CSC - 502

Assignment - 3

- Shaina Patel (V00949940)
- Samath Reddy Vemula (V00949217)

Q1 1) Cat : 234569 : 1
234578 : 1
234839 : 1

Dog : 234569 : 1
234578 : 1
234879 : 1

2) Dog : 234569 : 1
9 : 1
301 : 1

Binary : 234569 = b111001010 01001001
1 = b1
9 = b1001
1 = b1

VB Encoding : 234569 = 00011100 00101000
11001001
1 = 10000001
9 = 10001001
1 = 10000001
301 = 00000010 10101101
1 = 10000001

3) q: cat and dogs
cat dog

d1: 234569: phrase fight cat dog reflect natural tendency relationship two species antagonistic two species friends

d2: 234578: dog cat bad relationship

d3: 234839: cat furry

d4: 234879: dog man best friend

$$d1 = \sqrt{(1/2)^2 + (1/2)^2 + (1/2)^2 + (1/2)^2 + (1/2)^2 + (1/2)^2 + (2/2)^2 + (2/2)^2 + (1/2)^2 + (1/2)^2}$$

$$= \sqrt{4.5}$$

$$= 2.12$$

$$d2 = \sqrt{1+1+1+1} = 2$$

$$d3 = \sqrt{1+1} = 1.41$$

$$d4 = \sqrt{1+1+1+1} = 2$$

$$|q| = \sqrt{1+1} = \sqrt{2} = 1.41$$

$$\cos(q, d1) = \frac{1(1/2) + 1(1/2)}{1.41 \times 2.12} = 0.33$$

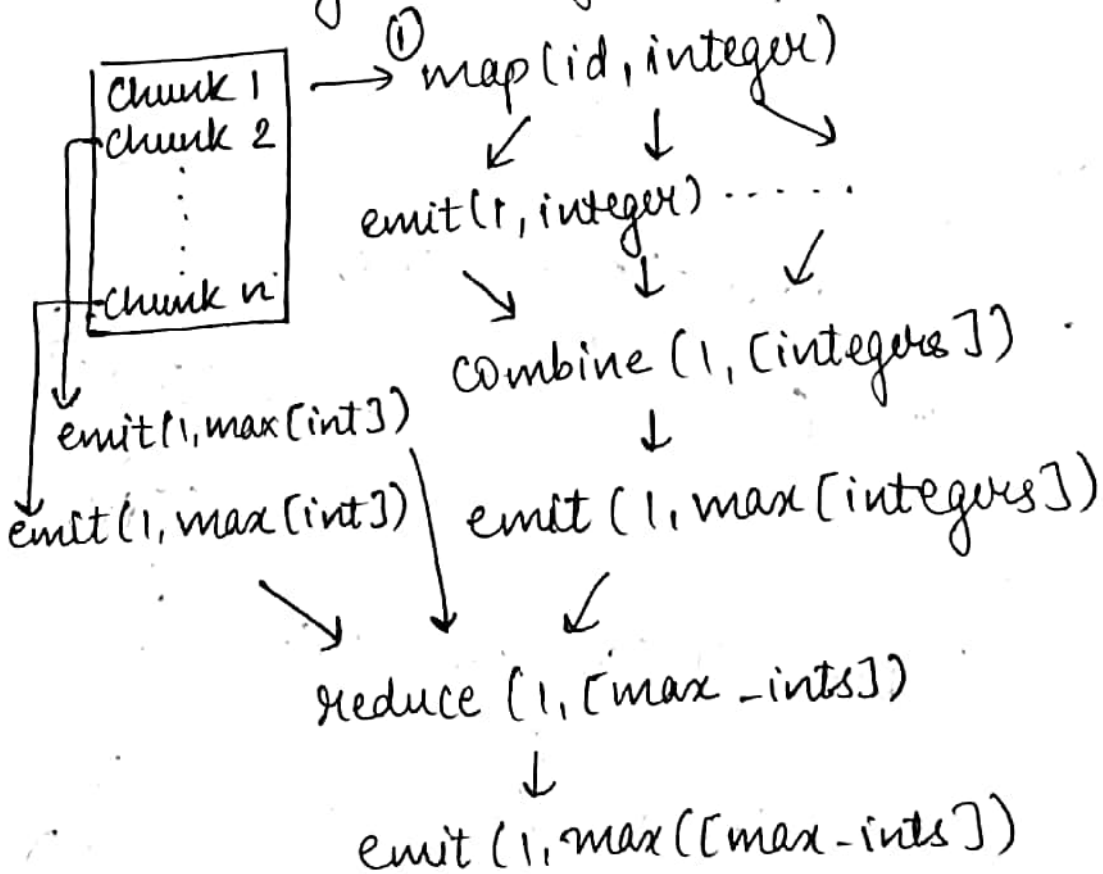
$$\cos(q, d_2) = \frac{1(1) + 1(1)}{1.41 \times 2} = 0.707$$

$$\cos(q, d_3) = \frac{1(1) + 1(0)}{1.41 \times 1.41} = 0.5$$

$$\cos(q, d_4) = \frac{1(0) + 1(1)}{1.41 \times 2} = 0.35$$

Q2

(a) The largest integer



① map(id, integers) : Takes a chunk with its id and integers

- map(id, integers)
  emit(1, integers)

