FP Growth Algorithm For Mining Frequent Pattern

Presented By:
Shihab Rahman
Dolon Chanpa
Department Of Computer Science And Engineering,
University of Dhaka

What is FP Growth?

- FP Growth Stands for frequent pattern growth
- It is a scalable technique for mining frequent pattern in a database

FP Growth

- FP growth improves Apriority to a big extent
- Frequent Item set Mining is possible without candidate generation
- Only “two scan” to the database is needed

BUT HOW?

FP Growth

- Simply a two step procedure
  - Step 1: Build a compact data structure called the FP-tree
    - Built using 2 passes over the data-set.
  - Step 2: Extracts frequent item sets directly from the FP-tree

FP Growth

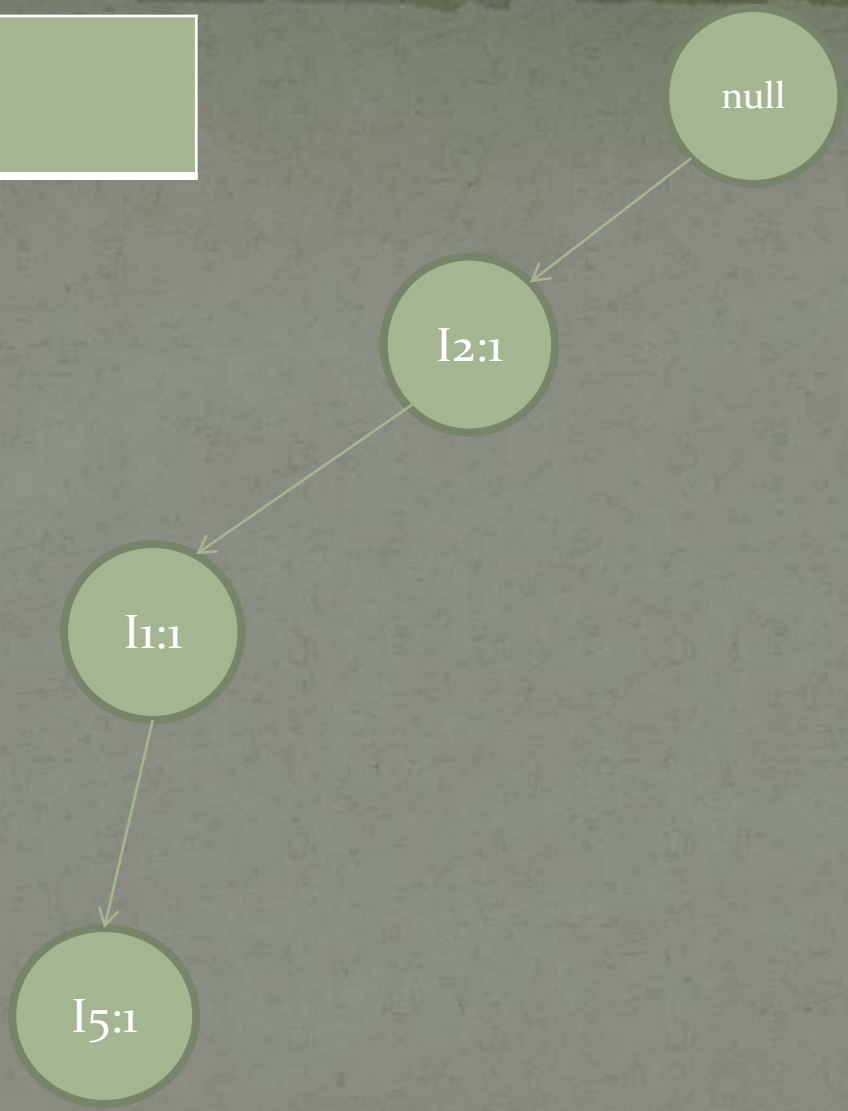
- Now Lets Consider the following transaction table

TID	List of item IDs
T100	I1,I2,I3
T200	I2,I4
T300	I2,I3
T400	I1,I2,I4
T500	I1,I3
T600	I2,I3
T700	I1,I3
T800	I1,I2,I3,I5
T900	I1,I2,I3

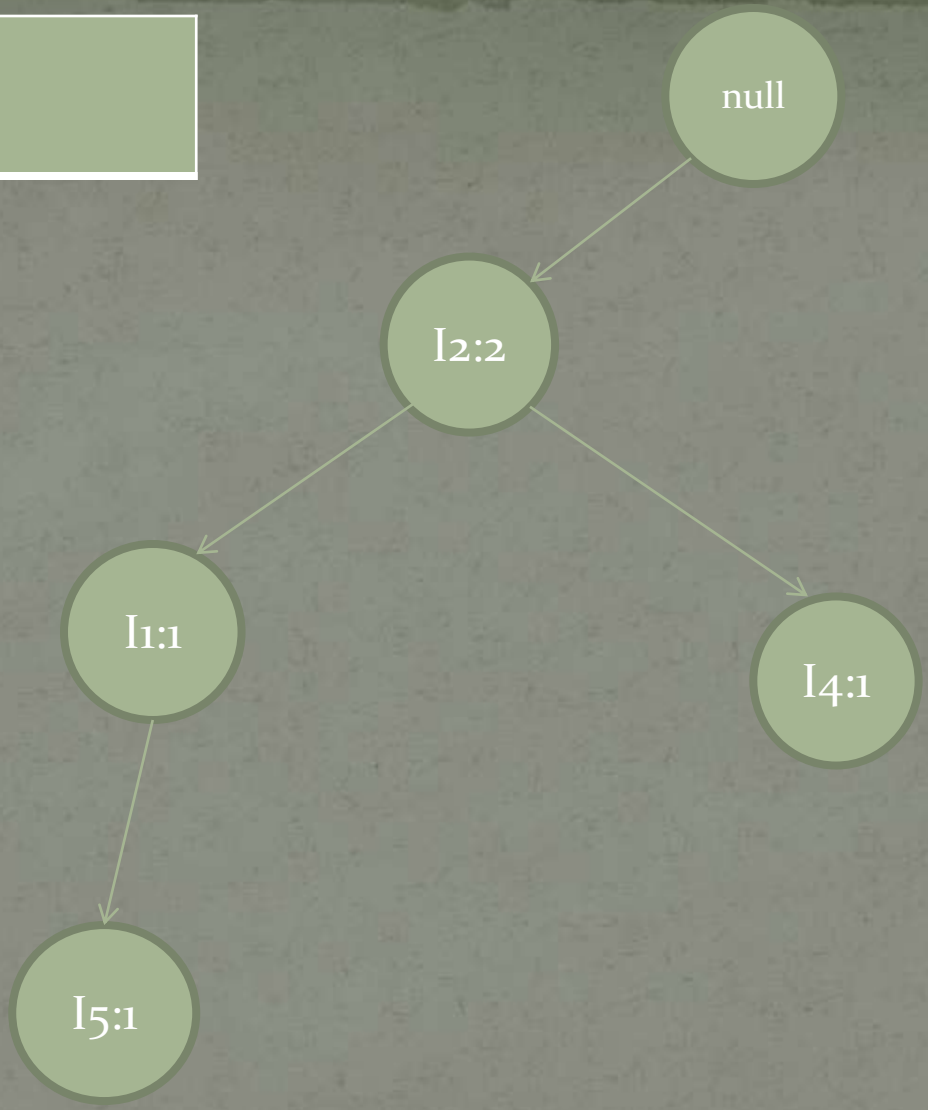
FP Growth

- Now we will build a FP tree of that database
- Item sets are considered in order of their descending value of support count.

