

Q1. Build inverted index using the modified algorithm

101

|     |     |     |
|-----|-----|-----|
| Cat | Pat | mat |
| Sat | cat | eat |

201

|     |     |     |
|-----|-----|-----|
| pat | mat | Sat |
| pat | mat | eat |

301

|     |     |     |
|-----|-----|-----|
| Sat | mat | Cat |
| pat | fat | mat |

|     |     |     |     |     |
|-----|-----|-----|-----|-----|
| Cat | Pat | mat | Sat | eat |
| 2   | 1   | 1   | 1   | 1   |

|     |     |     |     |
|-----|-----|-----|-----|
| Pat | mat | Sat | eat |
| 2   | 2   | 1   | 1   |

|     |     |     |     |     |
|-----|-----|-----|-----|-----|
| Sat | mat | Cat | Pat | fat |
| 1   | 2   | 1   | 1   | 1   |

Mapper 1  
 ((Cat, 101), 2)  
 ((Pat, 101), 1)  
 ((mat, 101), 1)  
 ((Sat, 101), 1)  
 ((eat, 101), 1)

Mapper 2  
 ((Pat, 201), 2)  
 ((mat, 201), 2)  
 ((Sat, 201), 1)  
 ((eat, 201), 1)

Mapper 3  
 ((Sat, 301), 1)  
 ((mat, 301), 2)  
 ((Cat, 301), 1)  
 ((pat, 301), 1)  
 ((fat, 301), 1)

S-S Partition

((Cat, 101), [2])  
 ((Cat, 301), [1])  
 ((mat, 101), [1])  
 ((mat, 201), [2])  
 ((mat, 301), [2])  
 ((Sat, 101), [1])  
 ((Sat, 201), [1])  
 ((Sat, 301), [1])

((eat, 101), [1])  
 ((eat, 201), [1])  
 ((Pat, 101), [1])  
 ((Pat, 201), [2])  
 ((Pat, 301), [1])  
 ((fat, 301), [1])

Reducer 1 output /

$t_{prev} = null$

$P \leftarrow \text{new List()}$

① ((cat, 101), [2])

$P \leftarrow [101 | 2]$

$t_{prev} = cat$

② ((cat, 301), [1])

$P \leftarrow [101 | 2] \leftrightarrow [301 | 1]$

$t_{prev} = cat$

③ ((mat, 101), [1])

$emit(cat, [101 | 2] \rightarrow [301 | 1])$

$P \leftarrow \text{new List()}$

$P \leftarrow [101 | 1]$

$t_{prev} = mat$

④ ((mat, 201), [2])

$P \leftarrow [101 | 1] \rightarrow [201 | 1]$

$t_{prev} = mat$

⑤ ((mat, 301), [1])

$P \leftarrow [101 | 1] \rightarrow [201 | 1] \rightarrow [301 | 1]$

$t_{prev} = mat$

⑥ ((sat, 101), [1])

$emit(mat, [101 | 1] \rightarrow [201 | 1] \rightarrow [301 | 1])$

$P \leftarrow \text{new List()}$

$t_{prev} = sat$   $P \leftarrow [101 | 1]$

⑦ ((sat, 201), [1])

$P \leftarrow [101 | 1] \rightarrow [201 | 1]$

$t_{prev} = sat$

⑧ ((sat, 301), [1])

$P \leftarrow [101 | 1] \rightarrow [201 | 1] \rightarrow [301 | 1]$

⑨

close method

$emit(sat, [101 | 1] \rightarrow [201 | 1] \rightarrow [301 | 1])$

Reducer 2 output.

$t_{prev} = \text{null}$

$P \leftarrow \text{new List}()$

①  $((\text{eat}, 101), [1])$

$P \leftarrow \boxed{101 | 1}$

$t_{prev} = \text{eat}$

②  $((\text{eat}, 201), [1])$

$P \leftarrow \boxed{101 | 1} \rightarrow \boxed{201 | 1}$

$t_{prev} = \text{eat}$

③  $((\text{fat}, 301), [1])$

$\text{Emit}(\text{eat}, \boxed{101 | 1} \rightarrow \boxed{201 | 1})$

$P \leftarrow \text{new List}()$

$P \leftarrow \boxed{301 | 1}$

$t_{prev} = \text{fat}$

④  $((\text{pat}, 101), [1])$

$\text{Emit}(\text{pat}, \boxed{301 | 1})$

$P \leftarrow \text{new List}() \quad P \leftarrow \boxed{101 | 1}$

$t_{prev} = \text{pat}$

⑤  $((\text{pat}, 201), [1])$

$P \leftarrow \boxed{101 | 1} \rightarrow \boxed{201 | 1}$

$t_{prev} = \text{pat}$

⑥  $((\text{pat}, 301), [1])$

$P \leftarrow \boxed{101 | 1} \rightarrow \boxed{201 | 1} \rightarrow \boxed{301 | 1}$

7) close method

$\text{Emit}(\text{pat}, \boxed{101 | 1} \rightarrow \boxed{201 | 1} \rightarrow \boxed{301 | 1})$