

4.2.5.4 FP Algorithm - Q1

Q (1) FP min sup = 2

(2) Apriori

TID	Items
1	{A,B}
2	{B,C,D}
3	{A,C,D,E}
4	{A,D,E}
5	{A,B,C}
6	{A,B,C,D}
7	{A}
8	{A,B,C}
9	{A,B,D}
10	{B,C,E}

Solution ① Scan DB + remove.

itemset	σ
A	8
B	7
C	6
D	5
E	3

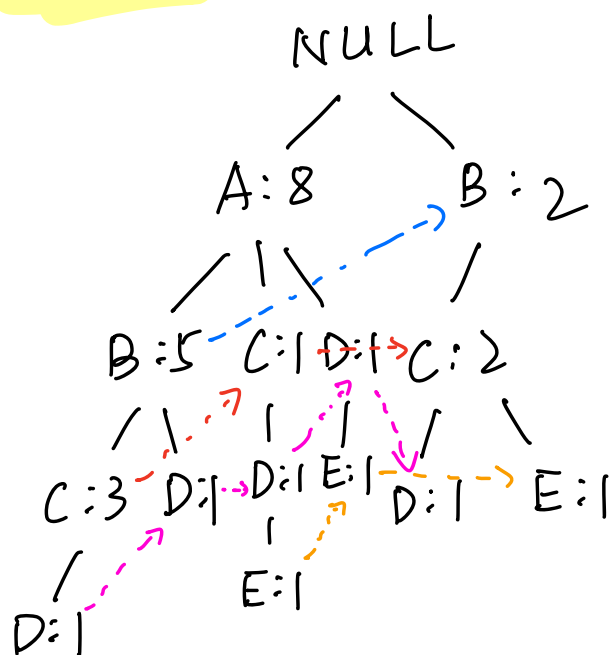
↑ look for
CFP - E
CFP - D
CFP - C
CFP - B
CFP - A

TID	Items
1	{A,B}
2	<u>{B,C,D}</u>
3	{A,C,D,E}
4	{A,D,E}
5	<u>{A,B,C}</u>
6	<u>{A,B,C,D}</u>
7	{A}
8	{A,B,C}
9	{A,B,D}
10	{B,C,E}

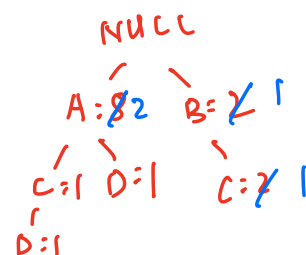
FP: $\{A=8\} \{B=7\} \{C=6\} \{D=5\} \{E=3\}$

② sorted frequency list

③ Construct FP-Tree.



TID	Items
1	{A,B} ✓
2	{B,C,D} ✓
3	{A,C,D,E} ✓
4	{A,D,E} ✓
5	{A,B,C} ✓
6	{A,B,C,D} ✓
7	{A} ✓
8	{A,B,C} ✓
9	{A,B,D} ✓
10	{B,C,E} ✓



