a)

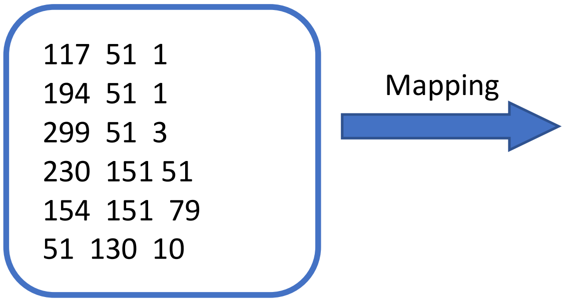
Map function:

Input: src tgt weight

Output: tgt as key and weight as value

Map function take in data and treat the second parameter (tgt) every line as a key, and third parameter (weight) as value.

In this case:



117 51 1
194 51 1
299 51 3
230 151 51
154 151 79
51 130 10

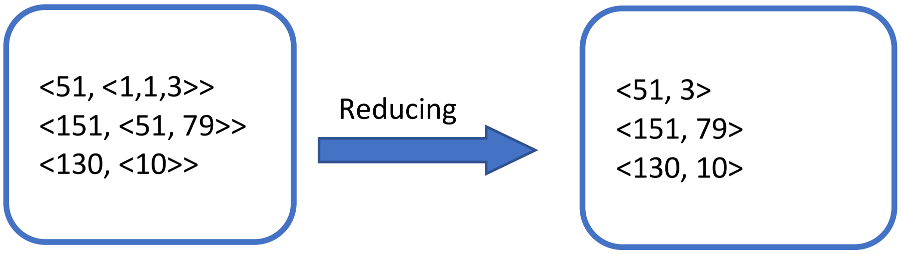
Mapping

<51, 1>
<51, 1>
<51, 3>
<151, 51>
<151, 79>
<130, 10>

Reduce function:

Input: unique tgt as key and corresponding weight list

Output: tgt as key max value in weight list



<51, <1,1,3>>
<151, <51, 79>>
<130, <10>>

Reducing

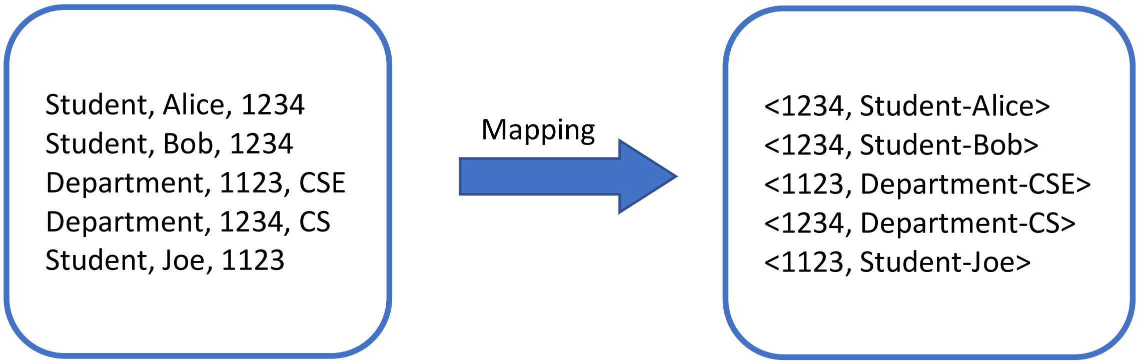
<51, 3>
<151, 79>
<130, 10>

b)

Mapper:

In mapper function, I will use Department_ID as the key for both Department and Student. And for value, I choose to combine the title (student or department) and name to form a value.

Given the record in HW3 PDF as an example:

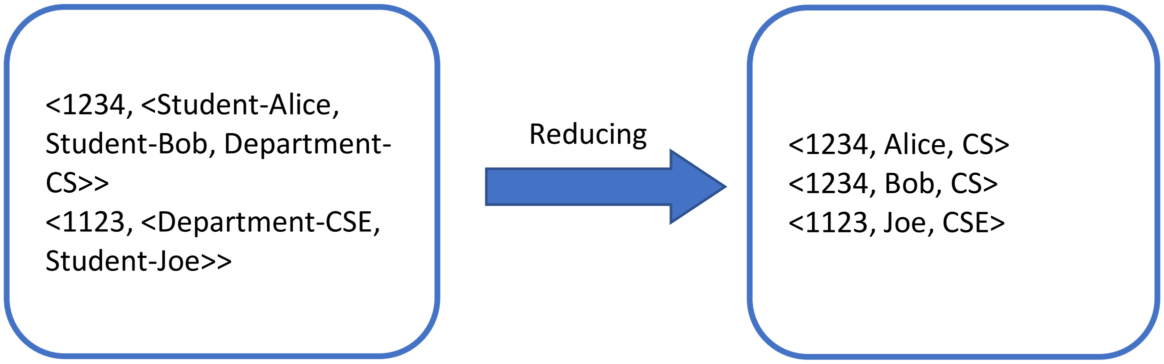


Student, Alice, 1234
Student, Bob, 1234
Department, 1123, CSE
Department, 1234, CS
Student, Joe, 1123

<1234, Student-Alice>
<1234, Student-Bob>
<1123, Department-CSE>
<1234, Department-CS>
<1123, Student-Joe>

Reducer:

After shuffle phase, we already have values grouped by key (Department_ID). In this case, we have values <Student-Alice, Student-Bob, Department-CS> for key 1234 and <Department-CSE, Student-Joe> for key 1123. So, in our reducer function, for every key, we firstly find the department value and save the department name. And then we iterate over all student values and emit <Department_ID, Student_Name, Department_Name> as final tuple.



<1234, <Student-Alice, Student-Bob, Department-CS>>
<1123, <Department-CSE, Student-Joe>>

<1234, Alice, CS>
<1234, Bob, CS>
<1123, Joe, CSE>

Pseudo code:

Mapper(file):

```
for line in file:
   if (line[0] == “Student”)
      emit <line[2], line[0] + “-” + line[1]>
   else
      emit <line[1], line[0] + “-” + line[2]>
```

Reducer(key, values):

```
for value in values:
   if value.startWith(“Department”)  Department_Name = value.split(‘-’)[1]
for value in values:
   if value.startWith(“Student”) emit <key, value.split(‘-‘)[1], Department_Name>
```