

MIDTERM SAMPLE QUESTIONS

Engr. Renato R. Maaliw III, *DIT*
Professor 1, College of Engineering
Southern Luzon State University
Lucban, Quezon, Philippines

Based on the following Transactions answer the following:

Transactions

[A, B, C, D, E]

[B, C, D]

[A, C, E]

[D, E]

[C, B]

1. What is the support of C?

A. 0.6

C. 0.8

B. 0.7

D. 0.9

Transactions

[A, B, C, D, E]

[B, C, D]

[A, C, E]

[D, E]

[C, B]

1. What is the support of C?

A. 0.6

C. 0.8

B. 0.7

D. 0.9

Transactions

[A, B, C, D, E]

[B, C, D]

[A, C, E]

[D, E]

[C, B]

2. What is the support count of {C, D}

A. 1

C. 3

B. 2

D. 4

Transactions

[A, B, C, D, E]

[B, C, D]

[A, C, E]

[D, E]

[C, B]

2. What is the support count of {C, D}

A. 1

C. 3

B. 2

D. 4

Transactions

[A, B, C, D, E]

[B, C, D]

[A, C, E]

[D, E]

[C, B]

3. What is the support count of {B, E}

A. 0.1

C. 0.3

B. 0.2

D. 0.4

Transactions

[A, B, C, D, E]

[B, C, D]

[A, C, E]

[D, E]

[C, B]

3. What is the support count of {B, E}

A. 0.1

C. 0.3

B. 0.2

D. 0.4

Transactions

[A, B, C, D, E]

[B, C, D]

[A, C, E]

[D, E]

[C, B]

4. What is the lift of $\{B\} \rightarrow \{D\}$

A. 1

C. 1.11

B. 1.22

D. 1.50

Transactions

[A, B, C, D, E]

[B, C, D]

[A, C, E]

[D, E]

[C, B]

4. What is the lift of $\{B\} \rightarrow \{D\}$	
A. 1	C. 1.11
B. 1.22	D. 1.50

Transactions

[A, B, C, D, E]

[B, C, D]

[A, C, E]

[D, E]

[C, B]

5. What is the lift of $\{B, C\} \rightarrow \{D\}$

A. 1.11

C. 1.22

B. 2.22

D. 11.1

Transactions

[A, B, C, D, E]

[B, C, D]

[A, C, E]

[D, E]

[C, B]

5. What is the lift of $\{B, C\} \rightarrow \{D\}$

A. 1.11

C. 1.22

B. 2.22

D. 11.1

6. What is the key principle of the Apriori algorithm?

A. Use of neural network to find patterns in data

C. Randomly generating item sets and evaluating their support

B. The “downward closure” property, where subsets of a frequent itemset are also frequent

D. Using decision trees to classify transactions.

6. What is the key principle of the Apriori algorithm?

A. Use of neural network to find patterns in data

C. Randomly generating item sets and evaluating their support

B. The “downward closure” property, where subsets of a frequent itemset are also frequent

D. Using decision trees to classify transactions.

7. What is the support in the context of Apriori?

A. The percentage of transactions that contain at least one item

C. The fraction of transactions that contain a specific item or itemset

B. The likelihood of an item appearing in a rule

D. The confidence level of an association rule

7. What is the support in the context of Apriori?

A. The percentage of transactions that contain at least one item

C. The fraction of transactions that contain a specific item or itemset

B. The likelihood of an item appearing in a rule

D. The confidence level of an association rule

8. Which parameter is used to reduce the search space in Apriori?

A. Lift

C. Confidence threshold

B. Support threshold

D. Correlation coefficient

8. Which parameter is used to reduce the search space in Apriori?

A. Lift

C. Confidence threshold

B. Support threshold

D. Correlation coefficient

9. What does “lift” measure in association rule mining

A. The probability of an itemset occurring

C. The ratio of support to confidence

B. The strength of an association compared to random chance

D. The frequency of the antecedent

9. What does “lift” measure in association rule mining

A. The probability of an itemset occurring

C. The ratio of support to confidence

B. The strength of an association compared to random chance

D. The frequency of the antecedent

10. Why is Apriori more efficient than a brute-force approach?

A. It processes the entire dataset in one pass

C. It only evaluates itemsets that have a high confidence

B. It evaluates the lift values for all rules in advance

D. It uses the downward closure property to prune the search space.

10. Why is Apriori more efficient than a brute-force approach?

A. It processes the entire dataset in one pass

C. It only evaluates itemsets that have a high confidence

B. It evaluates the lift values for all rules in advance

D. It uses the downward closure property to prune the search space.

**Given the following transactions
construct the Final FP-Tree**

Transactions

{A, B, C, D}

{A, C, D}

{B, D}

{A, D}

{B, C}

Transactions (Sorted)

