# Complements

September 10, 2026

## 1 On unique decodability, by Pacôme (2025)

Let $C$ be a code, and let

$$\mathrm{Rev}(C) = \{\mathrm{rev}(w) : w \in C\},$$

where $\mathrm{rev}(w)$, with $w = (w_1 \dots w_k)$, denotes $(w_k w_{k-1} \cdots w_1)$.

**Theorem 1.** *A code $C$ is uniquely decodable if and only if* $\mathrm{Rev}(C)$ *is uniquely decodable.*

**Consequence.** If $C$ is prefix-free (hence uniquely decodable), then $\mathrm{Rev}(C)$ is uniquely decodable, even though $\mathrm{Rev}(C)$ need not be prefix-free. For instance, the code $\mathcal{C} = \{0, 01, 011\}$ is uniquely decodable.

*Proof.* Rev has the following two properties:

$$P_1 : \mathrm{rev}(\mathrm{rev}(u)) = u \quad \text{for all words } u.$$

$$P_2 : \mathrm{rev}(uv) = \mathrm{rev}(v)\ \mathrm{rev}(u) \quad \text{for all words } u, v.$$

By $P_1$ we have $\mathrm{Rev}(\mathrm{Rev}(C)) = C$ and so we only need to prove one direction, say "⇒." Suppose $C$ is uniquely decodable and assume, for contradiction, that $\mathrm{Rev}(C)$ is not. Then there exist two distinct finite sequences of codewords $a_1, a_2, \dots, a_m$ and $b_1, b_2, \dots, b_n$ in $\mathrm{Rev}(C)$ such that

$$a_1 a_2 \cdots a_m = b_1 b_2 \cdots b_n.$$

Let $x_i = \mathrm{rev}(a_i)$ and let $y_j = \mathrm{rev}(b_j)$. Apply rev to the equality and use $P_2$ iteratively to get

$$\begin{aligned} \mathrm{rev}(a_1 \cdots a_m) &= \mathrm{rev}(a_m) \cdots \mathrm{rev}(a_1) \\ &= x_m \cdots x_1, \end{aligned}$$

and

$$\mathrm{rev}(b_1 \cdots b_n) = y_n \cdots y_1.$$

Since $a_1 \cdots a_m = b_1 \cdots b_n$,

$$x_m \cdots x_1 = y_n \cdots y_1.$$

This is an equality of two different concatenations of codewords of $C$, contradicting the unique decodability of $C$.

□

## 2 On $H(X) \le \log|\mathcal{X}|$, by Mathys (2025)

We saw in class that given a random variable $X$ with alphabet $\mathcal{X}$ we have

$$H(X) \le \log|\mathcal{X}| \tag{1}$$

with equality if and only if $X \sim \mathrm{Uniform}(\mathcal{X})$.

An alternative way to derive (1) is by leveraging the fact that an adaptive questioning strategy for identifying $X$ is equivalent to a prefix-free code; a strategy yields a prefix-free code and vice-versa, and the expected codeword length corresponds to the expected number of questions.

From this equivalence, we know that the entropy $H(X)$ is a lower bound to the minimum expected number $M$ of questions needed to identify $X$. In turn, we know through dichotomy search that $M \le \lceil \log|\mathcal{X}| \rceil$. [1] Hence,

$$H(X) \le \lceil \log|\mathcal{X}| \rceil,$$

which recovers (1) except for the roundings—notice that $\lceil \log|\mathcal{X}| \rceil$ is a worst-case upper bound for dichotomy.

[1] Each question splits the remaining set into two subsets of (as close as possible) equal cardinality, ignoring probabilities.