Eighth Lecture

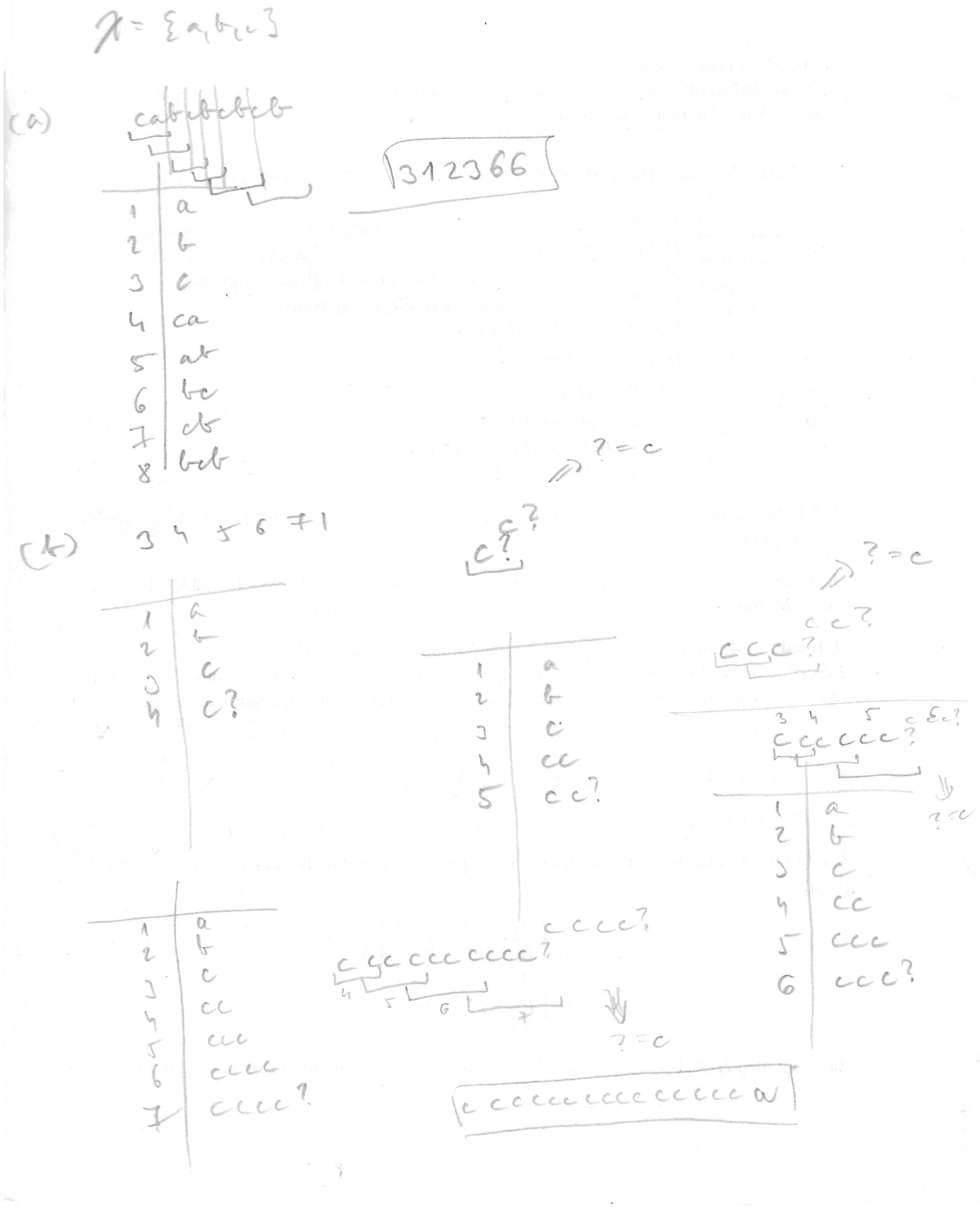
November 6, 2023

Solution of the Home work

Let the source alphabet be $\mathcal{X} = \{a, b, c\}$ and the initial dictionary contain the letters a, b and c with their indexes (1, 2 and 3 respectively). Using the Lempel-Ziv-Welch algorithm

(a) encode the sequence *cabcbcbcb*

(b) decode the sequence $3, 4, 5, 6, 7, 1$



Quantization

In many practical situations the source variables are real numbers, thus have a continuum range. If we want to use digital communication we have to discretize, which means that some kind of "rounding" is necessary.

Def. Let $X = X_1, X_2, \dots$ be a stationary source, where the X_i's are real-valued random variables. A (1-dimensional) *quantized* version of this source is a sequence of discrete random variables (another source) $Q(X_1), Q(X_2), \dots$ obtained by a map $Q : R \to R$ where the range of the map is finite. The function $Q(.)$ is called the *quantizer.*

Goal: Quantize a source so that the caused distortion is small.

How can we measure the distortion? We will do it by using the quadratic distortion measure $D(Q)$ defined for n-length blocks as

$$D(Q) = \frac{1}{n} E\left(\sum_{i=1}^{n} (X_i - Q(X_i))^2\right),$$

where $E(.)$ means expected value.
Since our X_i's are identically distributed we have

$$D(Q) = E((X - Q(X))^2).$$

(Here X is meant to have the same distribution as all the X_i's.)

Let the range of $Q(.)$ be the set $\{x_1, \dots, x_N\}$, where the x_i's are real numbers. $Q(.)$ is uniquely defined by the values $x_1, \dots, x_N$ and the sets $B_i = \{x : Q(x) = x_i\}$. Once we fix $x_1, \dots, x_N$, we will have the smallest distortion $D(Q)$ if every x is "quantized" to the closest x_i, i.e.,

$$B_i = \{x : |x - x_i| \le |x - x_j| \ \forall j \ne i\}.$$

(Note that this rule puts some values into two neighboring B_i's (considering $x_1 < x_2 < \dots < x_N$, we have $x = \frac{1}{2}(x_i + x_{i+1})$ in both B_i and B_{i+1}). This can easily be resolved by saying that all these values go to (say) the smaller indexed B_i.)