{D, A, B, C}

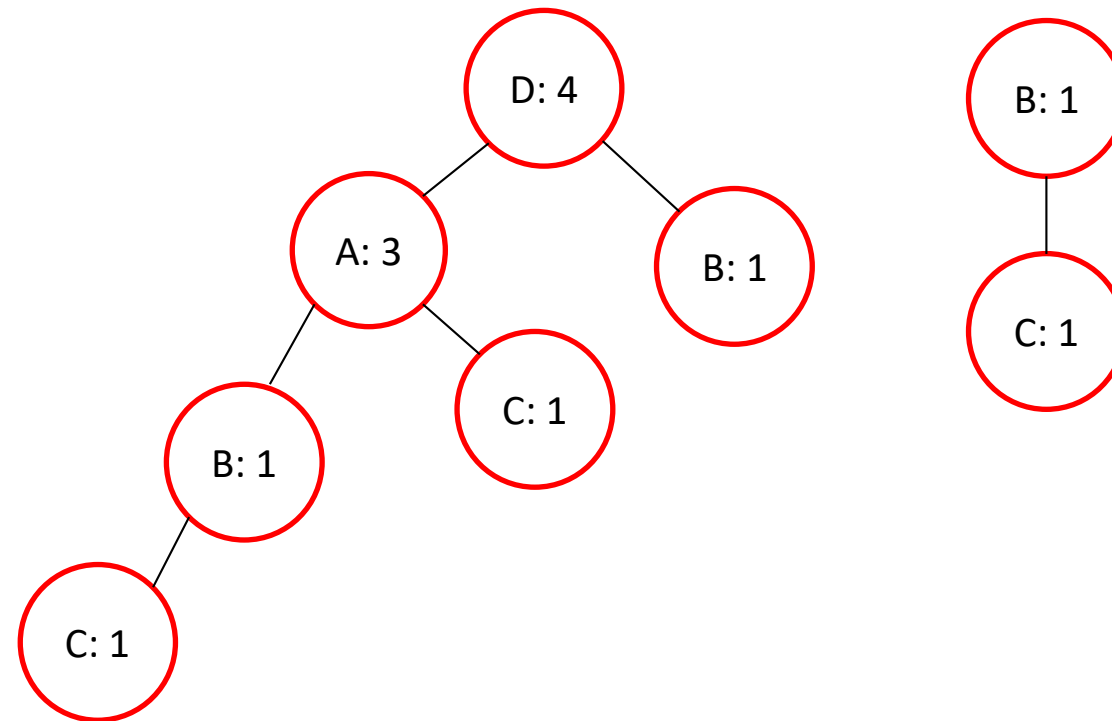
{D, A, C}

{D, B}

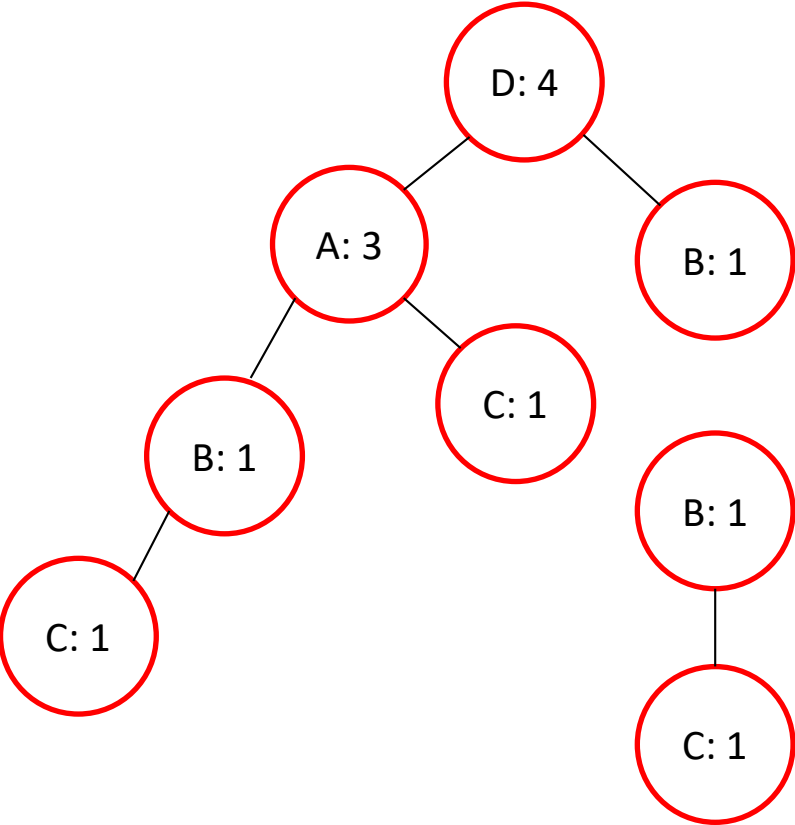
{D, A}

{B, C}