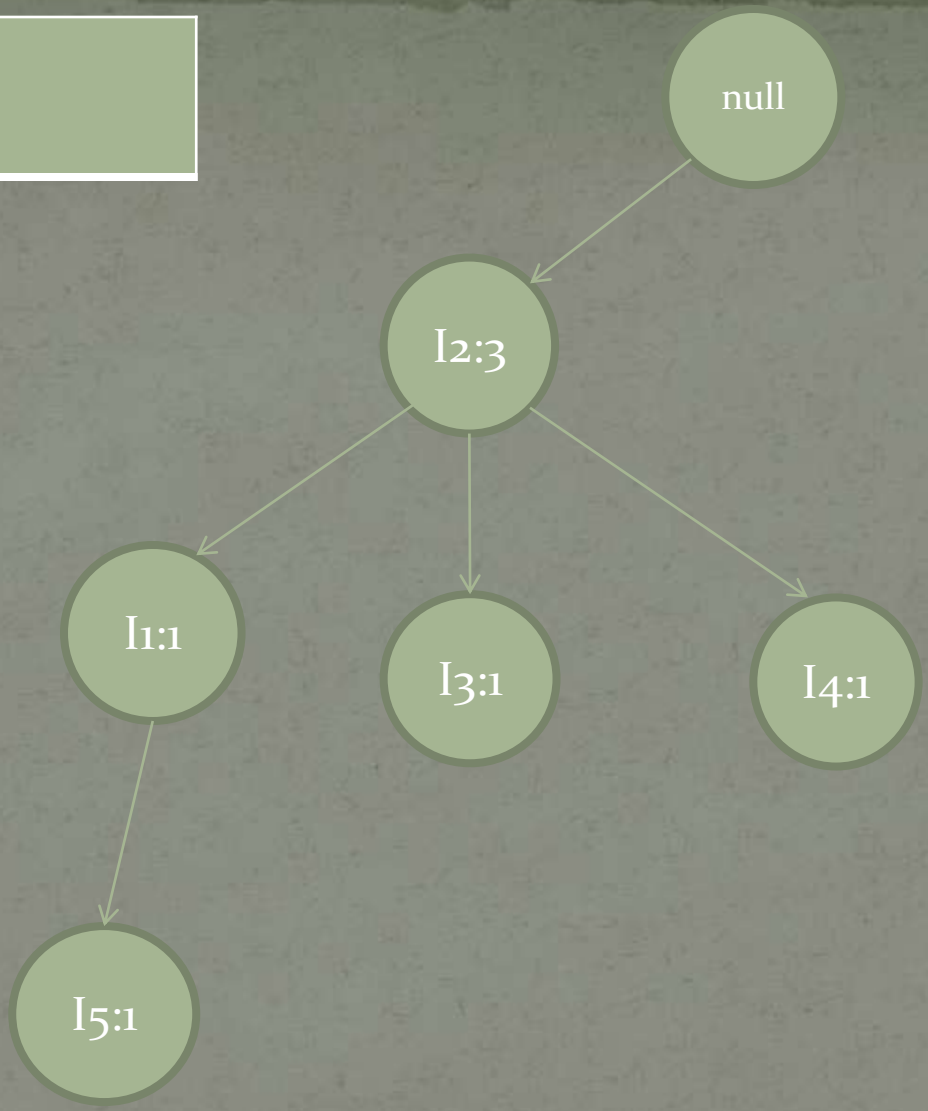
For Transaction:
I2,I1,I5



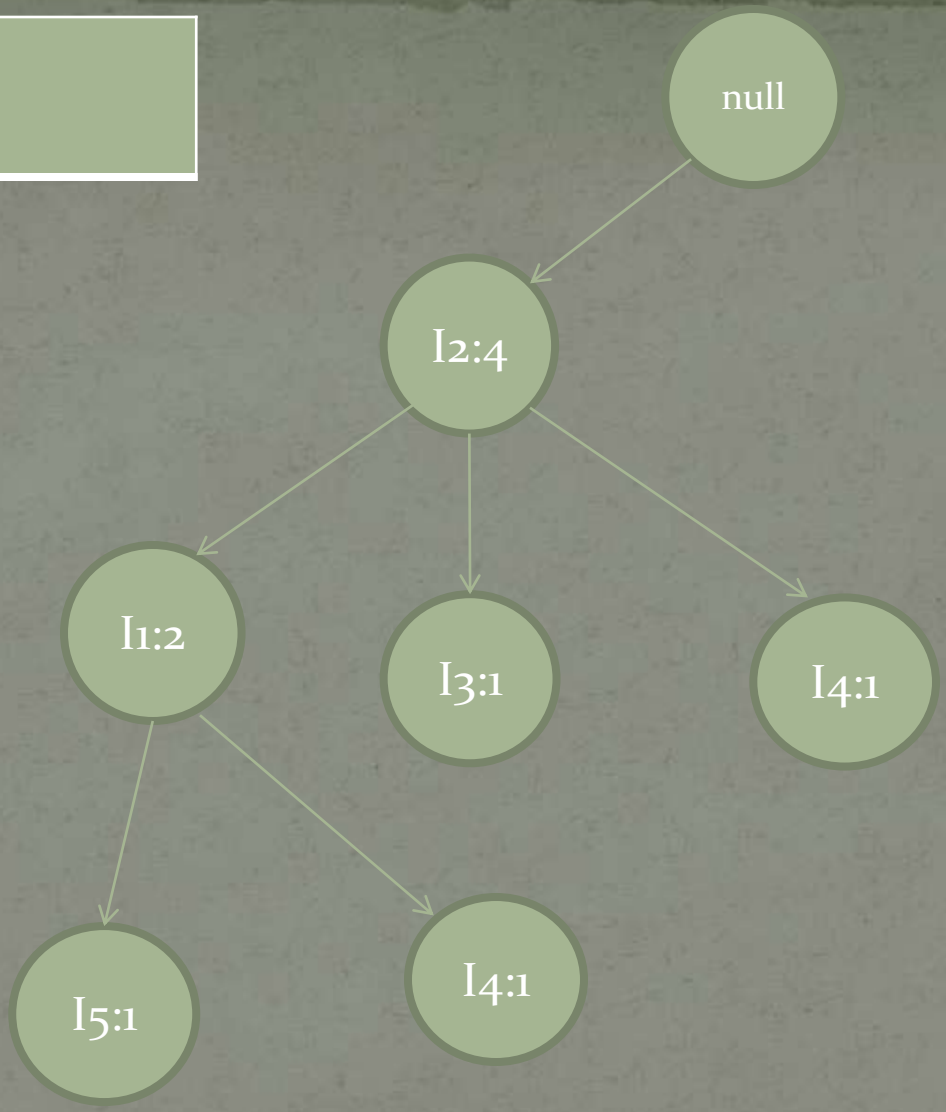
For Transaction:
I2,I4



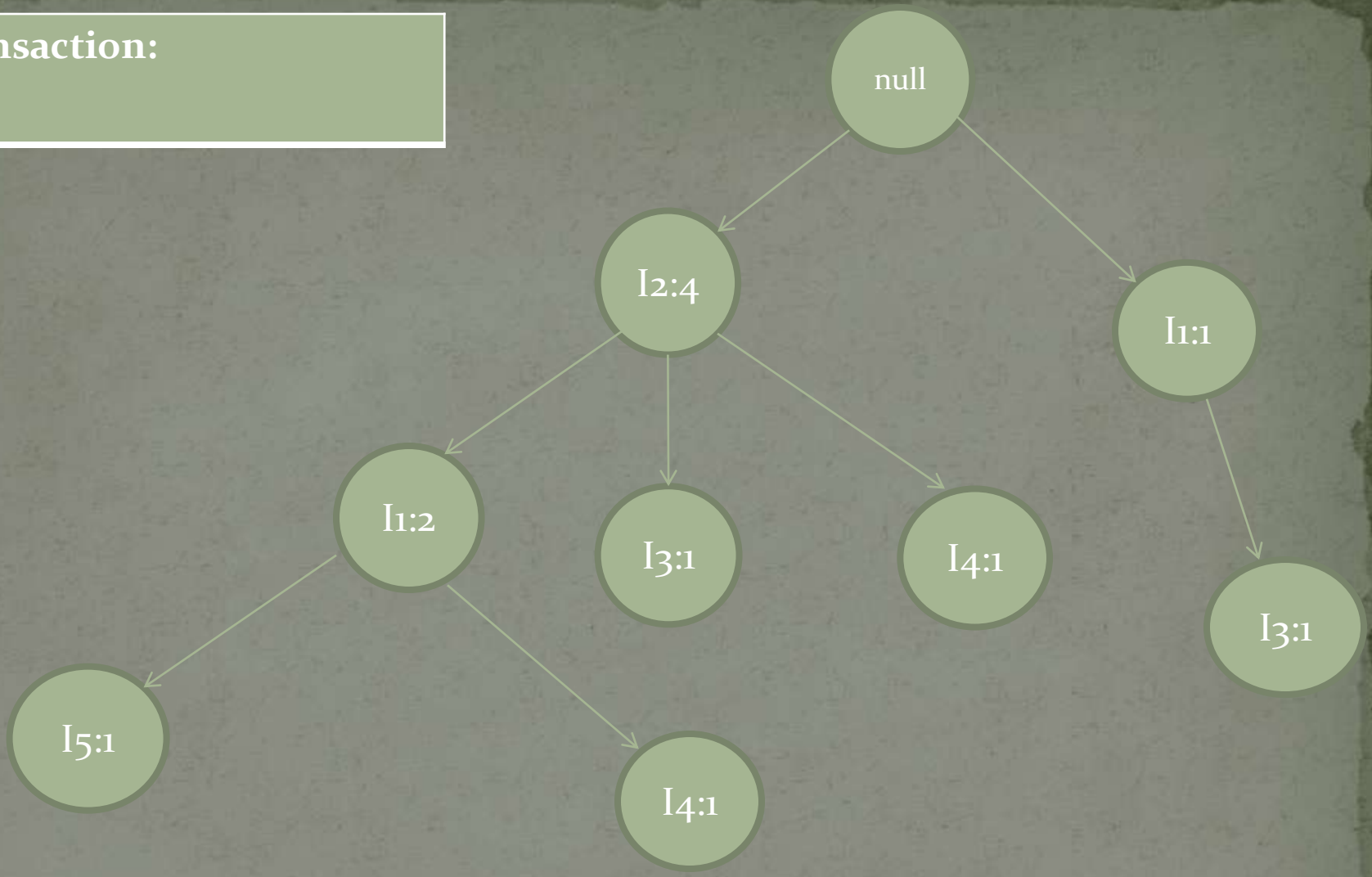
For Transaction:
