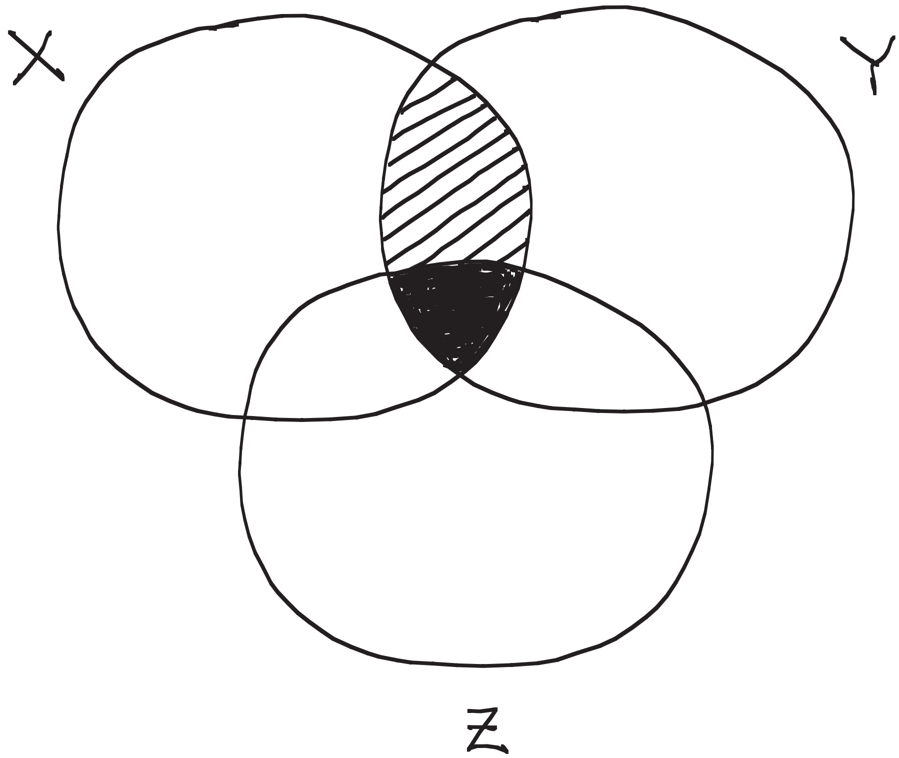
1.5



(a) The solid area is the proposed mutual information $H(X:Y:Z)$. The striped area is the mutual information of X and Y, given Z, i.e., $H(X:Y|Z) = H(X|Z) - H(X|Y,Z)$. The sum of these areas is the mutual information of X and Y, i.e., $H(X:Y) = H(X) - H(X|Y)$. So

$$H(X:Y:Z) = H(X:Y) - H(X:Y|Z) = H(X) - H(X|Y) - H(X|Z) + H(X|Y,Z)\ . \quad (1)$$

Expanded out entirely to get rid of all the conditional informations, this becomes

$$\begin{aligned} H(X:Y:Z) &= H(X) - H(X|Y) - H(X|Z) + H(X|Y,Z) \\ &= H(X) - [H(X,Y) - H(Y)] - [H(X,Z) - H(Z)] + [H(X,Y,Z) - H(Y,Z)] \\ &= H(X,Y,Z) - H(X,Y) - H(X,Z) - H(Y,Z) + H(X) + H(Y) + H(Z)\ . \end{aligned} \quad (2)$$

The final form is manifestly symmetric in the three systems, as is already clear from the Venn diagram.

(b) Consider three discrete "random variables" X, Y, and Z. According to the diagram of intersecting circles, the information shared in common by all three quantities—the three-variable mutual information—ought to be

$$H(X:Y:Z) \equiv H(X:Y) - H(X:Y|Z) = H(X) - H(X|Y) - H(X|Z) + H(X|Y,Z)\ . \quad (3)$$

This three-variable mutual information is manifestly symmetric under interchange of any two variables and thus can also be written as (in accordance with the diagram)

$$H(X:Y:Z) = H(X:Z) - H(X:Z|Y) = H(Y:Z) - H(Y:Z|X)\ . \tag{4}$$

The problem with $H(X:Y:Z)$ is that it can be negative when the amount of information X and Y share (i.e., the tightness of their correlation) is greater when Z is known than when it is not. Let's see how this works out for the probabilities given in the problem.

Suppose X, Y, and Z are binary variables, each taking on values 0 and 1. Let the three-variable joint probabilities be given by

$$p_{000} = p_{111} = 0\ , \quad p_{001} = p_{010} = p_{100} = \tfrac{1}{3}p\ , \quad p_{110} = p_{101} = p_{011} = \tfrac{1}{3}q\ , \tag{5}$$

where $p+q=1$. This joint distribution is symmetric under interchange of any two variables and symmetric under the simultaneous interchange of 0 and 1 and p and q. The marginal probabilities are given by

$$p_{00} = \tfrac{1}{3}p\ , \quad p_{11} = \tfrac{1}{3}q\ , \quad p_{01} = p_{10} = \tfrac{1}{3}\ , \tag{6}$$

$$p_0 = \tfrac{2}{3}p + \tfrac{1}{3}q = \tfrac{1}{3}(1+p)\ , \quad p_1 = \tfrac{1}{3}p + \tfrac{2}{3}q = \tfrac{1}{3}(2-p)\ , \tag{7}$$

