

对各商品进行编号如下:

Milk: T₁, Banana: T₂, Bread: T₃,

Coffee: T₄, Coke: T₅, Pie: T₆,

Orange: T₇, Cheese: T₈, Cake: T₉

①原始表格转换为:

Tid	Items
1	T ₁ , T ₂ , T ₃ , T ₄ , T ₅
2	T ₄ , T ₁ , T ₃ , T ₆
3	T ₇ , T ₃ , T ₄ , T ₈
4	T ₁ , T ₃ , T ₄ , T ₈
5	T ₁ , T ₃ , T ₈ , T ₆
6	T ₂ , T ₁ , T ₉

③再次扫描, 剔除非频繁项, 排序.

Tid	Items
1	T ₁ , T ₃ , T ₄
2	T ₁ , T ₃ , T ₄
3	T ₃ , T ₄ , T ₈
4	T ₁ , T ₃ , T ₄ , T ₈
5	T ₁ , T ₃ , T ₈
6	T ₁

②根据 FP-tree 算法, 找到频繁 1-项集并排序: (降序)

项集	支持度计数
T ₁	5
T ₂	2
T ₃	5
T ₄	4
T ₅	1
T ₆	2
T ₇	1
T ₈	3
T ₉	1

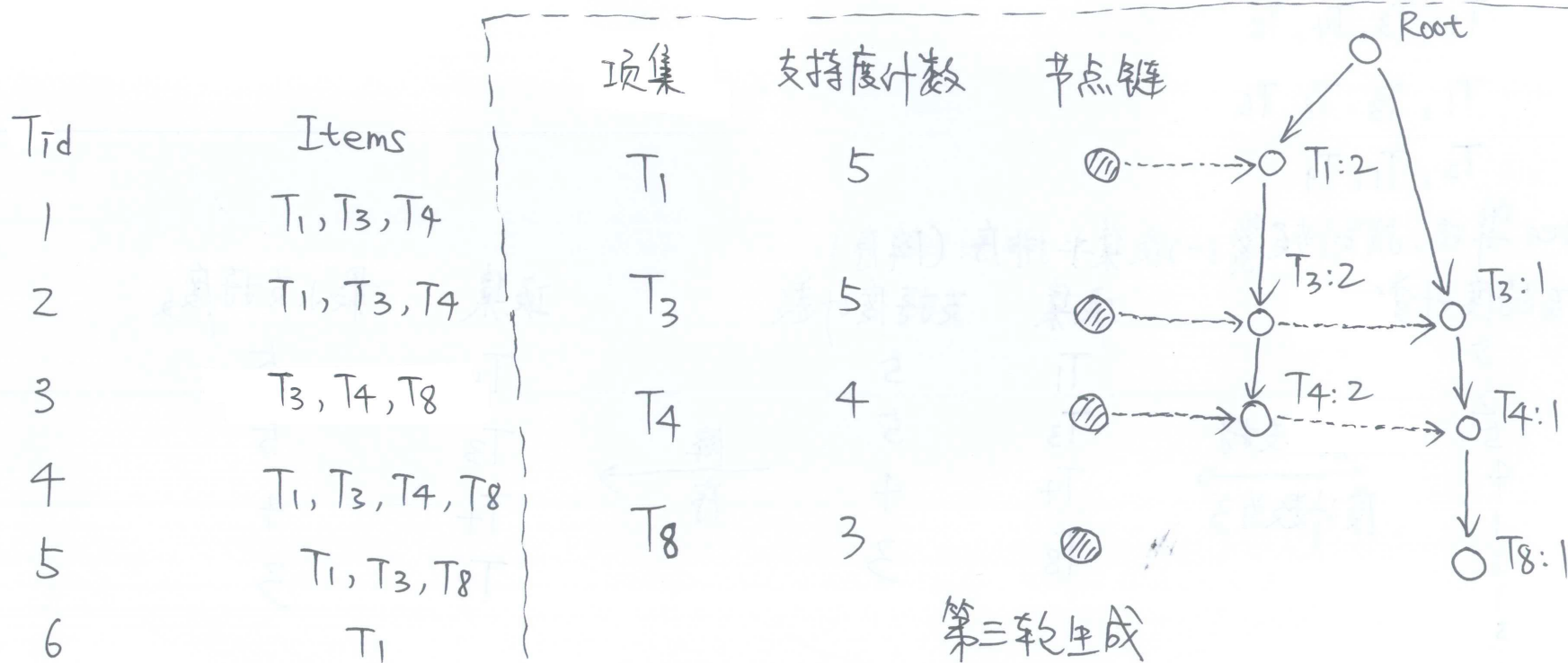
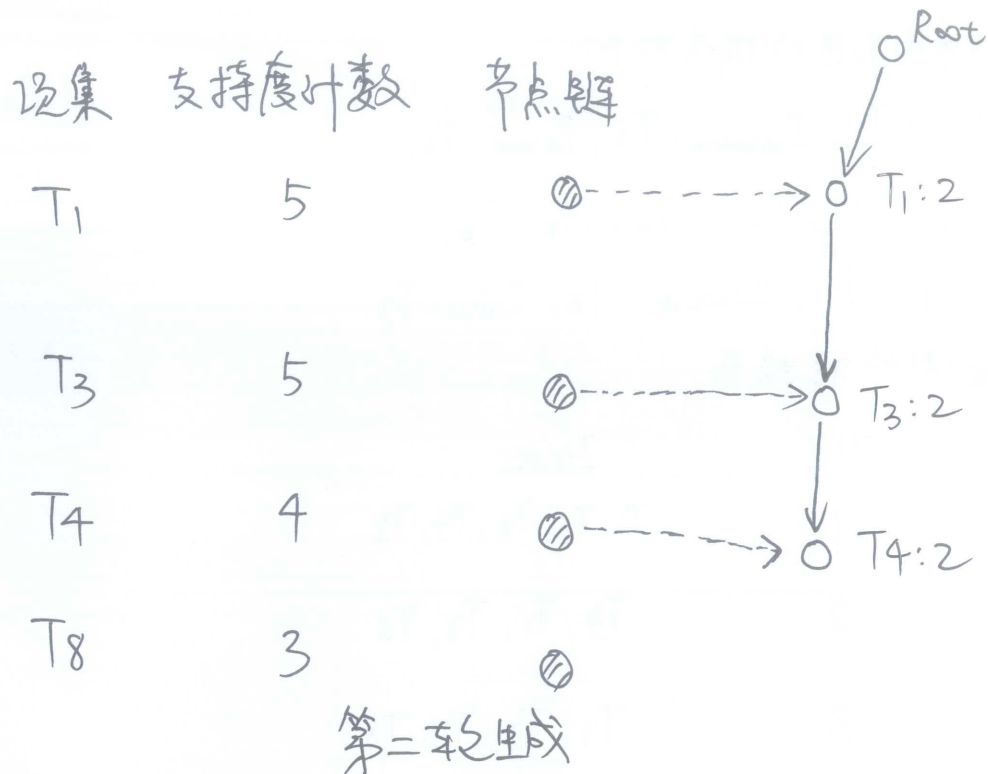
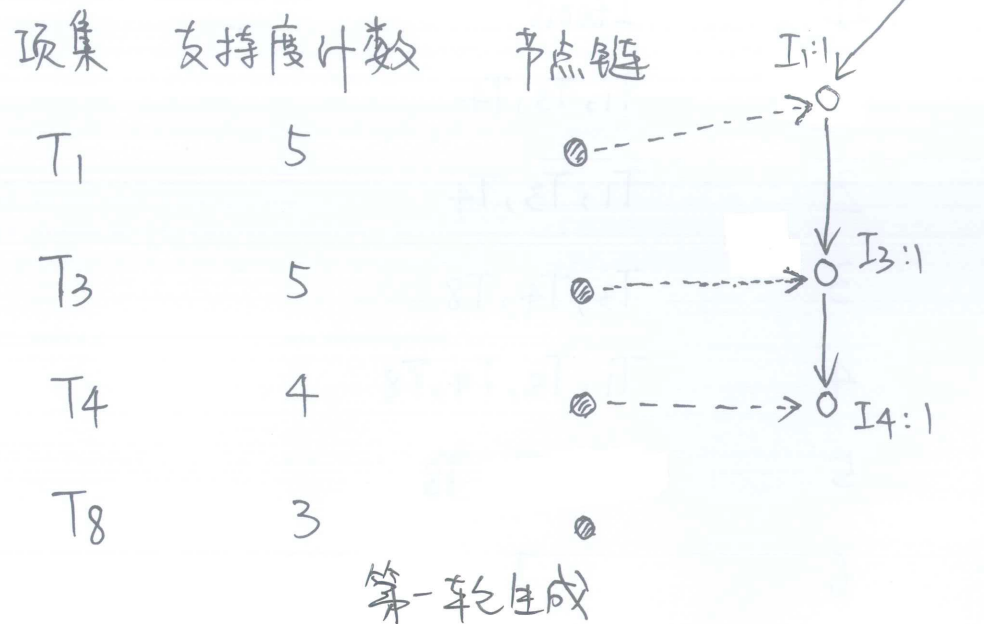
最小支持
度计数为3

