

CIS 572 Assignment 1

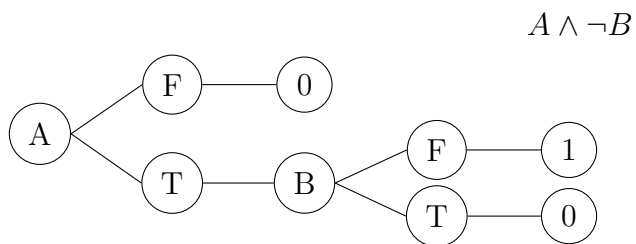
Steven Walton

April 6, 2019

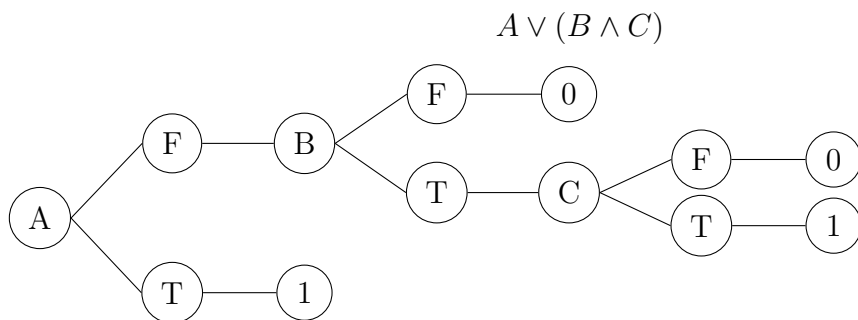
Problem 1

Answer Exercise 3.1 from Chapter 3 of Mitchell's machine learning book.

Part a)

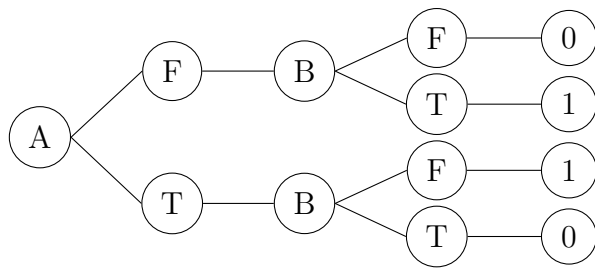


Part b)

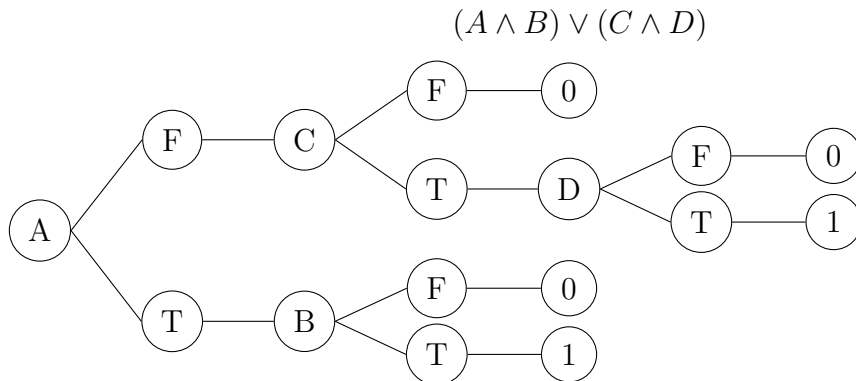


Part c)

$A \text{ XOR } B$



Part d)



Problem 2

Consider the samples in the Play-tennis dataset from Table 3.2 in Mitchell's textbook. If you calculate the information-gain for all of the attributes of this set, you will observe that the attribute "Outlook" has the largest information-gain, which is equal to 0.246. Therefore, the attribute "Outlook" is the best heuristic choice for the root node.

- List the labels of the new tree branches below the root node.
- Which partition of the data will be assigned to each branch by ID3? Please list the sample IDs that will be assigned to each branch.
- Calculate the information gain for the remaining attributes in each branch, and determine which attribute will be chosen as the root of the sub-tree in each branch.

Part a)

Part b)

Part c)