22 The solution of this problem was not satisfactory for some of the students, so we are posting it here again. You are advised to go through the book and the wiki page for conditional mutual information, in order to get familiar with the equations used here.

$$I(X;E|Y) = H(X|Y) - H(X|E,Y) \text{ and} \tag{1}$$
$$I(X;E|Y) = H(E|Y) - H(E|X,Y) \tag{2}$$
$$\text{if } X \text{ and } Y \text{ are known, } H(E|X,Y) = 0 \tag{3}$$
$$\text{hence, } H(X|Y) = H(E|Y) + H(X|E,Y) \tag{4}$$
$$\text{now, conditionality reduces entropy, } H(E|Y) \leq H(E) \tag{5}$$

For the second term in equation (4),

$$H(X|E,Y) = Pr(E=0)H(X|E=0,Y) + Pr(E=1)H(X|E=1,Y) \tag{6}$$

The first term in the above equation is 0, because given error is 0, X and Y are same. So the entropy is 0. In the second term, $Pr(E=1) = P_e$, and $H(X|E=1,Y) \leq \log(n-1)$. The reason of taking $n-1$ and not n is because X does not take the value Y does, so the entropy is bounded by $\log(n-1)$ (one symbol less).

Plugging all into equation (4) gives

$$H(X|Y) = H(P_e) + P_e \log(n-1) \tag{7}$$