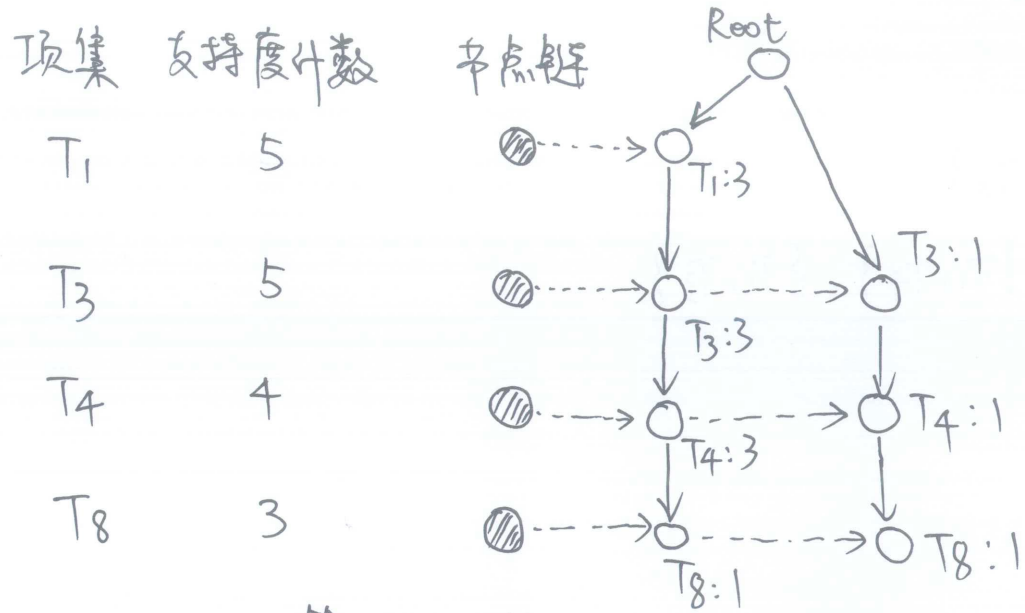
项集	支持度计数
T ₁	5
T ₃	5
T ₄	4
T ₈	3

降序

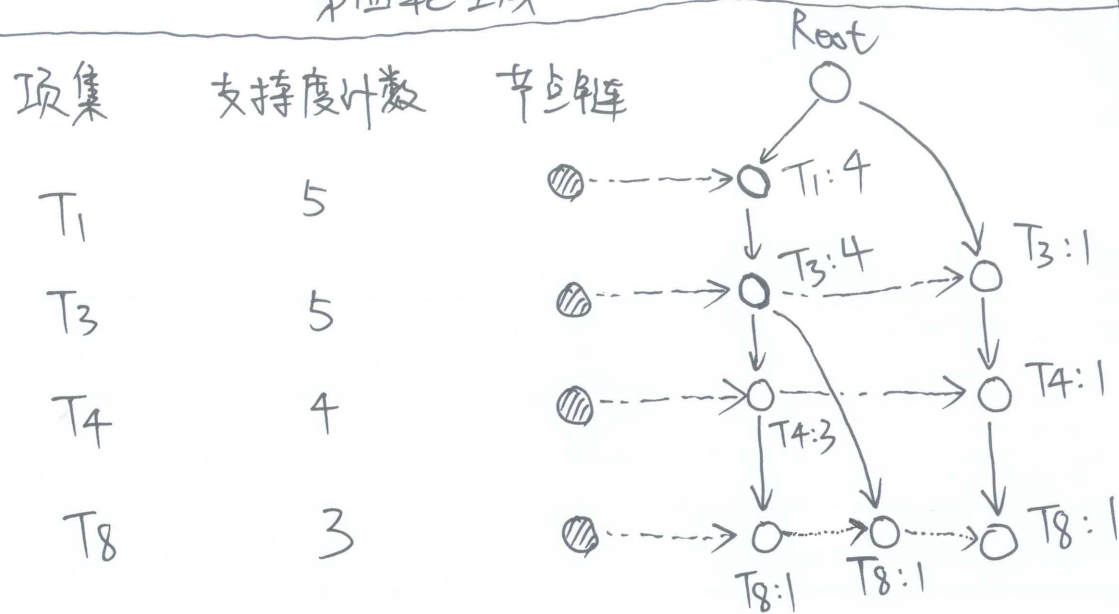
项集	支持度计数
T ₁	5
T ₃	5
T ₄	4
T ₈	3

④ 构建 FP-tree

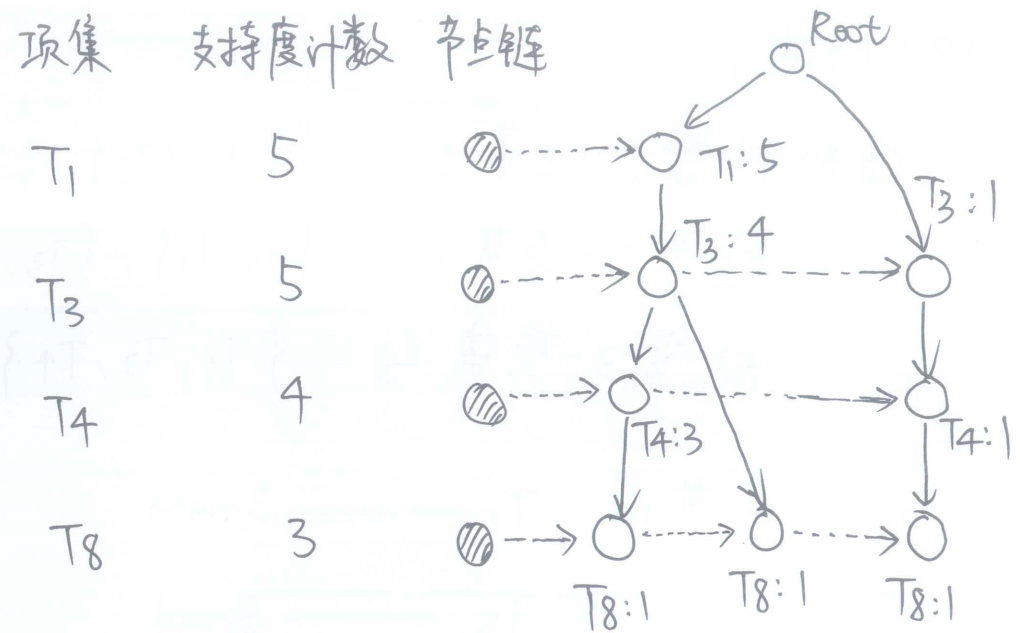




第四轮生成



第五轮生成



第六轮生成

后缀项	条件数据库	挖掘到的频繁项集
$\{T_8\}$	$\{T_1, T_3, T_4\}:1$	$\{T_3, T_8\}$
$\{T_4\}$	$\{T_1, T_3\}:3$ $\{T_3\}:1$	$\{T_3, T_4\}$ $\{T_1, T_4\}$ $\{T_1, T_3, T_4\}$
$\{T_3\}$	$\{T_1\}:4$	$\{T_1, T_3\}$
$\{T_1\}$		

综上所述, 使用FP-tree算法发现的频繁项集有:

频繁1-项集 L_1 : $\{T_1\}, \{T_3\}, \{T_4\}, \{T_8\}$

频繁2-项集 L_2 : $\{T_3, T_8\}, \{T_3, T_4\}, \{T_1, T_4\}, \{T_1, T_3\}$

频繁3-项集 L_3 : $\{T_1, T_3, T_4\}$

其中 T_1 为 Milk, T_3 为 Bread, T_4 为 Coffee, T_8 为 Cheese.