EE376A Information Theory Lecture 5 - 01/20/2015

Lecture 5: Near-Lossless Compression

Lecturer: Tsachy Weissman *Scribe: Chuan-Zheng Lee, Jocelin Ho, Dawei Ding*

1 Review of Typical Sequences and AEP

Let $U_1, U_2, \ldots, U_n$ be i.i.d. with discrete distribution U. Then

$$p\left(u^n\right) = \prod_{i=1}^{n} p_U\left(u_i\right). \tag{1}$$

We defined u^n to be ε**-typical** if

$$2^{-n[H(U)+\varepsilon]} \leq p(u^n) \leq 2^{-n[H(U)-\varepsilon]} \tag{2}$$

We defined $A_\varepsilon^{(n)}$ to be the set of all u^n's that are ε-typical.

Theorem 1 (AEP)**.** *For all $\varepsilon > 0$,*

$$\lim_{n\to\infty} P\left(U^n \in A_\varepsilon^{(n)}\right) = 1 \tag{3}$$

Theorem 2. *For all $\varepsilon > 0$ and sufficiently large n,*

$$\left(1-\varepsilon\right) 2^{n[H(U)-\varepsilon]} \leq \left|A_\varepsilon^{(n)}\right| \leq 2^{n[H(U)+\varepsilon]} \tag{4}$$

Theorem 3. *For all $\delta > 0$ and all sequences of sets $B^{(n)} \subseteq \mathcal{U}^n$ such that $\left|B^{(n)}\right| \leq 2^{n[H(U)-\delta]}$,*

$$\lim_{n\to\infty} P\left(U^n \in B^{(n)}\right) = 0 \tag{5}$$

A visualization of Theorem 3 is shown in Figure 1.

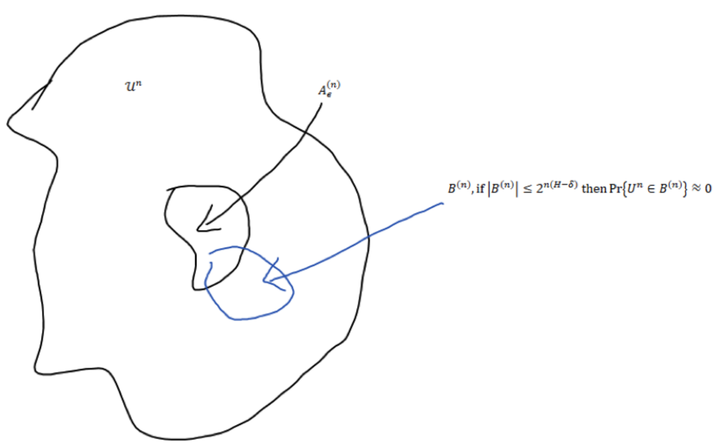

Figure 1: Visualization of all source sequences and ε-typical sequences.

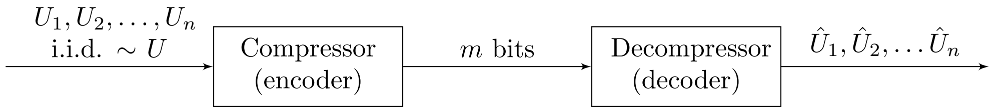

Figure 2: Block diagram of lossless compression

2 Fixed Length (Near) Lossless Compression

Suppose we have a source $U_1, \ldots, U_n$ i.i.d. with distribution U. We wish to devise a compression scheme, as shown in Figure 2. The compressor takes a block of n source symbols and converts them into m binary bits. The decompressor does the inverse process. The rate of such a scheme (compression and decompression) is defined to be $\frac{m}{n}$ bits/source symbol.

We relax our requirements slightly: rather than insisting on strictly lossless compression, we will simply require the probability of error to be small. That is,

$$P_e = P\left(\hat{U}^n \neq U^n\right) \ll 1 \tag{6}$$

Definition 4 (Achievable rate)**.** *R is an achievable rate if for all $\varepsilon > 0$, there exists a scheme (n, m, compressor, decompressor) whose rate $\frac{m}{n} \leq R$ and whose probability of error $P_e < \varepsilon$.*

We are interested in the question: What is the lowest achievable rate? Theorems 5 and 6 tell us the answer.

Theorem 5 (Direct theorem)**.** *For all $R > H(U)$, R is achievable.*

Proof Fix $R > H(U)$ and $\varepsilon > 0$. Set $\delta = R - H(U) > 0$ and note that for all n sufficiently large, by Theorem 1,

$$P\left(U^n \notin A_\delta^{(n)}\right) < \varepsilon, \tag{7}$$

and by Theorem 2,

$$\left|A_\delta^{(n)}\right| \leq 2^{n[H(U)+\delta]} = 2^{nR}. \tag{8}$$

Consider a scheme that enumerates sequences in $A_\delta^{(n)}$. That is, the compressor outputs a binary representation of the index of U^n if $U^n \in A_\delta^{(n)}$; otherwise, it outputs $(0, 0, \ldots, 0)$. The decompressor maps this binary representation back to the corresponding sequence in $A_\delta^{(n)}$. For this scheme, the probability of error is bounded by

$$P_e \leq P\left(U^n \notin A_\delta^{(n)}\right) < \varepsilon \tag{9}$$

and the rate is equal to

$$\frac{\log\left|A_\delta^{(n)}\right|}{n} \leq \frac{nR}{n} = R \tag{10}$$

Hence, R is an achievable rate. $\square$

Theorem 6 (Converse theorem)**.** *If $R < H(U)$, R is not achievable.*

Proof For a given scheme of rate $r \leq R$ (and block length n), let $B^{(n)}$ denote the set of possible reconstruction sequences $\hat{U}_n$. Note that $\left|B^{(n)}\right| \leq 2^m = 2^{nr} \leq 2^{nR}$. So if $R < H(U)$, by Theorem 3,

$$P_e \geq P\left(U^n \notin B^{(n)}\right) \to 1, \text{ as } n \to \infty \tag{11}$$

Hence, increasing n cannot make the probability of error arbitrarily small. Furthermore, there is clearly a nonzero probability of error for any finite n, so R is not achievable. Conceptually, if the rate is too small, it can't represent a large enough set. □