Final FP-Tree



Final FP-Tree



11. What is the correct frequent pattern for C?

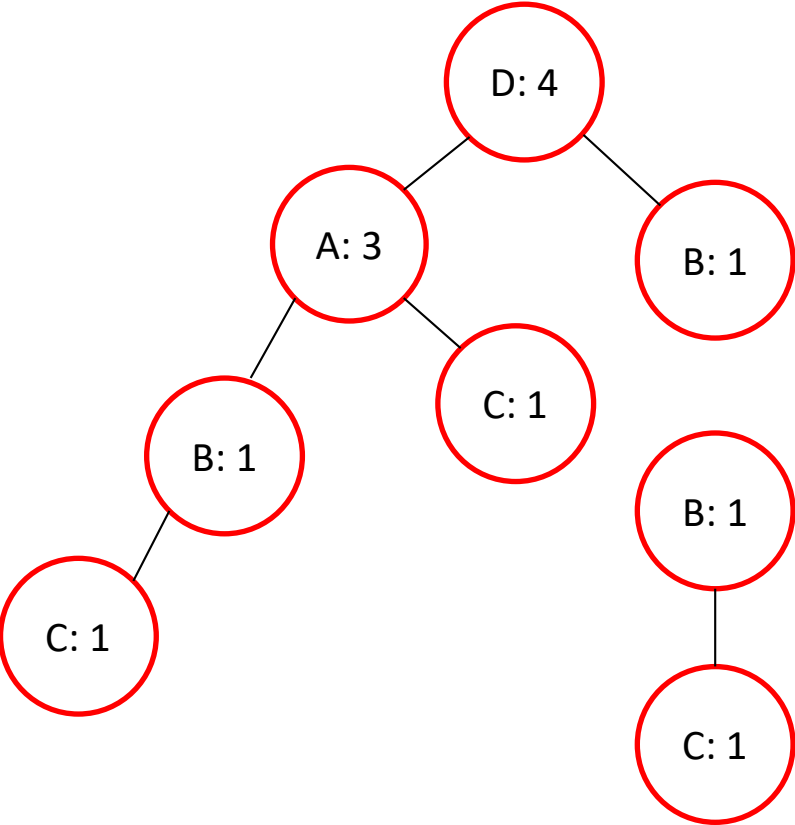
A. {D, A, B} (S = 2)