and the conditional probabilities by

$$\begin{aligned} &p_{00|0} = 0 && p_{11|1} = 0 \\ &p_{01|0} = p_{10|0} = \frac{p}{2p+q} = \frac{p}{1+p}\ , && p_{10|1} = p_{01|1} = \frac{q}{2q+p} = \frac{1-p}{2-p}\ , \\ &p_{11|0} = \frac{q}{2p+q} = \frac{1-p}{1+p} && p_{00|1} = \frac{p}{2q+p} = \frac{p}{2-p} \end{aligned} \tag{8}$$

$$\begin{matrix} p_{0|00} = 0 \\ p_{1|00} = 1 \end{matrix}\ , \quad \begin{matrix} p_{0|01} = p \\ p_{1|01} = q \end{matrix}\ , \quad \begin{matrix} p_{0|10} = p \\ p_{1|10} = q \end{matrix}\ , \quad \begin{matrix} p_{0|11} = 1 \\ p_{1|11} = 0 \end{matrix}\ , \tag{9}$$

$$\begin{aligned} p_{0|0} &= \frac{p}{2p+q} = \frac{p}{1+p} \\ p_{1|0} &= \frac{1}{2p+q} = \frac{1}{1+p} \end{aligned}\ , \qquad \begin{aligned} p_{0|1} &= \frac{1}{p+2q} = \frac{1}{2-p} \\ p_{1|1} &= \frac{q}{p+2q} = \frac{1-p}{2-p} \end{aligned}\ . \tag{10}$$

The symmetry under exchange of any two variables means that these probabilities apply in the obvious way to any choice of the random variables.

The correlations between X and Y when Z is known are given by the probabilities in Eq. (8), and the correlations between X and Y when Z is averaged over are given by the probabilities in Eq. (6). When $p=1$, X and Y are perfectly correlated for $z=1$, i.e., $p_{00|1}=1$, and they are half that strongly correlated when $z=0$, i.e., $p_{01|0}=p_{10|0}=\frac{1}{2}$. When we average over Z to get the marginal probabilities (6), these correlations get spread out over the three possibilities, 00, 01, and 10, thus giving a weaker correlation. The tighter correlation when Z is known than when it is not gives rise to the negative value for the proposed mutual information of the three variables.

The information quantities that go into $H(X:Y:Z)$ become

$$H(X) = H(p_0, p_1) = H\big(\tfrac{1}{3}(1+p), \tfrac{1}{3}(2-p)\big) \ , \tag{11}$$

$$\begin{aligned} H(X|Y) = H(X|Z) &= p_0 H(p_{0|0}, p_{1|0}) + p_1 H(p_{0|1}, p_{1|1}) \\ &= \tfrac{1}{3}(1+p) H\left(\frac{p}{1+p}, \frac{1}{1+p}\right) + \tfrac{1}{3}(2-p) H\left(\frac{1}{2-p}, \frac{1-p}{2-p}\right) \ , \end{aligned} \tag{12}$$

$$\begin{aligned} H(X|Y,Z) &= p_{00} H(p_{0|00}, p_{1|00}) + p_{01} H(p_{0|01}, p_{1|01}) \\ &\quad + p_{10} H(p_{0|10}, p_{1|10}) + p_{11} H(p_{0|11}, p_{1|11}) \\ &= \tfrac{2}{3} H(p, 1-p) \ , \end{aligned} \tag{13}$$

$$\begin{aligned} H(X:Y) &= H(X) - H(X|Y) \\ &= H\big(\tfrac{1}{3}(1+p), \tfrac{1}{3}(2-p)\big) \\ &\quad - \tfrac{1}{3}(1+p) H\left(\frac{p}{1+p}, \frac{1}{1+p}\right) - \tfrac{1}{3}(2-p) H\left(\frac{1}{2-p}, \frac{1-p}{2-p}\right) \ , \end{aligned} \tag{14}$$

$$\begin{aligned} H(X:Y|Z) &= H(X|Z) - H(X|Y,Z) \\ &= \tfrac{1}{3}(1+p) H\left(\frac{p}{1+p}, \frac{1}{1+p}\right) + \tfrac{1}{3}(2-p) H\left(\frac{1}{2-p}, \frac{1-p}{2-p}\right) - \tfrac{2}{3} H(p, 1-p) \ . \end{aligned} \tag{15}$$

For the special case $p = 1$ and $q = 0$, these expressions reduce to

$$\begin{aligned} H(X) &= H\big(\tfrac{2}{3}, \tfrac{1}{3}\big) = \log 3 - \tfrac{2}{3} = 0.9183 \ , \\ H(X|Y) = H(X|Z) &= \tfrac{2}{3} H(\tfrac{1}{2}, \tfrac{1}{2}) = \tfrac{2}{3} \ , \\ H(X|Y,Z) &= 0 \ , \end{aligned} \tag{16}$$

$$H(X:Y) = H\big(\tfrac{2}{3}, \tfrac{1}{3}\big) - \tfrac{2}{3} = \log 3 - \tfrac{4}{3} = 0.2516 \ , \quad H(X:Y|Z) = \tfrac{2}{3} \ , \tag{17}$$

implying that

$$H(X:Y:Z) = H\big(\tfrac{2}{3}, \tfrac{1}{3}\big) - \tfrac{4}{3} = \log 3 - 2 = -0.4150 \ . \tag{18}$$