# 3 Kraft's inequality

Let $\{0,1\}^*$ be the set of (finite) strings over the alphabet $\{0,1\}$, which we call *words*. A collection of words is a *code*.

**Definition 3.1** (Prefix-free code)**.** A finite set $S$ of words is *prefix-free* if whenever $x, y \in S$ and $x$ is a prefix of $y$ (that is, we can write $y = xz$ for some word $z$), then $x = y$.

**Theorem 3.1** (Kraft's inequality)**.** *If $S$ is a finite prefix-free code then*

$$\sum_{w \in S} 2^{-|w|} \leq 1,$$

*where $|w|$ is the length of $w$.*

*Proof.* Let $\ell = \max_{w \in S} |w|$ be the maximum length of a word in $S$. Consider the set $\{0,1\}^\ell$ of all words of length $\ell$.

For every $w \in S$, let $E(w) = \{wz : z \in \{0,1\}^{\ell - |w|}\} \subseteq \{0,1\}^\ell$. We claim that the sets $E(w)$ are disjoint.

Indeed, suppose that $x \in E(w_1) \cap E(w_2)$. Then $x = w_1 z_1 = w_2 z_2$. Let $s = \min(|w_1|, |w_2|)$. The first $s$ letters of $w_1, w_2$ are identical (since both are equal to the first $s$ letters of $x$). If $|w_1| = |w_2|$ then this implies that $w_1 = w_2$. If $|w_1| < |w_2|$ then this implies that $w_1$ is a proper prefix of $w_2$, which is impossible; and similarly $|w_1| > |w_2|$ leads to a contradiction.

Since the sets $E(w)$ are disjoint, we have

$$2^\ell = |\{0,1\}^\ell| \geq \sum_{w \in S} |E(w)| = \sum_{w \in S} 2^{\ell - |w|}.$$

Dividing by $2^\ell$ completes the proof. □

*Remark* 3.1.

1. The inequality holds also for infinite $S$, since any subset of a prefix-free code is prefix-free.
2. The inequality holds for any finite alphabet, with 2 replaced by the size of the alphabet.
3. The converse of the inequality also holds: if $\sum_i 2^{-\ell_i} \leq 1$ then there exists a prefix-free code $\{w_i\}$ with $|w_i| = \ell_i$. This holds even in the infinite case.

## 3.1 Alternative argument

Here is an alternative proof of Kraft's inequality, by induction. We use the following notation: $\epsilon$ is the empty word.

*Proof of Theorem 3.1.* We can assume that $S$ is non-empty, since otherwise the inequality trivially holds.

The proof is by strong induction on the length of the longest codeword in $S$.

If the length of the longest codeword in $S$ is 0, then $S = \{\epsilon\}$. Consequently

$$\sum_{w \in S} 2^{-|w|} = 2^{-0} = 1.$$

Suppose next that the length of the longest codeword in $S$ is $\ell + 1$. We claim that $\epsilon \notin S$. Indeed, by assumption, there exists $w \in S$ of length $\ell + 1$. If also $\epsilon \in S$, then $S$ is not prefix-free, since $\epsilon$ is a prefix of $w$ but $\epsilon \neq w$ since $|\epsilon| = 0$ while $|w| = \ell + 1 \geq 1$.

Define

$$S_0 = \{w \in \{0,1\}^* : 0w \in S\},$$
$$S_1 = \{w \in \{0,1\}^* : 1w \in S\}.$$

We claim that $S_0, S_1$ are prefix-free. Indeed, suppose $x, y \in S_b$, where $b \in \{0,1\}$, and $x$ is a prefix of $y$, say $y = xz$. By definition of $S_b$, we have $bx, by \in S$. Since $by = bxz$, we see that $bx$ is a prefix of $by$. Since $S$ is prefix-free, $bx = by$, hence $x = y$.