If now we consider the B_i's fixed then the smallest distortion $D(Q)$ is obtained if the x_i values lie in the barycenter of the B_i, which is $E(X|B_i) := E(X|X \in B_i) = \frac{\int_{B_i} x f(x) dx}{\int_{B_i} f(x) dx}$, where $f(x)$ is the density function of the random variable X. (We will always assume that $f(x)$ has all the "nice" properties needed for the existence of the integrals we mention.)

We proved the previous claim, i.e. smallest distortion is achieved for given quantization intervals B_i when $Q(x) = E(X|B_i)$ for $x \in B_i$. Here you can find a different proof for the statement:

This holds for all B_i separately, so it is enough to show it for one of them. By the linearity of expectation

$$E((X - c)^2) = E(X^2) - c(2E(X) - c),$$

and this is smallest when $c(2E(X) - c)$ is largest. Since the sum of c and $2E(X) - c$ does not depend on c, one can see simply from the inequality between the arithmetic and geometric mean ($\frac{a+b}{2} \ge \sqrt{ab}$ with equality iff $a = b$) that this product is largest when $c = E(X)$. (At least this is the case if we can assume that both c and $2E(X) - c$ are non-negative and so the inequality $\frac{a+b}{2} \ge \sqrt{ab}$ can be used. If this is not the case, we can still easily obtain that $c(2E(X) - c)$ is maximized by $c = E(X)$ by looking at the derivatives.)

Lloyd-Max algorithm

The above suggests an iterative algorithm to find a good quantizer: We fix some quantization levels $x_1 < \cdots < x_N$ first and optimize for them the B_i domains by defining them as above: let $y_i = \frac{x_i + x_{i+1}}{2}$ for $i = 1, \ldots, N-1$ and

$$B_1 := (-\infty, y_1], \quad B_i := (y_i, y_{i+1}],\ i = 2, \ldots, N-1, \quad B_N = (y_{N-1}, \infty).$$

Notice that in general there is no reason for the x_i's to be automatically the barycenters of the domains B_i obtained in the previous step. So now we can consider these domains B_i fixed and optimize the quantization levels with respect to them by redefining them as the corresponding barycenters:

$$x_i := \frac{\int_{B_i} x f(x) dx}{\int_{B_i} f(x) dx}.$$

Now we can consider again the so-obtained x_i's fixed and redefine the B_i's for them, and so on. After each step (or after each "odd" step when we optimize the B_i domains for the actual x_i's) we can check whether the current distortion is below a certain threshold. If yes we stop the algorithm, if no, then we continue with further iterations.

The distortion is non-increasing in each step, therefore it converges to somewhere since it is non-negative. The problem is that not necessarily to the global optimum, the limit might as well be a local optimum. To solve this the algorithm can be started from different initial quantization levels and then the one with smallest distortion is chosen.

It should be clear from the above that if either of the two steps above changes the x_i quantization levels or the B_i domains, then the quantizer before that step was not optimal. It is possible, that no such change is attainable already and the quantizer is still not optimal.

A quantizer is called a Lloyd-Max quantizer if the two steps of the Lloyd-Max algorithm have no effect on them.

Example
Let X be a random variable that takes its values on the finite set $\{1, 2, 3, 4\}$ with uniform distribution. (That is $P(X = 1) = P(X = 2) = P(X = 3) = P(X = 4) = 1/4$.) Let $N = 2$ that is we are allowed to use two values for the quantization. There are three different quantizers for which neither of the above steps can cause any improvement, but only one of them is optimal.
These three quantizers can be described by

$$Q_1(1) = 1, Q_1(2) = Q_1(3) = Q_1(4) = 3;$$

$$Q_2(1) = Q_2(2) = 1.5, Q_2(3) = Q_2(4) = 3.5;$$

$$Q_3(1) = Q_3(2) = Q_3(3) = 2, Q_3(4) = 4.$$

Home-work 1:
Calculate the distortion of the above three quantizers.

Remember that we call a quantizer a Lloyd-Max quantizer if the two steps of the Lloyd-Max algorithm have no effect on them. In the previous example we have seen that a Lloyd-Max quantizer is not necessarily optimal. Fleischer gave a sufficient condition for the optimality of a Lloyd-Max quantizer. It is in terms of the density function $f(x)$ of the random variable to be quantized. In particular, it requires that $\log f(x)$ is concave.

Home-work 2: Let X be a random variable with density function

$$f(x) = \begin{cases} \frac{3x^2}{8}, & \text{if } x \in [0, 2] \\ 0, & \text{otherwise.} \end{cases}$$

The source is quantized by a 2-level quantizer. Starting from the initial levels $\frac{1}{2}$ and $\frac{3}{2}$, give the first iteration (first two steps) of the Lloyd-Max algorithm.

The above condition of Fleischer is satisfied by the density function of a random variable uniformly distributed in an interval $[a, b]$. Thus a corollary of Fleischer's theorem is that there is only one Lloyd-Max quantizer with N levels for the uniform distribution on $[a, b]$. It is not hard to see that this should be the uniform quantizer: the one belonging to $B_i = \{x : a+(i-1)\frac{b-a}{N} \leq x \leq a+i\frac{b-a}{N}\}$ and quantization levels at the middle of these intervals. (The extreme points of the intervals belonging to two neighboring B_i's can be freely decided to belong to either of them.)