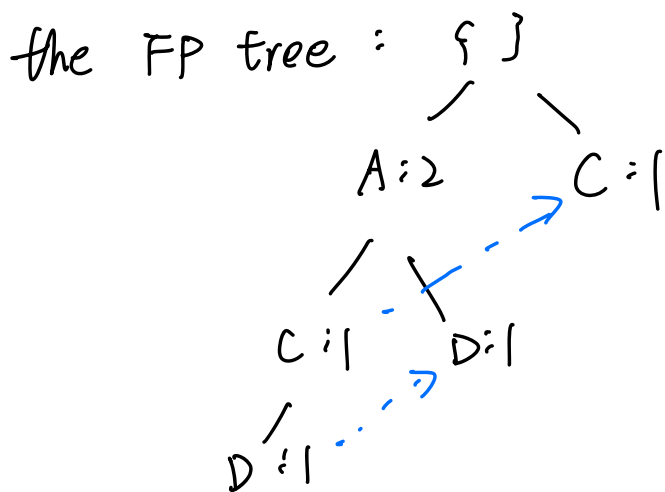
④ Conditional - Pattern - Base of E

$\{ ACD:1, AD:1, BC:1 \}$

包含 prefix tree
+ 节点..update

$\Rightarrow \{ A:2, C:2, D:2, B:1 \}$

Sorted. $\{ ACD:1, AD:1, C:1 \}$



FP: $\{ DE=2 \} \{ CE=2 \} \{ AE=2 \}$

So look for

CPB - DE

CPB - CE

CPB - AE

⑤ CPB - DE

$\{AC:1, A:1\}$

$\Rightarrow \{A:2, \cancel{C:1}\}$

Sorted $\{A:2\}$

FP Tree: $\{\}$
|
A:2

So FP: $\{ADE=2\}$

⑥ CPB - CE

$\{\cancel{A:1}\}$

$\Rightarrow \{\} \Rightarrow$ no FP-Tree

⑦ CPB - AE

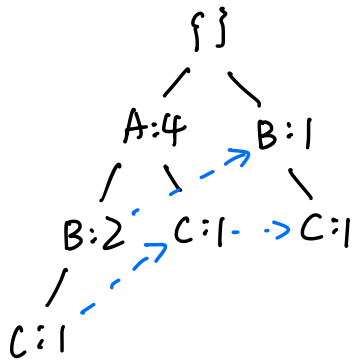
$\{\}$

⑧ CFP - D

CPB - D : $\{ \underset{\triangle}{A} \underset{\triangle}{B} \underset{\triangle}{C} : 1, \underset{\triangle}{A} \underset{\triangle}{B} : 1, \underset{\triangle}{A} \underset{\triangle}{C} : 1, \underset{\triangle}{A} : 1, \underset{\triangle}{B} \underset{\triangle}{C} : 1 \}$

$\Rightarrow \{A:4, B:3, C:3\}$

\Rightarrow CFP - D



look for

CFP - CD

CFP - BD

CFP - AD

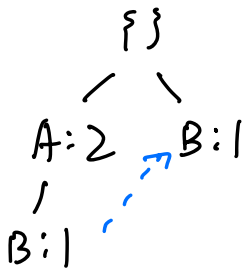
FP: {AD=4} {BD=3} {CD=3}

⑨ CFP - CD

CPB: {AB:1, A:1, B:1}

\Rightarrow {A:2, B:2}

CPF - CD



look for

CPF - BCD

CPF - ACD

FP: {BCD=2} {ACD=2}

⑩ CFP - BD

CPB: {A:2} \Rightarrow {A:2}

CFD - BD



FP: {ABD=2}

⑪ CFP-AD

CPB: { } \Rightarrow { } \Rightarrow NO FP Tree

⑫ CFP-BCD

CPB: { ~~A:1~~ } \Rightarrow { } \Rightarrow NO FP Tree

⑬ CFP-ACD

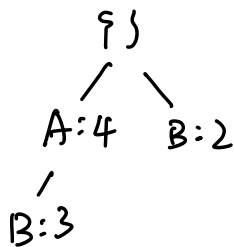
CPB: { } \Rightarrow { } \Rightarrow NO FP Tree.

⑭ CFP-C

CPB: { AB:3, A:1, B:2 }

\Rightarrow { A:4, B:1 }

CFP-C

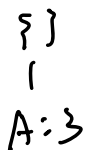


look for
CFP-BC
CFP-AC

FP: { AC=4, BC=1 }

⑮ CFP-BC

CPB: { A:3 } \Rightarrow { A:3 }



FP: { ABC=3 }

⑯ CFP-AC

CPB: { 3 } \Rightarrow { } \Rightarrow NO

⑪ CFP - B

$$CPB: \{A:5\} \Rightarrow \{A:5\} \quad \begin{matrix} \{ \} \\ | \\ A:5 \end{matrix}$$

$$FP: \{AB=5\}$$

⑫ CFP - A

$$CPB = \{ \} = \{ \} \Rightarrow NO.$$

All in All

$$FP \{A=8\} \{B=7\} \{C=6\} \{D=5\} \{E=3\} \quad 5$$

$$\left. \begin{array}{l} \{DE=2\} \{CE=2\} \{AE=2\} \\ \{ADE=2\} \end{array} \right\} \quad 4$$

