

MDA hw1 report  
106062132趙仰生

以下照順序執行：

1.第一個mapper(假設兩個陣列為 $M_{ij}$  ,  $N_{jk}$  , 最後結果為 $P_{ik}$ )：

Key :  $M$ 的 $j$ 與 $N$ 的 $j$

Value: 設成list的形式 $[M, i, M_{ij}]$  or  $[N, k, N_{jk}]$

然後append到maplist上。

2.第一個reducer：

依據Key來reduce，因為 $M_{ij}$ 跟 $N_{jk}$ 必須相乘為 $P_{ik}$ (所有的 $P_{ik}$ 會在最後相加)。

直接將Value相加，會以字串的方式接起來(為了給第二個mapper處理)。

3.第二個mapper：

Key :  $(i, k)$ ，剛剛的資料型態為串接起來並且已經以 $j$ 來分類，所以只要找到任何一對 $M$ 與 $N$ (利用雙層迴圈來找)，就可以找到相對應的 $i$ 跟 $k$ 。(在 $M$ 與 $N$ 的後面一項)

Value:  $(M_{ij} * N_{jk})$ ，剛剛的資料型態為串接起來並且已經以 $j$ 來分類，所以只要找到任何一對 $M$ 與 $N$ (利用雙層迴圈來找)，就可以找到相對應的 $M_{ij}$ 跟 $N_{jk}$ 。(在 $M$ 與 $N$ 的後面兩項)

然後append到maplist上

4.第二個reducer:

依據Key來reduce，直接將Value相加就可以得到最後結果。(與第一個reducer其實是同一個)  
最後結果的型態: $((i, k), \text{Value})$

5.第三個mapper：

這個mapper用來write file，這樣比較容易符合題目輸出的格式。

小心得：利用list的形式來設計Value比較容易取值、mapper用來write file是一個好方法。