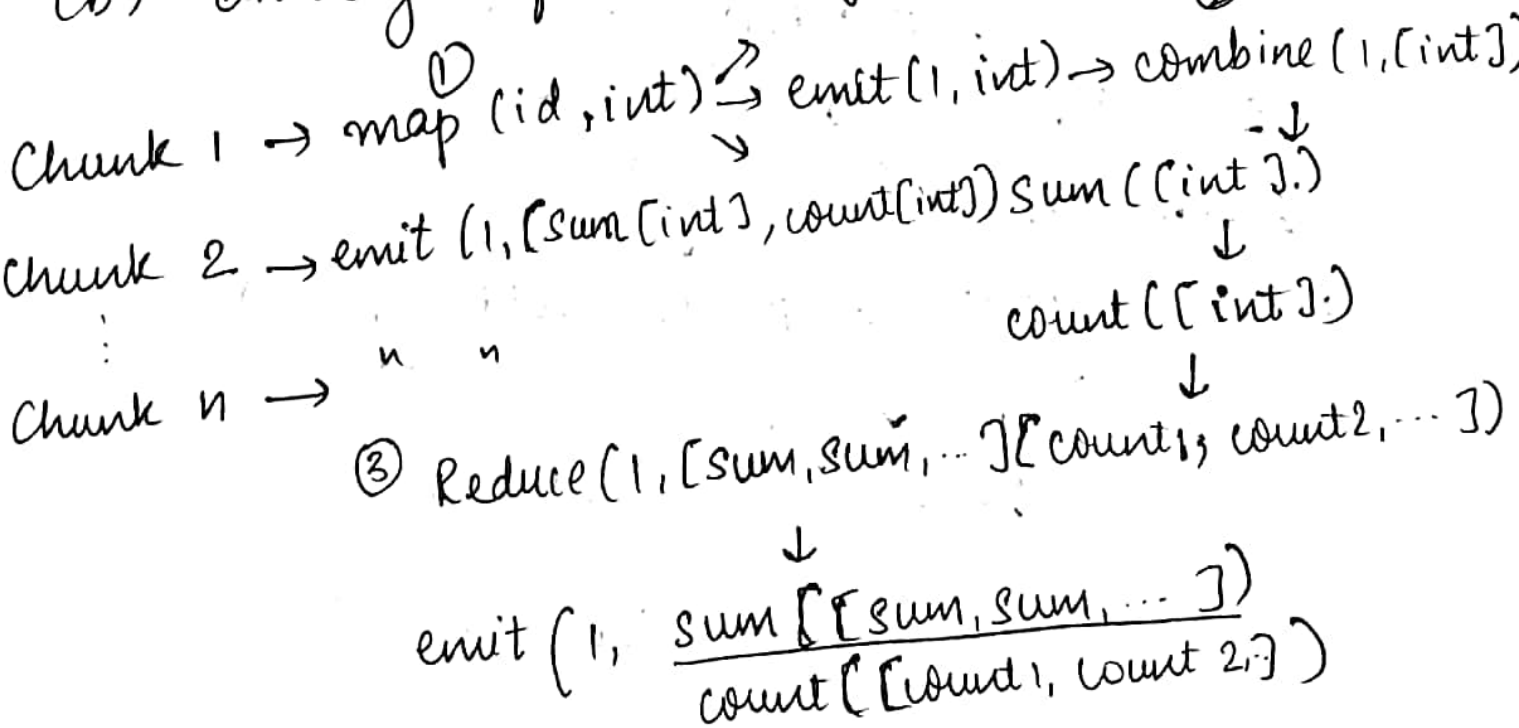
② combine (j, list): Given list of integers, emits the largest integer of chunk

- combine (j, list)
  emit (1, max (list))

③ Reduce (j, list): Takes maximum integer of each chunk of list, emits maximum of list

- Reduce (j, list)
  emit (max (list, null))