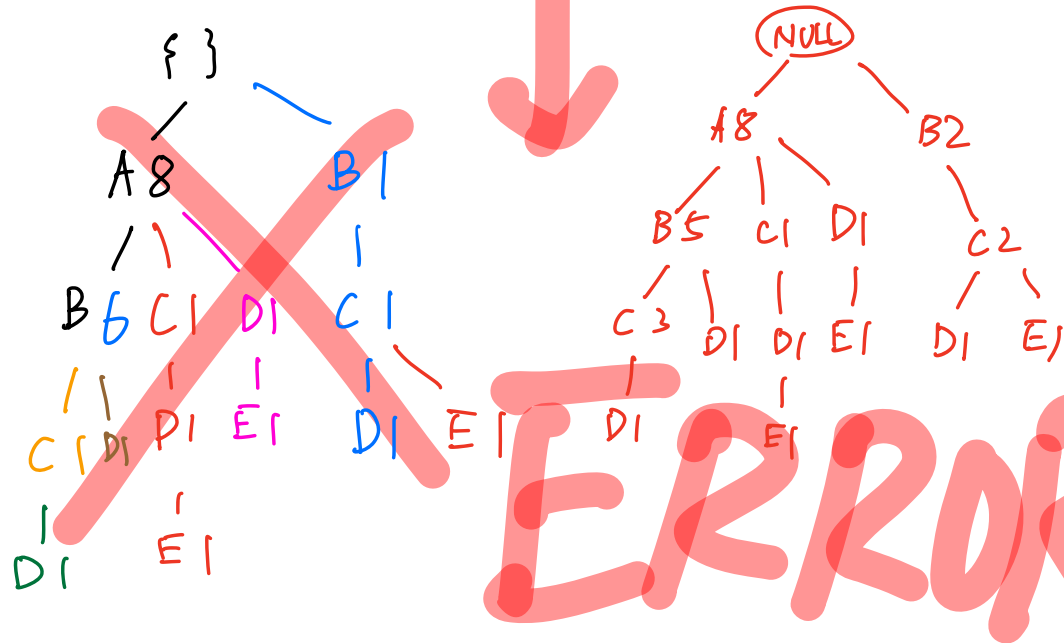
$$\left. \begin{array}{l} \{AD=4\} \{BD=3\} \{CD=3\} \\ \{BCD=2\} \{ACD=2\} \\ \{ABD=2\} \end{array} \right\} \quad 6$$

$$\left. \begin{array}{l} \{AC=4, BC=5\} \\ \{ABC=3\} \end{array} \right\} \quad 3$$

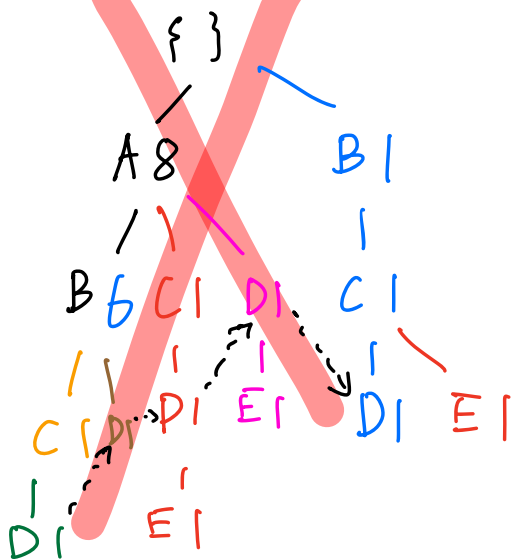
$$\{AB=5\} \quad 1$$

$$5 + 4 + 6 + 3 + 1 = 19$$

② Fp Tree



③ CPB : D



CPB of E

$\{ACD, AD:1, BC:1\}$

CPB-E 1-FP: $\{A_2, C_2, D_2, B_1\}$

Sorted CPB-E $\{ACD:1, AD:1, C:1\}$



④ conditional pattern base for D

$PB = \{ (A:1, B:1, C:1) \}$

$(A:1, B:1)$

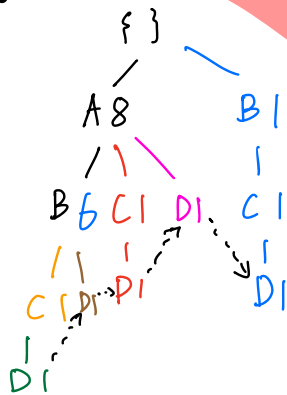
$(A:1, C:1)$

$(A:1)$

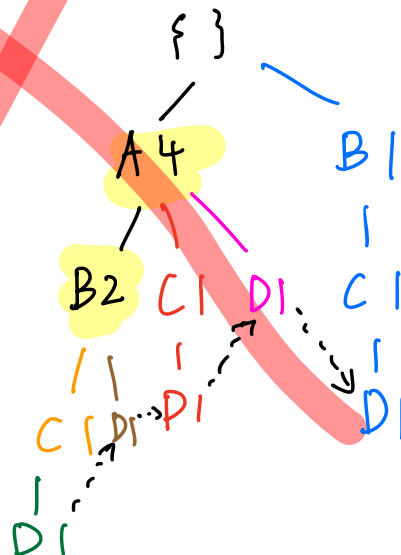
$(B:1, C:1) \}$

sup of D = 5 \geq 2 So {D} is freq. itemset

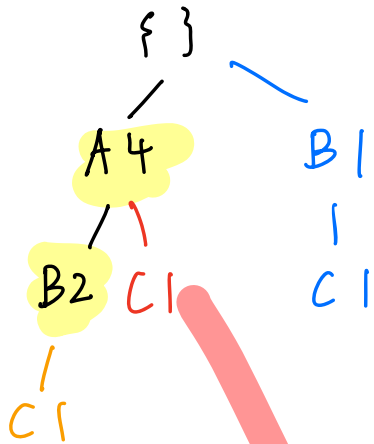
⑤ find D. CPT (a) prefix tree.