I2,I3



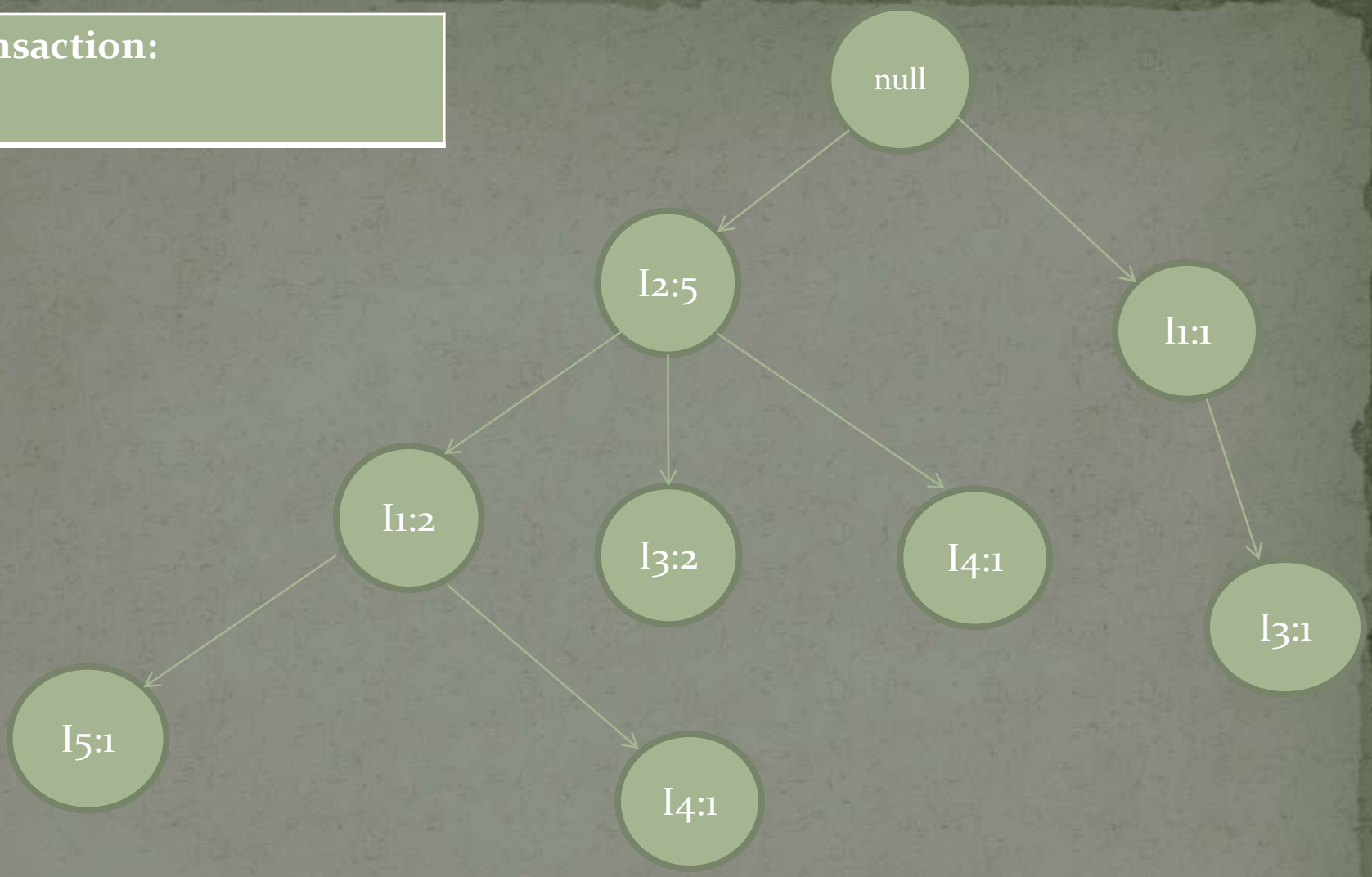
For Transaction:
I2,I1,I4



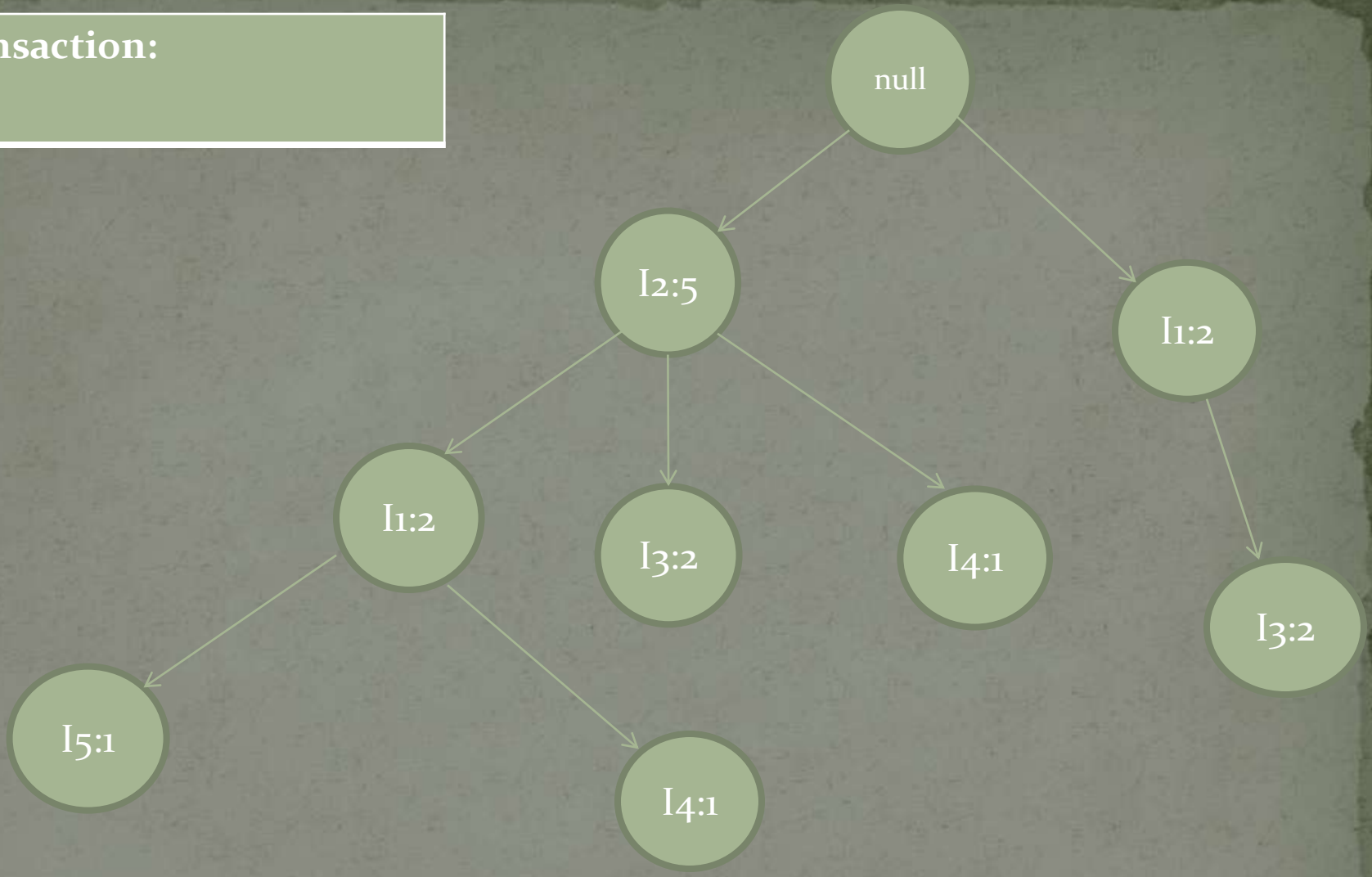
For Transaction:
I1,I3



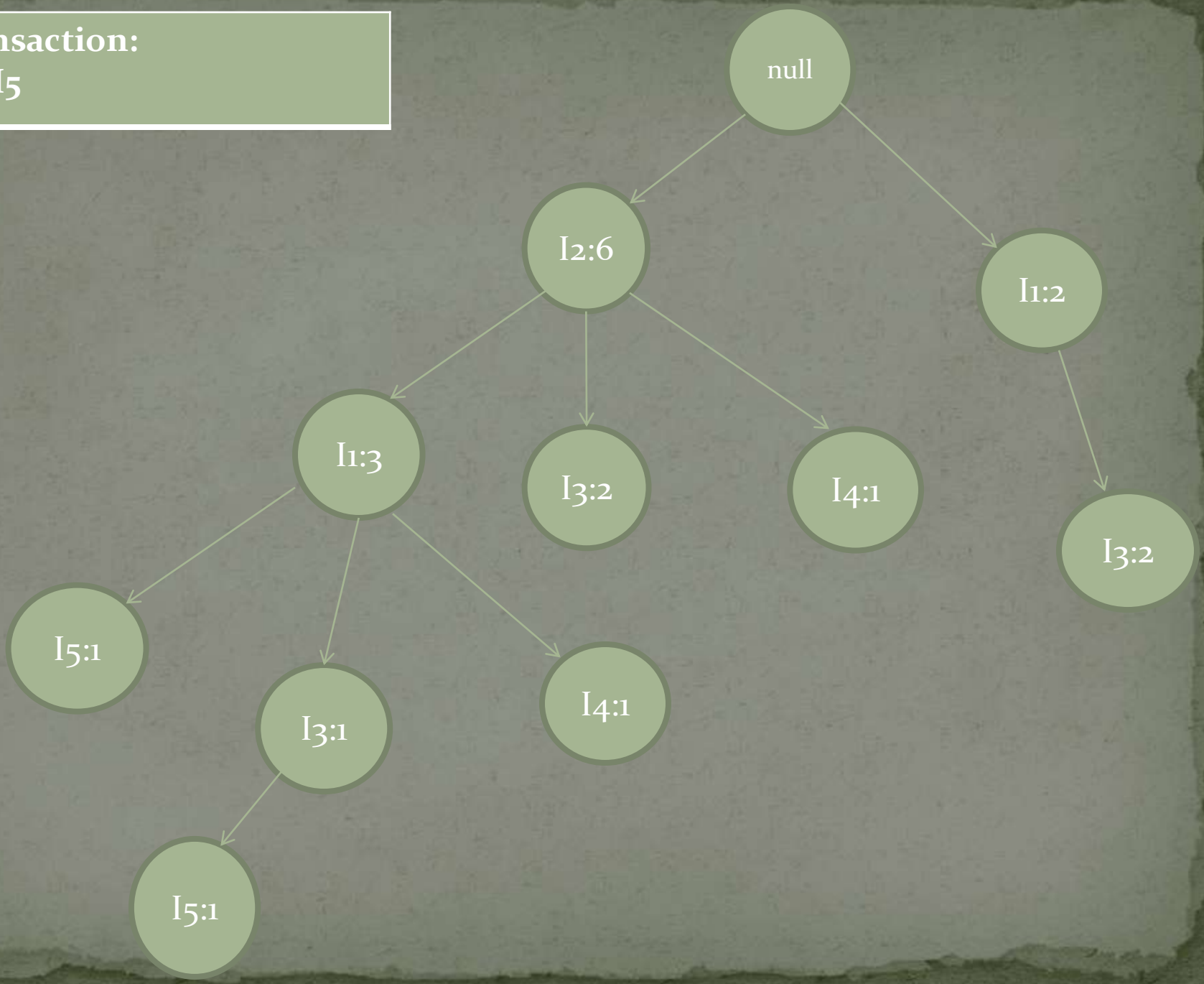