(b) Average of all integers



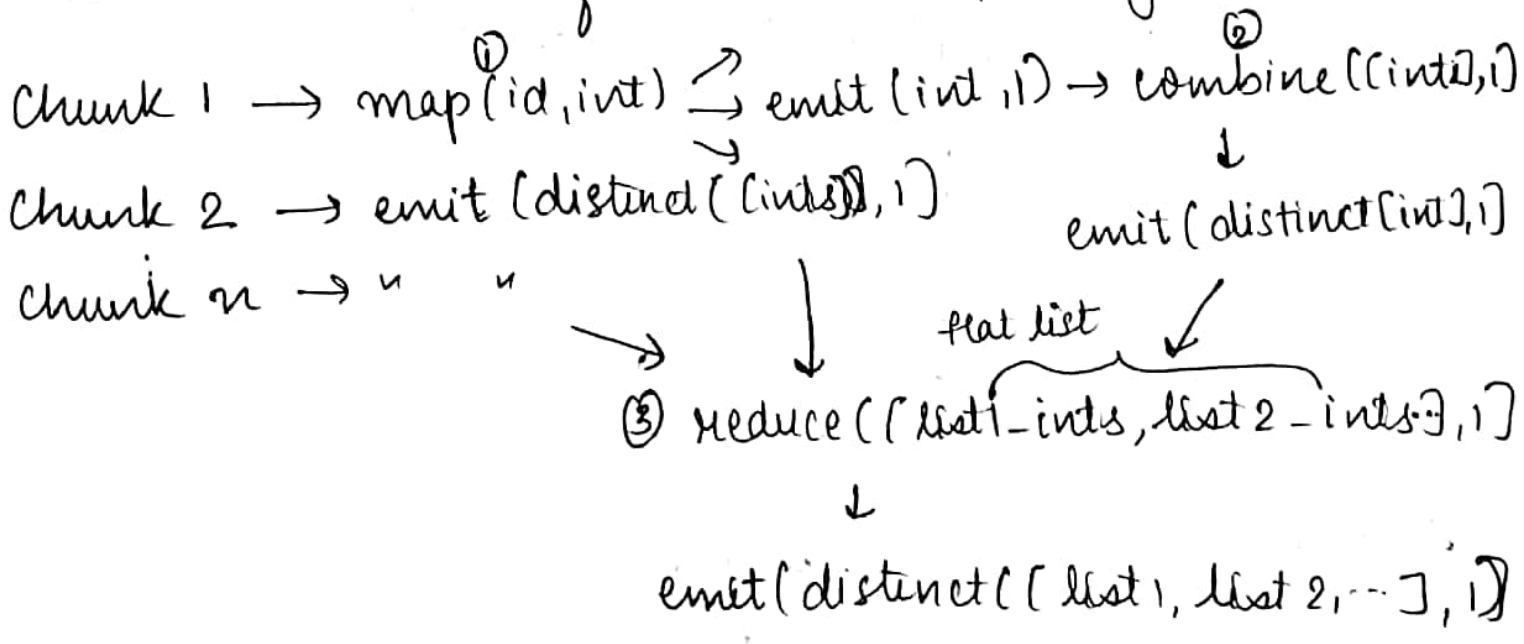
① map (id, integers)
emit (1, i)

② combine (): With the list from mapped & count,

it emits sum of int in chunk

③ Reduce(): Takes all sum & count from combine() as 2 flat lists (sum & count)
Emits the average which is computed as sum/count for list.

(c) Same set of int non repeating



① Map (id, integers): Takes chunk of id & int as i/p

② combine ([ints], 1): With list of int, combine() emits distinct ints of list for mapper

③ Reduce(): Takes flat list of intermediate o/p from combine().
Emit list of distinct integers

(d) Count of no. of distinct int in i/p

Using (c) intermediate O/P from Combine

↓

(2) combine([ints], 1)

↓

emit(distinct([int]), 1)

↓ ↓ ↓

(3) reduce([list1_ints, list2_ints, ...])

flat list

↓

emit(count([list1, list2, ...]), 1)

(1) Map(id, int):
emit(1, 1)

Reduce : Emits the count of int which are distinct

Q 3

(a) Map: For each t in R & S
Emit(t, t)

Reduce: Input(t, t)...
Emit(t, t)

(b) M- the bit showing relation betn R-S

```
Map(tid, (t, M)):
    Emit(t, M)
Reduce(t, list):
    length_R = [M for M in list of M==R]
    length_S = [M for M in list of M==S]
    x = min(len(length_R), len(length_S))
    if x > 0:
        Emit(t, x)
```

(c) (1) Map:

```
for each tuple t in R
    Emit((t,R),1)
for each tuple t in S
    Emit((t,S),1)
```

Reduce:

```
input((t,R),[1,...]) or
input((t,S),[1,...])
Emit((t,R), sum[1,...]=x) or
Emit((t,S), sum[1,...]=y)
```

② Map: Input from ①

((t, R), x) or ((t, S), P)

↓ ↓

Emit (t, x) Emit (t, P)

Reduce: Input (t, [x, y])

Emit (t, x - y = z)

③ Map: Input (t, z)

Emit (t, z)

Reduce: Input (t, z)

if z ≥ 0:

Emit z tuple (t, t)