(b) update count



(c) delete D node



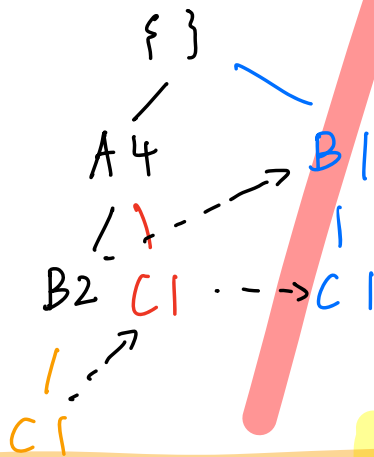
cd) delete no f.i.

$$A: 4 \geq 2$$

$$B: 3 \geq 2$$

$$C: 3 \geq 2$$

ce) d condition tree is

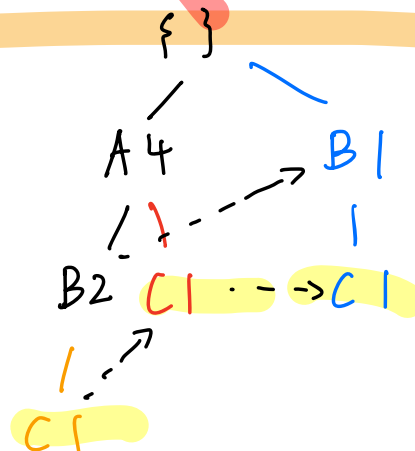


So we need to find
 (cd) (bd) (ad) Question

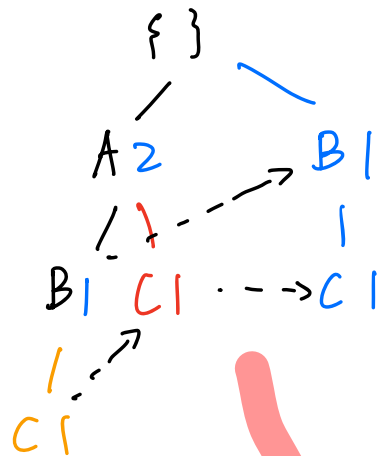
(6) prefix path of c

sup of cd: $3 \geq 2$

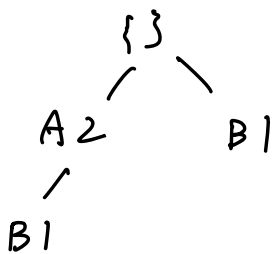
So fcd is f.i



b) update node



c) delete C

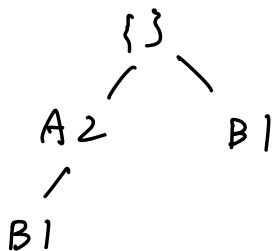


d) delete sup

$$A:2 \geq 2$$

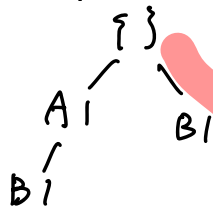
$$B:2 \geq 2$$

e) cd : condition FP tree.

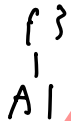


⑦ {bcd} sup : 2 \geq 2 So {bcd} is f.i

(a) update



(b) delete B



(c) delete sup



$A: 1 < 2$ delete A

(d) $\{BCD\}$

CFP

That's why $abcd$ isn't frequently it's a set



⑧ consider $\{ACD\}$ prefix path from cd CFP



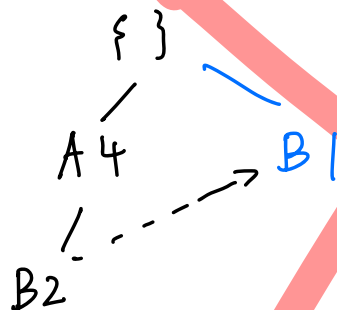
$A: 2$

$A: 2 \geq 2$ So ACD is f. i

CD

BD

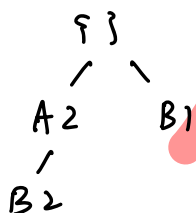
⑨ find BD prefix path from D CFP



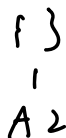
$B: 3 \geq 2$

So BD is f. i.

(a) update



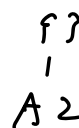
(b) delete B



(c) sup

$A: 2 \geq 2$

(d) BD CFP



So $\{A BD\}$ is f. i

BD

AD

⑩ find AD prefix path from D CPF

{ }
1
A4

$A:4 \geq 2$

so AD is f.i

⑪ Frequent Itemset found

D AD BD CD ABD ACD BCD