For $b \in \{0,1\}$, let $S'_b$ consist of these words in $S$ which start with $b$. Since $\epsilon \notin S$, every word in $S$ is either in $S'_0$ or in $S'_1$. Also, $S'_0$ and $S'_1$ are disjoint, since if $w \in S'_0 \cap S'_1$ then $w_1 = 0$ and $w_1 = 1$, which is impossible. Hence

$$\sum_{w \in S} 2^{-|w|} = \sum_{w \in S'_0} 2^{-|w|} + \sum_{w \in S'_1} 2^{-|w|}.$$

The mapping $w \mapsto bw$ is a bijection between $S_b$ and $S'_b$. Indeed, it is one-to-one: if $bx = by$ then $x = y$; and it is onto by definition of $S_b, S'_b$. Therefore

$$\sum_{w \in S} 2^{-|w|} = \sum_{w \in S_0} 2^{-|0w|} + \sum_{w \in S_1} 2^{-|1w|} = \frac{1}{2} \sum_{w \in S_0} 2^{-|w|} + \frac{1}{2} \sum_{w \in S_1} 2^{-|w|}.$$

We claim that $S_b$ is either empty, or the length of the longest word in it is strictly smaller than the length of the longest word in $S$. Indeed, if $w \in S_b$ then $bw \in S$, which is longer. Therefore we can apply the inductive hypothesis (considering separately the case in which $S_b$ is empty) to deduce that

$$\sum_{w \in S} 2^{-|w|} \leq \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 1 = 1. \qquad \square$$

## 3.2 Converse

Kraft's inequality is tight in the following sense.

**Theorem 3.2.** *Let $I$ be a finite set and let $\ell\colon I \to \mathbb{N}$ satisfy*

$$\sum_{i\in I} 2^{-\ell(i)} \leq 1.$$

*There exists an injective mapping $w\colon I \to \{0,1\}^*$ whose image is prefix-free, and furthermore $|w(i)| = \ell(i)$.*

The proof will require the following auxiliary result.

**Lemma 3.1.** *Let $I$ be a finite set and let $\ell\colon I \to \mathbb{N}$ satisfy*

$$\sum_{i\in I} 2^{-\ell(i)} \geq 1.$$

*There exists a subset $S \subseteq I$ such that*

$$\sum_{i\in S} 2^{-\ell(i)} = 1.$$

*Proof.* Let $n = |I|$ and let $\ell'_1, \ldots, \ell'_n$ consist of $(\ell(i))_{i\in I}$ arranged in nonincreasing order: there exists a bijection $\pi\colon \{1,\ldots,n\} \to I$ such that $\ell'_i = \ell(\pi(i))$, and for every $1 \leq i \leq j \leq n$ we have $\ell'_i \geq \ell'_j$.

Observe that

$$\sum_{i=1}^{n} 2^{-\ell'_i} = \sum_{i\in I} 2^{-\ell(i)} \geq 1.$$

Let $i_0$ be the minimal $i$ such that

$$\sum_{i=1}^{i_0} 2^{-\ell'_i} \geq 1.$$

Clearly $i_0 > 0$, since $\sum_{i=1}^{0} 2^{-\ell'_i} = 0$. Therefore

$$\sum_{i=1}^{i_0-1} 2^{-\ell'_i} < 1.$$

Multiply both inequalities by $2^{\ell'_{i_0}}$:

$$\sum_{i=1}^{i_0} 2^{\ell'_{i_0}-\ell'_i} \geq 2^{\ell'_{i_0}},$$
$$\sum_{i=1}^{i_0-1} 2^{\ell'_{i_0}-\ell'_i} < 2^{\ell'_{i_0}}.$$

Since $\ell'_1, \ldots, \ell'_n$ is nonincreasing, we have $\ell'_{i_0} \geq \ell'_i$ for all $1 \leq i \leq i_0$, and so $2^{\ell'_{i_0}-\ell'_i}$ is an integer for all such expressions appearing above. The second inequality above therefore can be strengthened to

$$\sum_{i=1}^{i_0-1} 2^{\ell'_{i_0}-\ell'_i} \leq 2^{\ell'_{i_0}} - 1,$$

which implies that

$$\sum_{i=1}^{i_0} 2^{\ell'_{i_0}-\ell'_i} = \sum_{i=1}^{i_0-1} 2^{\ell'_{i_0}-\ell'_i} + 2^{\ell'_{i_0}-\ell'_{i_0}} = \sum_{i=1}^{i_0-1} 2^{\ell'_{i_0}-\ell'_i} + 1 \leq 2^{\ell'_{i_0}}.$$