For Transaction:
I2,I3



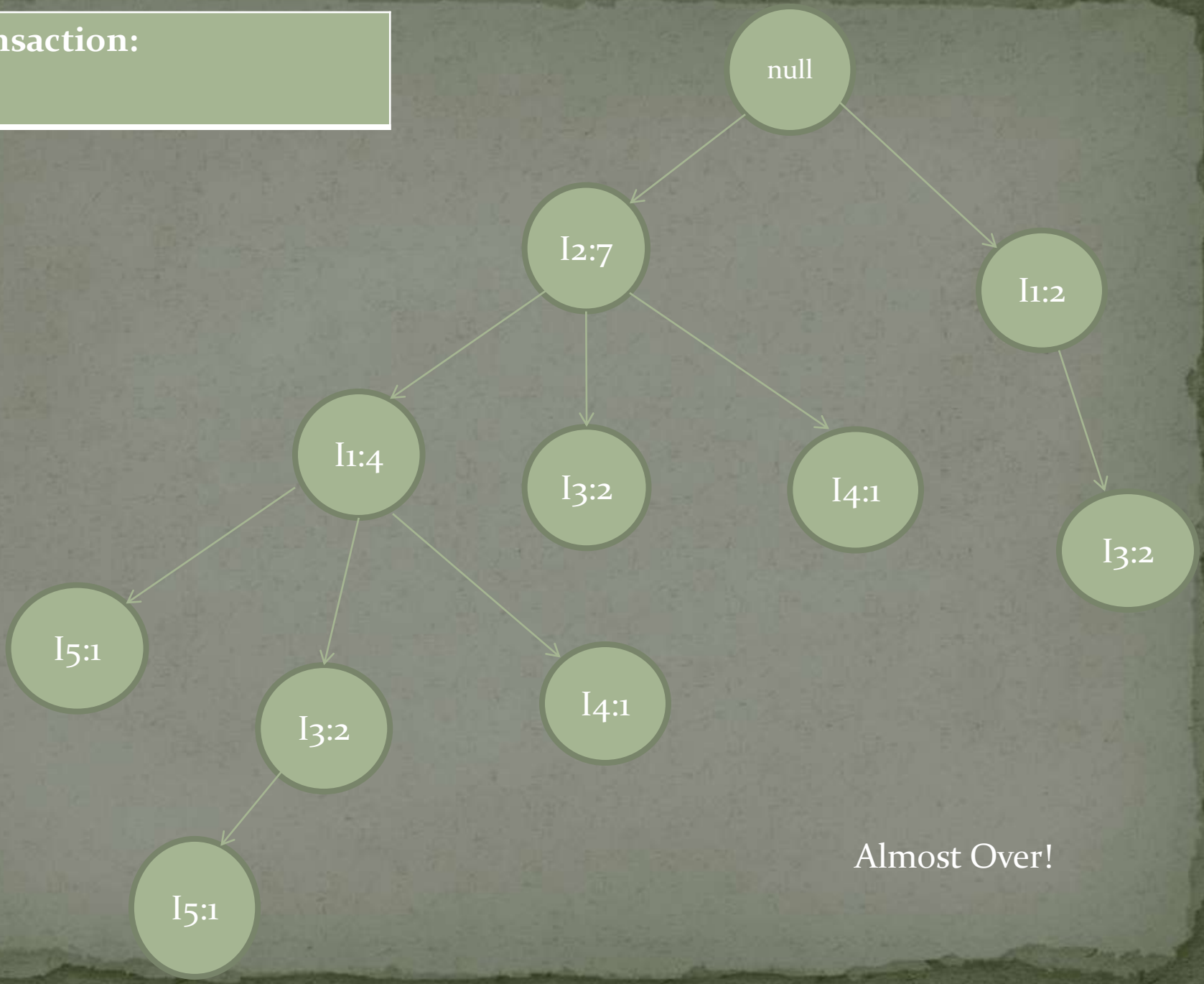
For Transaction:
I1,I3



For Transaction:
I2,I1,I3,I5



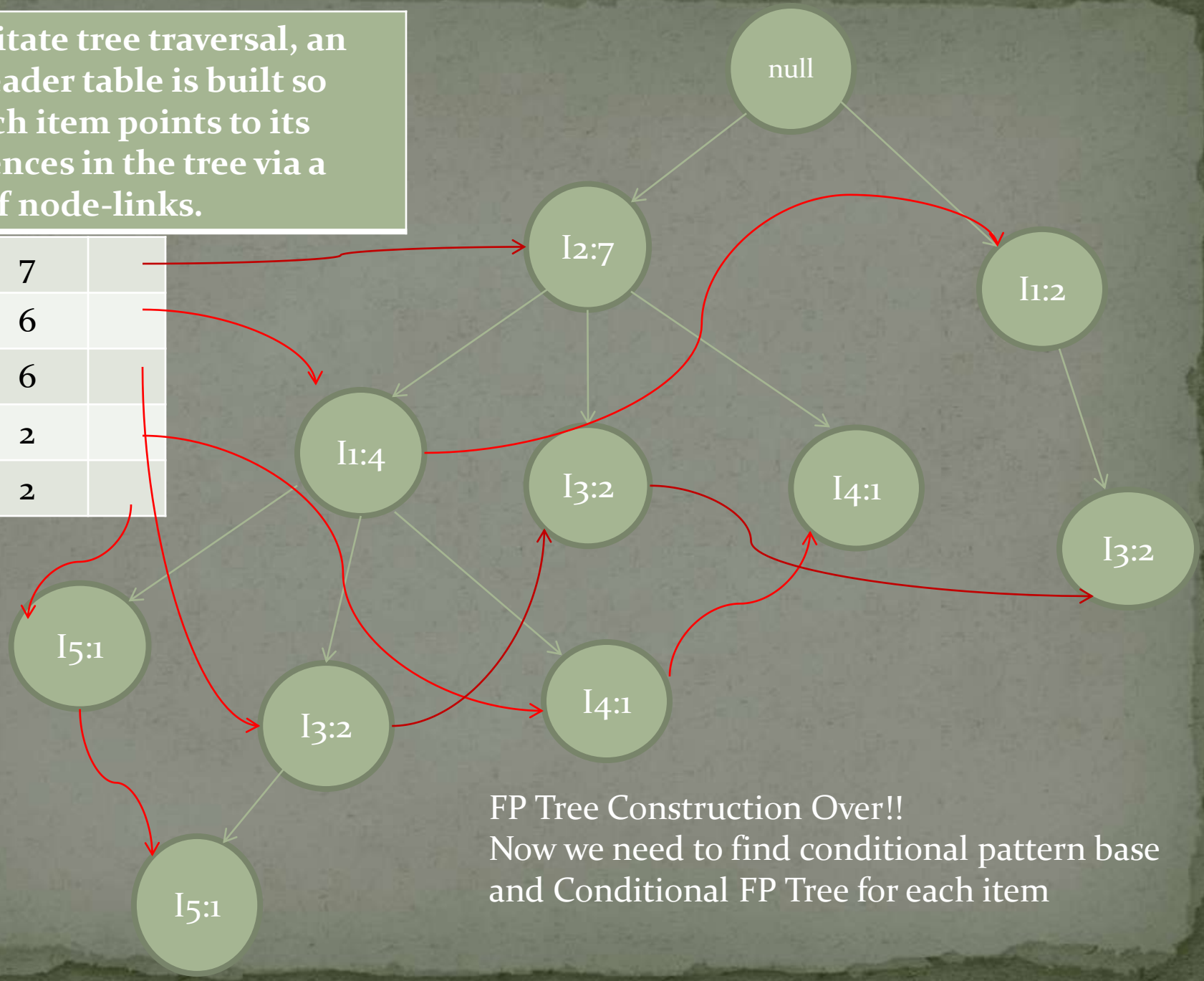
For Transaction:
I2,I1,I3



