SEM 451, CS 470, Exercise (learning by examples II)

A candy shop interviews its customers to know what they **like** about their candy based on two features (or attributes), namely, **color** and **flavor** and the table below shows the responses:

Examples:

Color	Flavor	Like?
Red	Vanilla	Yes
Red	Strawberry	Yes
Green	Vanilla	Yes
Green	Strawberry	No
Blue	Vanilla	No
Blue	Strawberry	No

Some Common Entropy values

Entropy	Value
$H(1,0) = H(0,1)$	0
$H\left(\frac{1}{2},\frac{1}{2}\right)$	1
$H\left(\frac{2}{3},\frac{1}{3}\right) = H\left(\frac{1}{3},\frac{2}{3}\right)$	0.9183

Use the examples in the table above to answer each of the following questions.

Note: You may find the above common entropy values useful.

a) What is the **entropy** of like, H(like)?

b) By computing the information gains, find which of the two attributes should be used to split the examples first to form the root of the tree.

c) Draw the decision tree based on your results in (b) above.