ECE 788: Network Information Theory
Assignment 3 (due on Sept. 28)

1. This problem guides you through the proof of the achievability part of the lossy source coding theorem. All the elements were given in class, along with an intuitive explanation. Here we formalize the proof using the fundamental lemma (see class notes).
We need to prove that there exists a coding/ decoding scheme $\mathcal{C}$ that is able to guarantee $E[d^n(X^n, \hat{X}^n)|\mathcal{C}] \leq D$, where the expectation is taken with respect to the distribution of the i.i.d. source X^n and $\hat{X}^n$ is the reconstruction point of X^n for the scheme $\mathcal{C}$ (note that $\hat{X}^n$ is a function of X^n). Moreover, we require such scheme to have rate R arbitrarily close to $I(X;\hat{X})$ for a joint distribution $P_{X\hat{X}}$ such that $E[d(X,\hat{X})] \leq D$ (here the expectation is taken with respect to $P_{X\hat{X}}$). To prove this, we use Shannon's random coding argument and generate the codebook of 2^{nR} reconstruction points $\hat{X}^n(w)$, $w \in [1, 2^{nR}]$, i.i.d. and independently, from the marginal distribution $P_{\hat{X}}$. The encoder, given a sequence x^n, looks for a jointly typical $\hat{x}^n(w)$ and "sends" the corresponding w. If not found, it sends $w = 1$. The decoder reconstructs $\hat{x}^n(w)$ (see also booklet, p. 20-21 for description of the scheme).
We fix a $P_{X\hat{X}}$ such that $E[d(X,\hat{X})] \leq D$. We first prove in steps 1.1-1.10 that

$$E_{\mathcal{C}}\left[E[d^n(X^n, \hat{X}^n)|\mathcal{C}]\right] \leq D, \tag{1}$$

where $E_{\mathcal{C}}$ is the expectation with respect to the codebooks generated as above. Define the above double expectation, with respect to the source and codebook distributions, as $E[d^n(X^n, \hat{X}^n)]$ for simplicity.
1.1. Argue that we can break the calculation of $E[d^n(X^n, \hat{X}^n)]$ as

$$E[d^n(X^n, \hat{X}^n)] = \sum_{i=1}^{3} \Pr[\mathcal{E}_i] E[d^n(X^n, \hat{X}^n)|\mathcal{E}_i],$$

where $\mathcal{E}_i$ are given events in terms of random vectors $(X^n, \hat{X}^n)$. What is the condition on events $\mathcal{E}_i$ such that the above formula holds (this refers to any expectation, not specifically to this problem)?
1.2. Define the events (2.3)-(2.5) in the booklet. Check that they satisfy the condition found above.
1.3. Show that $\Pr[\mathcal{E}_1] \leq \delta_\epsilon(n)$ using one of the theorems we discussed.
1.4. Show that $E[d^n(X^n, \hat{X}^n)|\mathcal{E}_3] \leq E[d(X,\hat{X})] + \epsilon d_{\max}$ (Hint: This was done in class! Find where)
1.5. Show that from the points 1.3. and 1.4. above we have that

$$\Pr[\mathcal{E}_1]E[d^n(X^n, \hat{X}^n)|\mathcal{E}_1] + \Pr[\mathcal{E}_3]E[d^n(X^n, \hat{X}^n)|\mathcal{E}_3] \leq (\delta_\epsilon(n) + \epsilon)d_{\max} + E[d(X,\hat{X})]$$

1.6. We are left with calculating $\Pr[\mathcal{E}_2]$. We want to prove that $\Pr[\mathcal{E}_2]$ can be made arbitrarily small, i.e., $\Pr[\mathcal{E}_2] \leq \tilde{\delta}_\epsilon(n)$ for some $\tilde{\delta}_\epsilon(n) \to 0$ as $n \to \infty$ (for $\epsilon > 0$). In fact, from 1.5 above, this will allow us to conclude that $E[d^n(X^n, \hat{X}^n)]$ can be made arbitrarily close to D. Why?
1.7. To prove that $\Pr[\mathcal{E}_2] \leq \tilde{\delta}_\epsilon(n)$, we proceed as follows. First, argue that

$$\Pr[\mathcal{E}_2] = \Pr\left[\bigcap_{w=1}^{2^{nR}} \{(x^n, \hat{X}^n(w)) \notin T_\epsilon^n(P_{X\hat{X}})\}\right],$$

where $x^n \in T_\epsilon^n(P_X)$.
1.8. Now, show that

$$\Pr[\mathcal{E}_2] = \prod_{w=1}^{2^{nR}} \Pr\left[(x^n, \hat{X}^n(w)) \notin T_\epsilon^n(P_{X\hat{X}})\right]$$

(Hint: It depends on how we built the codebook)
1.9. Argue that to calculate an upper bound to $\Pr\left[(x^n, \hat{X}^n(w)) \notin T_\epsilon^n(P_{X\hat{X}})\right]$ we can use the fundamental lemma (why *not* the conditional AEP?). Calculate such upper bound.
1.10. Justify the remaining steps in (2.8) of the booklet. From this, show that $\Pr[\mathcal{E}_2] \leq \tilde{\delta}_\epsilon(n)$, as desired.
1.11. Conclude the proof by explaining why having verified (1) guarantees that there exists at least one codebook with $E[d(X^n, \hat{X}^n)|\mathcal{C}] \leq D$ with the desired properties