Therefore

$$\sum_{i=1}^{i_0} 2^{\ell'_{i_0}-\ell_i} = 2^{\ell'_{i_0}},$$

and so

$$\sum_{i=1}^{i_0} 2^{-\ell'_i} = 1.$$

Taking $S = \{\pi(i) : 1 \le i \le i_0\}$ completes the proof, since

$$\sum_{i\in S} 2^{-\ell(i)} = \sum_{i=1}^{i_0} 2^{-\ell(\pi(i))} = \sum_{i=1}^{i_0} 2^{-\ell'_i} = 1. \qquad \square$$

**Corollary 3.1.** *Let $I$ be a finite set and let $\ell\colon I \to \mathbb{N}$ satisfy*

$$\sum_{i\in I} 2^{-\ell(i)} \ge \frac{1}{2}.$$

*If $\ell(i) > 0$ for all $i$ then there exists a subset $S \subseteq I$ such that*

$$\sum_{i\in S} 2^{-\ell(i)} = 1.$$

*Proof.* Let $\ell'(i) = \ell(i) - 1$. Since $\ell(i) > 0$, we have

$$\sum_{i\in I} 2^{-\ell'(i)} = 2\sum_{i\in I} 2^{-\ell(i)} \ge 2 \cdot \frac{1}{2} = 1.$$

Therefore the theorem implies that there exists a set $S \subseteq I$ such that

$$\sum_{i\in S} 2^{-\ell(i)} = \frac{1}{2}\sum_{i\in S} 2^{-\ell'(i)} = \frac{1}{2} \cdot 1 = \frac{1}{2}. \qquad \square$$

*Proof of Theorem 3.2.* We can assume that $I \neq \emptyset$, otherwise we can simply output the empty mapping.

The proof is by strong induction on $\max_{i\in I} \ell(i)$.

If $\max_{i\in I} \ell(i) = 0$ then $\ell(i) = 0$ for all $i \in I$, hence $\sum_{i\in I} 2^{-\ell(i)} = \sum_{i\in I} 1 = |I|$, implying that $|I| \le 1$. Since also $|I| \ge 1$, we have $|I| = 1$, say $I = \{i_0\}$. In this case we can take $w(i_0) = \epsilon$. This is clearly injective, a prefix-free code, and satisfies $|w(i_0)| = \ell(i_0)$.

From now on, we assume that $\max_{i\in I} \ell(i) > 0$. This implies that $\ell(i) > 0$ for all $i \in I$. Indeed, suppose that $\ell(i) = 0$. Since $\max_{i\in I} \ell(i) > 0$, there exists $j \in I$ such that $\ell(j) > 0$; note $i \neq j$. Then $\sum_{k\in I} 2^{-\ell(k)} \ge 2^{-\ell(i)} + 2^{-\ell(j)} > 2^{-\ell(i)} = 1$, which is impossible.

Next, we claim that we can find a subset $S \subseteq I$ such that

$$\sum_{i\in S} 2^{-\ell(i)} \le \frac{1}{2},$$
$$\sum_{i\in S^c} 2^{-\ell(i)} \le \frac{1}{2}.$$

For the proof, we consider two cases. If

$$\sum_{i\in I} 2^{-\ell(i)} \le \frac{1}{2}$$

then we take $S = I$. The first condition holds by assumption, the second since the sum is zero.

Suppose next that

$$\sum_{i \in I} 2^{-\ell(i)} > \frac{1}{2}.$$

In this case we apply Corollary 3.1 to obtain a set $S \subseteq I$ such that

$$\sum_{i \in S} 2^{-\ell(i)} = \frac{1}{2}.$$

Since

$$1 \geq \sum_{i \in I} 2^{-\ell(i)} = \sum_{i \in S} 2^{-\ell(i)} + \sum_{i \in S^c} 2^{-\ell(i)} = \frac{1}{2} + \sum_{i \in S^c} 2^{-\ell(i)},$$

this implies that

$$\sum_{i \in S^c} 2^{-\ell(i)} \leq \frac{1}{2},$$

as needed.