C. {D, A, B} (S = 1}

B. {B} (S = 1}

D. {A, B} (S = 2}

Final FP-Tree



11. What is the correct frequent pattern for C?

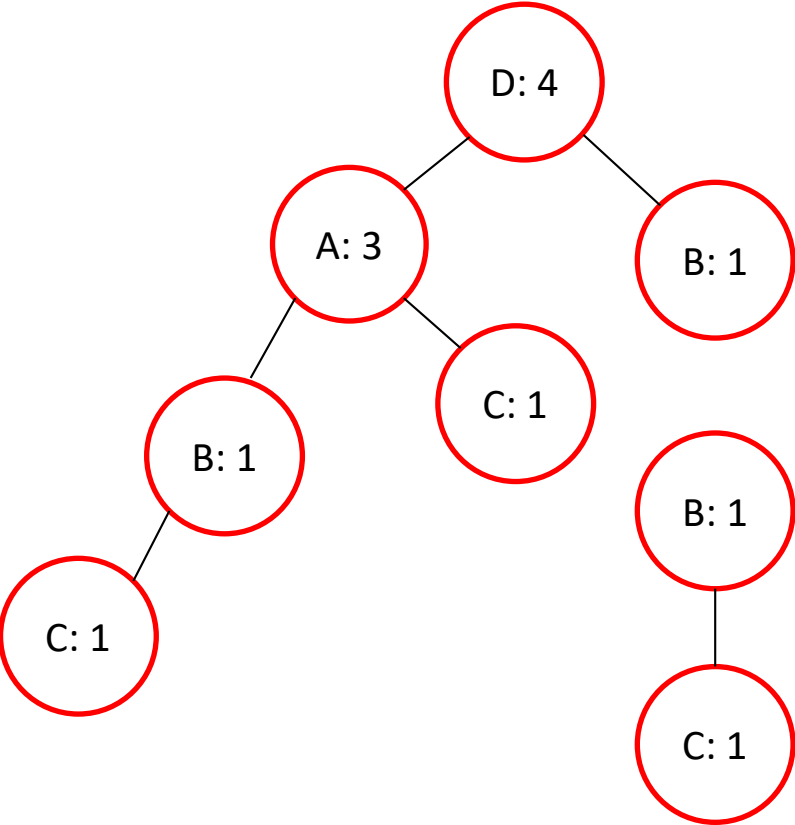
A. {D, A, B} ($S = 2$)

C. {D, A, B} ($S = 1$)

B. {B} ($S = 1$)

D. {A, B} ($S = 2$)

Final FP-Tree



12. What is the correct frequent pattern for B?

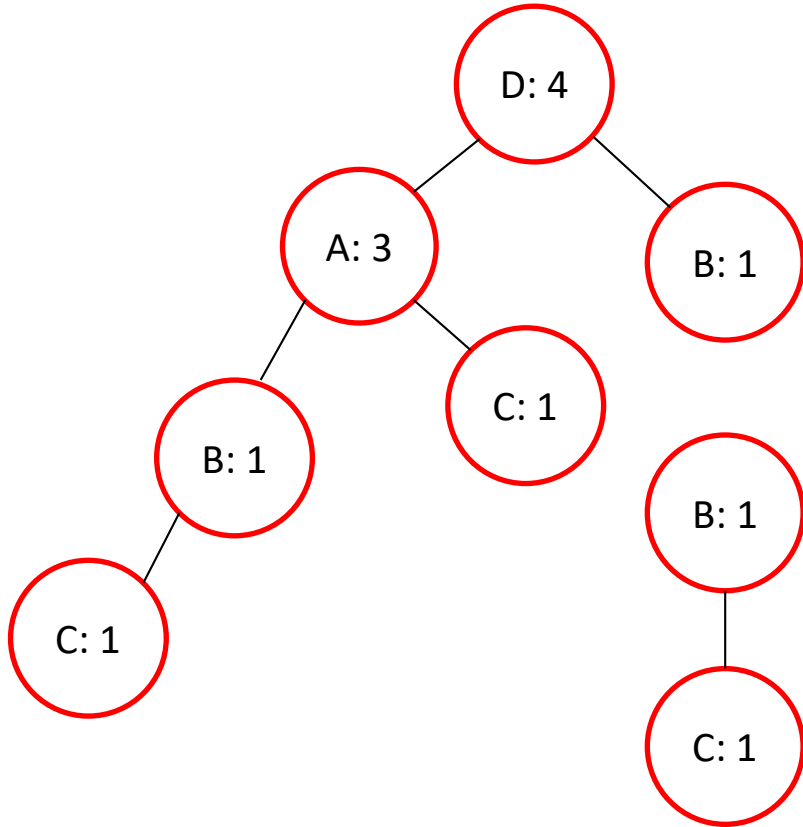
A. {B} (S = 2)

