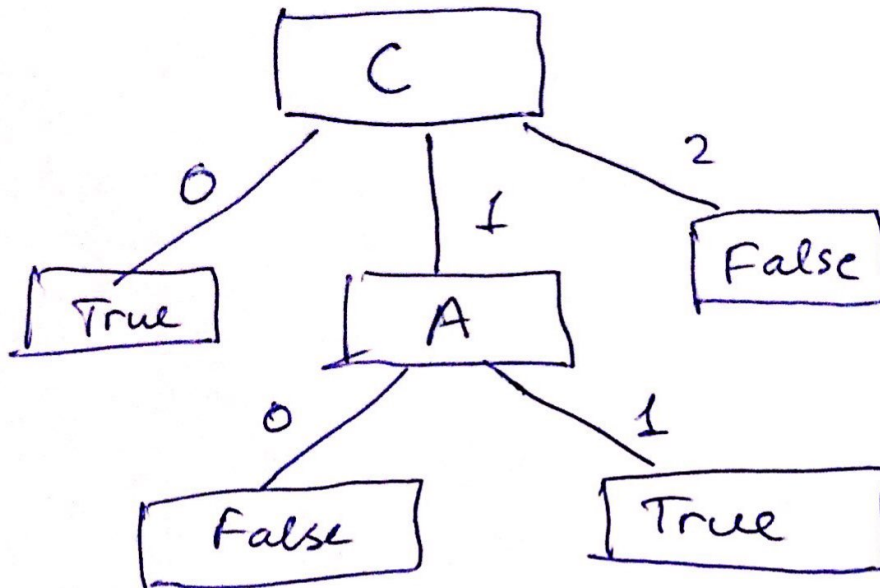a) C because the data distribution is equal in A and B, which means that information gain would be zero and also because C is a pure node for cases such as $C=0$ and $C=2$ in.

b)

c) • False
• True