Almost Over!

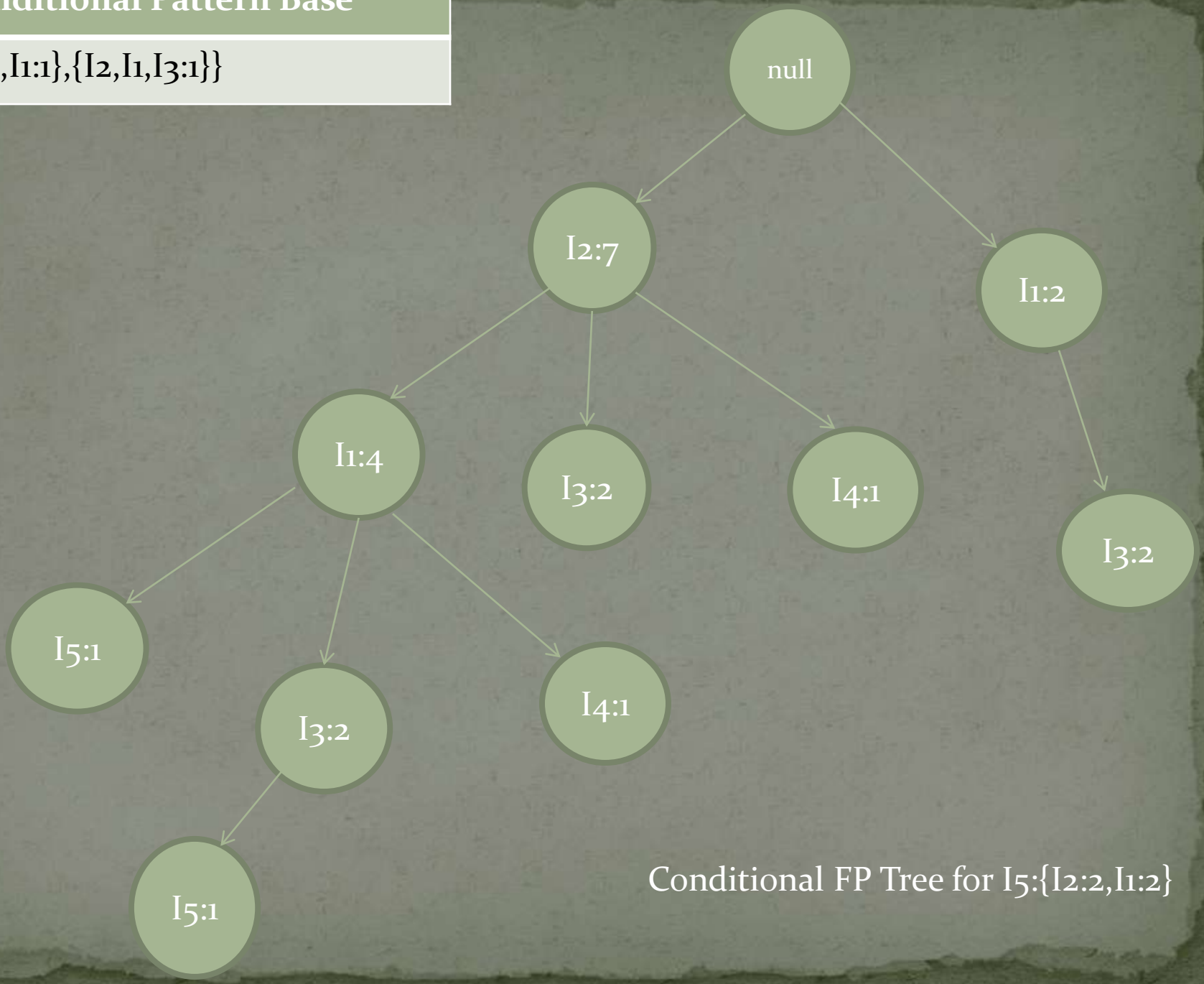
To facilitate tree traversal, an item header table is built so that each item points to its occurrences in the tree via a chain of node-links.