C. {D} (S = 1)

B. {D, A} (S = 3)

D. {D, A} (S = 4)

Final FP-Tree



12. What is the correct frequent pattern for B?

A. {B} (S = 2)

C. {D} (S = 1)

B. {D, A} (S = 3)

D. {D, A} (S = 4)

13. What does FP in FP-Growth stand for?

A. Frequent Pattern

C. Frequent Prediction

B. Fast Processing

D. Forward Propagation

13. What does FP in FP-Growth stand for?

A. Frequent Pattern

C. Frequent Prediction

B. Fast Processing

D. Forward Propagation

14. How does FP-Growth differ from Apriori?

A. FP-Growth generates candidate itemsets, while Apriori does not

C. FP-Growth requires more iterations than Apriori.

B. FP-Growth uses a tree structure to store data, whereas Apriori uses a level-wise approach.

D. FP-Growth only works for single-item transactions.

14. How does FP-Growth differ from Apriori?

A. FP-Growth generates candidate itemsets, while Apriori does not

C. FP-Growth requires more iterations than Apriori.

B. FP-Growth uses a tree structure to store data, whereas Apriori uses a level-wise approach.

D. FP-Growth only works for single-item transactions.

15. What is the main advantage of FP-Growth over Apriori?

A. FP-Growth finds all possible association rules.

C. FP-Growth works on smaller datasets only.

B. FP-Growth avoids the expensive candidate generation process.

D. FP-Growth uses lift instead of support.

15. What is the main advantage of FP-Growth over Apriori?

A. FP-Growth finds all possible association rules.

C. FP-Growth works on smaller datasets only.

B. FP-Growth avoids the expensive candidate generation process.

D. FP-Growth uses lift instead of support.

16. In FP-Growth, what happens after constructing the FP-Tree?

A. Candidate itemsets are generated for each branch of the tree.

C. The tree is used to identify frequent itemsets by recursively mining conditional patterns.

B. The tree is pruned to retain only the top-k patterns.

D. Rules are directly extracted from the tree without further computation.

16. In FP-Growth, what happens after constructing the FP-Tree?

A. Candidate itemsets are generated for each branch of the tree.

C. The tree is used to identify frequent itemsets by recursively mining conditional patterns.

B. The tree is pruned to retain only the top-k patterns.

D. Rules are directly extracted from the tree without further computation.

17. What type of dataset is FP-Growth best suited for?

A. Datasets with very large transactions.

C. Datasets where frequent patterns are dense and repetitive.

B. Datasets with a high number of unique items.

D. Datasets with low transaction counts.

17. What type of dataset is FP-Growth best suited for?

A. Datasets with very large transactions.

C. Datasets where frequent patterns are dense and repetitive.

B. Datasets with a high number of unique items.

D. Datasets with low transaction counts.

18. What is a drawback of FP-Growth?

A. It cannot handle datasets with more than 1,000 transactions.

C. It cannot find association rules with confidence values.

B. The FP-Tree may require a large amount of memory for sparse datasets.

D. It generates candidates for every transaction.

18. What is a drawback of FP-Growth?

A. It cannot handle datasets with more than 1,000 transactions.

C. It cannot find association rules with confidence values.

B. The FP-Tree may require a large amount of memory for sparse datasets.

D. It generates candidates for every transaction.

19. What is the main input parameter for FP-Growth?

A. Confidence threshold

C. Support threshold

B. Lift threshold

D. Depth of the FP-Tree

19. What is the main input parameter for FP-Growth?

A. Confidence threshold

C. Support threshold

B. Lift threshold

D. Depth of the FP-Tree

20. What does FP-Growth do with infrequent items during the tree construction phase?

A. Removes them before constructing the tree.

C. Keeps them in the tree but marks them as low priority

B. Groups them into a single node in the tree

D. Places them at the root of the tree

20. What does FP-Growth do with infrequent items during the tree construction phase?

A. Removes them before constructing the tree.

C. Keeps them in the tree but marks them as low priority

B. Groups them into a single node in the tree

D. Places them at the root of the tree

Good Luck!