 and 1 (D3)  0(D4)

L1 available days = 7-2-3=2( D5 and D6)
scanned books: 5 (D5)
```

Ln 26, Col 1 | 100% | Windows (CRLF) | UTF-8


Library 0
available books:
{0, 1, 2, 3, 4}
Library 1
available books:
{0, 2, 3, 5}
signup process
3
2
0
4
1
signup process
5
(deadline)
Days
0
1
2
3
4
5
6

07

Output



- There are 6 books,2 libraries and 7 days for scanning.
- Score of book 0 is 1, that of book 1 is 2 and so on
- Library 0 has 5 books, the signup process takes 2 days, and the library can ship 2 books per day.
- Books in library 0 are book 0,book 1 so on
- Library 1 has 4 books, the signup process takes 3 days, and the library can ship 1 books per day.
- Books in library 1 are book 0,book 2,book 3 so on

```
**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-*

Number of libraries scanned up for signing =  2

Library chosen                    Number of books scanned from library

        0                                    5

Book ID's of books chosen from library : 0 1 2 3 4

Library chosen                    Number of books scanned from library

        1                                    1

Book ID's of books chosen from library : 5


Total score: 21


**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-**-*
```

- Two libraries will be signed for scanning
- The first library to do the signup process is library 0.After the signup process it will send 5 books.
- Library 0 will send book 0, book 1 & so on in order
- The second library to do the signup process is library 1.After the signup process it will send book.
- Library 1 will send book 5.
- The Total score is the sum of all books that are scanned within 7 days.

Conclusion

- Hence we are able to explore the challenges of setting up a scanning process for millions of books.
- Here we are able to plan by viewing the problem from various angles.

FUTURE SCOPE

When calculating library scores we can take into account the repetition of books, since otherwise two libraries with similar books may have higher score compared to some other library with different books

Thanks