I2	7	
I1	6	
I3	6	
I4	2	
I5	2	



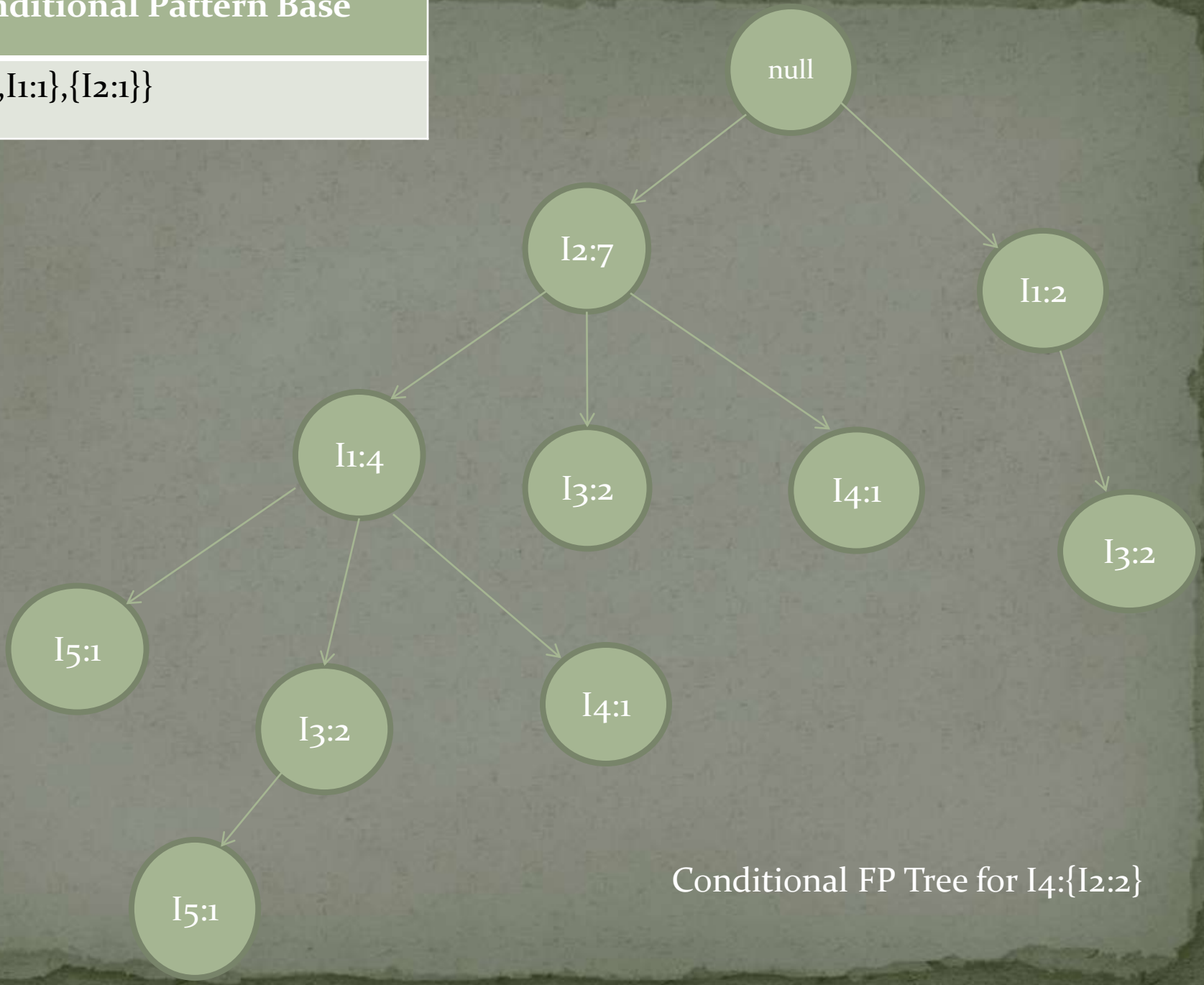
FP Tree Construction Over!!
Now we need to find conditional pattern base and Conditional FP Tree for each item

	Conditional Pattern Base
I5	{{I2,I1:1},{I2,I1,I3:1}}



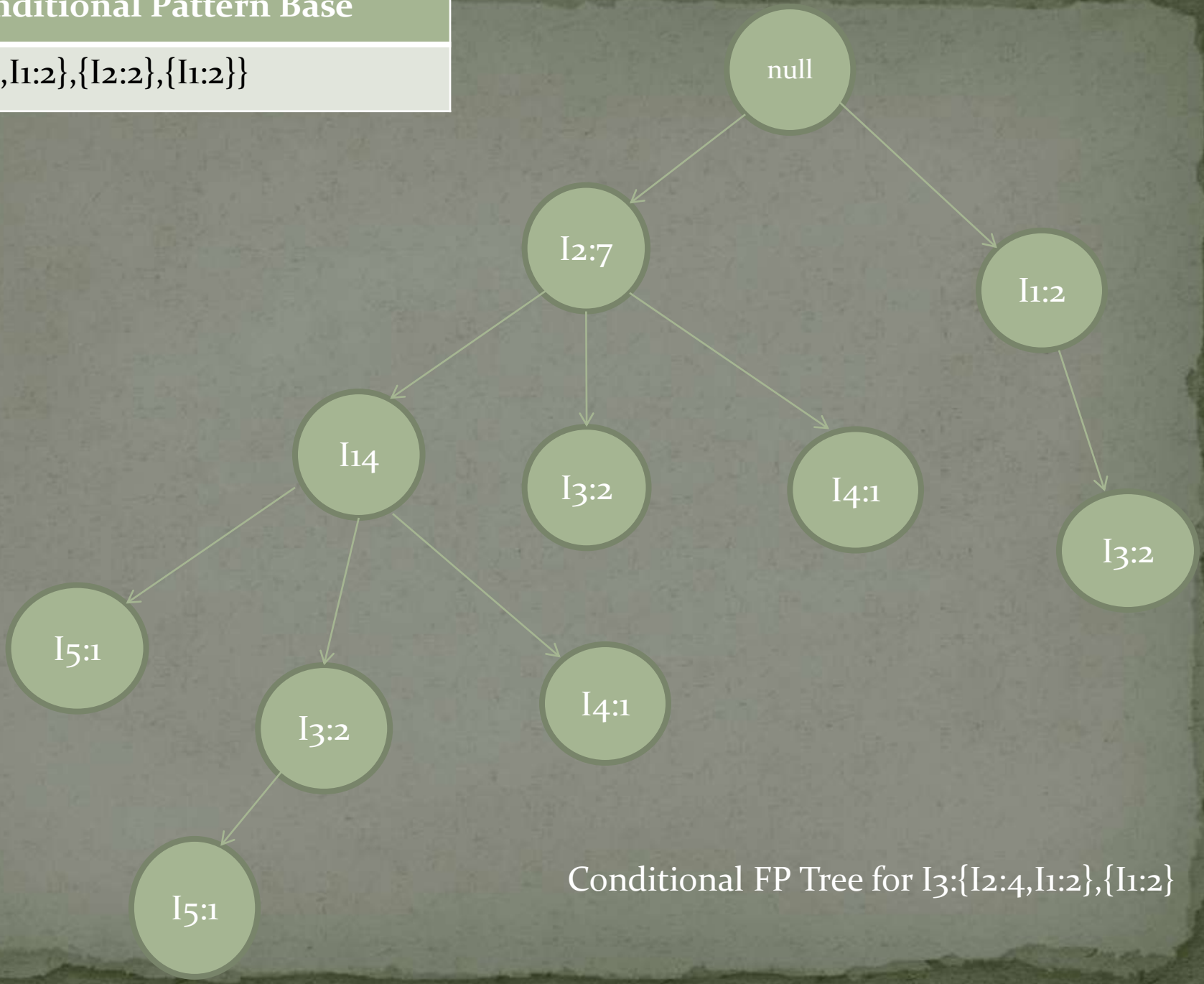
Conditional FP Tree for I5:{I2:2,I1:2}

	Conditional Pattern Base
I4	{{I2,I1:1},{I2:1}}



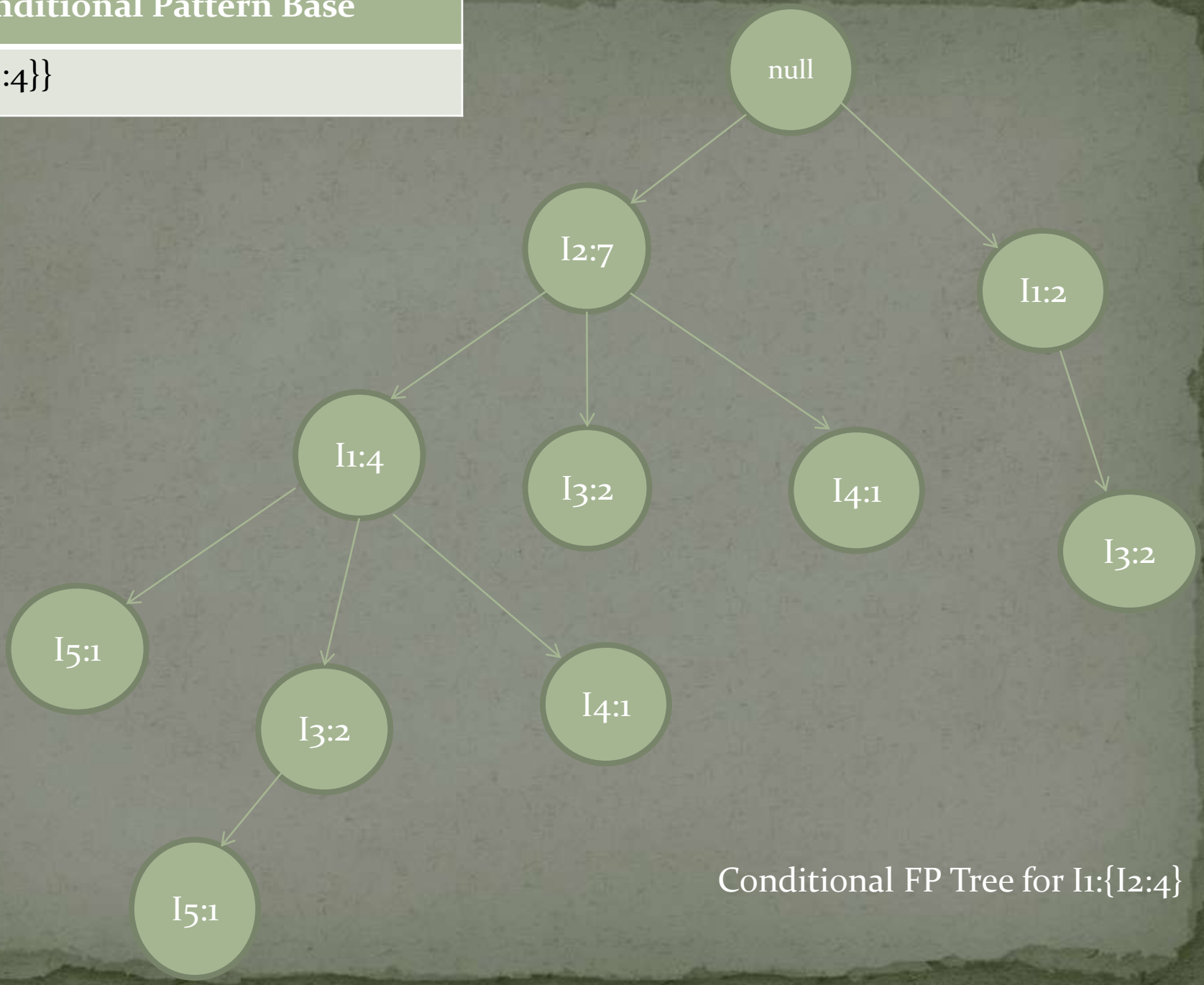
Conditional FP Tree for I4:{I2:2}

	Conditional Pattern Base
I3	{{I2,I1:2},{I2:2},{I1:2}}



Conditional FP Tree for I3:{I2:4,I1:2},{I1:2}

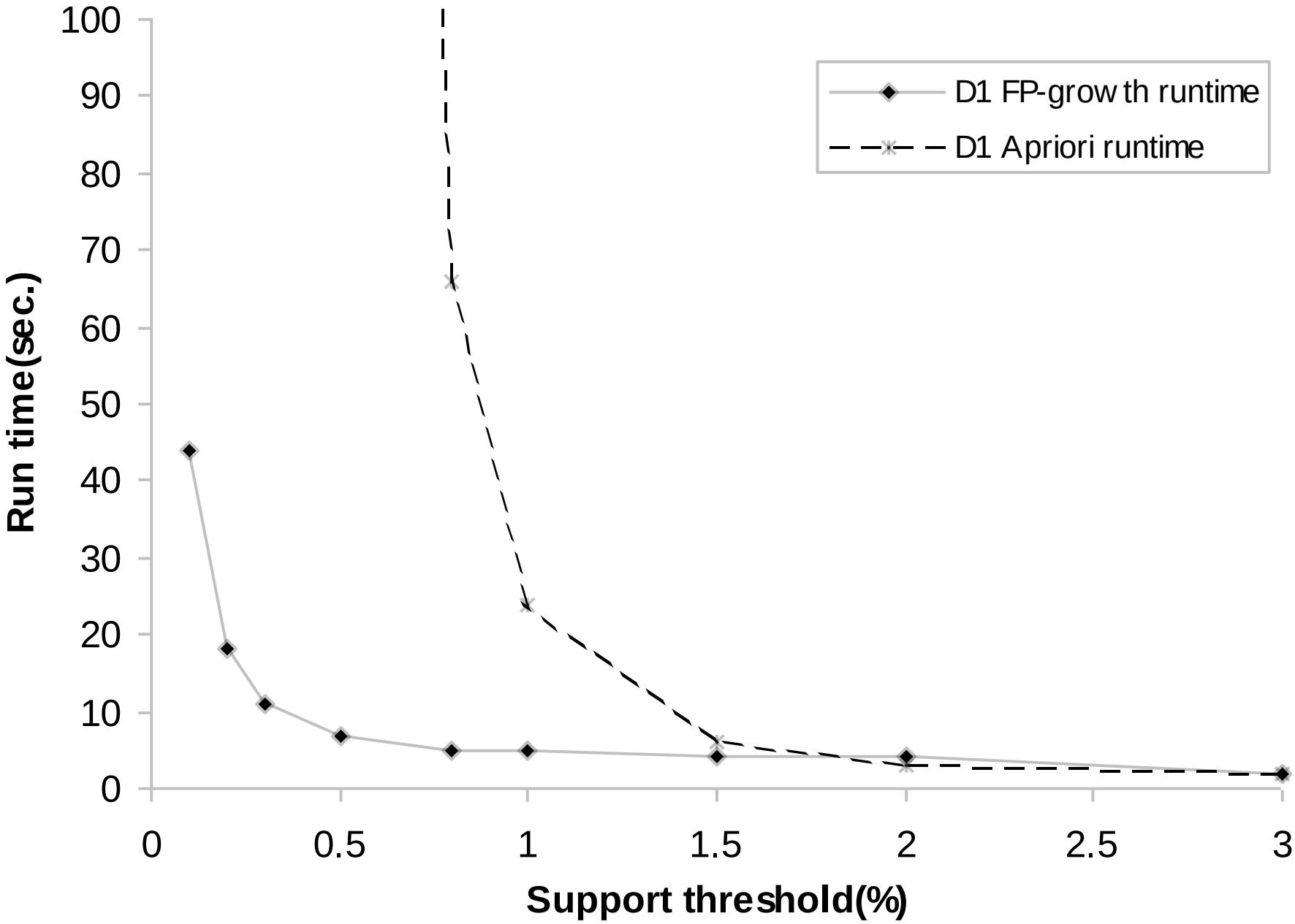
	Conditional Pattern Base
I1	{{I2:4}}



Conditional FP Tree for I1:{I2:4}

Frequent Patters Generated

	Frequent Pattern Generated
I5	{I2, I5: 2}, {I1, I5: 2}, {I2, I1, I5: 2}
I4	{I2, I4: 2}
I3	{I2, I3: 4}, {I1, I3: 4}, {I2, I1, I3: 2}
I1	{I2, I1: 4}


100
90
80
70
60
50
40
30
20
10
0
Run time(sec.)
0
0.5
1
1.5
2
2.5
3
Support threshold(%)
D1 FP-grow th runtime
D1 Apriori runtime

Thank You