Define $\ell' \colon I \to \mathbb{N}$ by $\ell'(i) = \ell(i) - 1$. Then

$$\sum_{i \in S} 2^{-\ell'(i)} = 2 \sum_{i \in S} 2^{-\ell(i)} = 1,$$
$$\sum_{i \in S^c} 2^{-\ell'(i)} = 2 \sum_{i \in S^c} 2^{-\ell(i)} \leq 1.$$

Also, $\max_{i \in I} \ell'(i) = \max_{i \in I} \ell(i) - 1$, implying that $\max_{i \in S} \ell'(i) \leq \max_{i \in I} \ell'(i) < \max_{i \in I} \ell(i)$ and $\max_{i \in S^c} \ell'(i) \leq \max_{i \in I} \ell'(i) < \max_{i \in I} \ell(i)$.

Therefore we can apply the induction hypothesis twice, on $\ell'|_S$ and on $\ell'|_{S^c}$, to obtain injective maps $w_0 \colon S \to \{0,1\}^*$ and $w_1 \colon S^c \to \{0,1\}^*$ such that $w_0(S)$ and $w_1(S)$ are prefix-free; $|w_0(i)| = \ell'(i)$ for all $i \in S$; and $|w_1(i)| = \ell'(i)$ for all $i \in S^c$.

Define $w \colon I \to \{0,1\}^*$ by

$$w(i) = \begin{cases} 0w_0(i) & \text{if } i \in S, \\ 1w_1(i) & \text{otherwise.} \end{cases}$$

We claim that $w$ is injective. Indeed, if $w(i) = w(j)$ then either $i, j \in S$ or $i, j \notin S$ (considering the first symbol). If $i, j \in S$ then $w_0(i) = w_0(j)$, and so $i = j$ since $w_0$ is injective. If $i, j \notin S$ then $w_1(i) = w_1(j)$, and so $i = j$ since $w_1$ is injective.

We claim that $w(I)$ is prefix-free. Indeed, suppose that $w(i)$ is a prefix of $w(j)$. Considering the first symbol, either $i, j \in S$ or $i, j \notin S$. If $i, j \in S$ then $w_0(i)$ is a prefix of $w_0(j)$, and so $i = j$ since $w_0(S)$ is prefix-free. If $i, j \notin S$ then $w_1(i)$ is a prefix of $w_1(j)$, and so $i = j$ since $w_1(S^c)$ is prefix-free.

Finally, let us show that $|w(i)| = \ell(i)$ for all $i \in I$. If $i \in S$ then $|w(i)| = |w_0(i)| + 1 = \ell'(i) + 1 = \ell(i)$. If $i \notin S$ then $|w(i)| = |w_1(i)| + 1 = \ell'(i) + 1 = \ell(i)$. $\square$

## 3.3 Kraft–McMillan inequality

Kraft's inequality extends to the more general setting of uniquely decodable codes.

**Definition 3.2** (Uniquely decodable code)**.** A finite set $S$ of words is *uniquely decodable* if every $x \in \{0,1\}^*$ can be written in at most one way as $x = w_1 \dots w_r$ (for any $r$), where $w_1, \dots, w_r \in S$.

**Theorem 3.3** (Kraft–McMillan inequality)**.** *If $S$ is a finite uniquely decodable code then*

$$\sum_{w \in S} 2^{-|w|} \leq 1,$$

*where $|w|$ is the length of $w$.*

*Proof.* Let $\ell$ again denote the maximum length of a word in $S$. Observe that $\ell \geq 1$. Indeed, $\epsilon \notin S$, since otherwise $\epsilon = \epsilon\epsilon$ would be two ways to "decode" $\epsilon$ (here $\epsilon$ is the empty word).

Define

$$C = \sum_{w \in S} 2^{-|w|}.$$

We would like to show that $|C| \leq 1$.

For every $r \geq 1$, consider

$$C^r = \left( \sum_{w \in S} 2^{-|w|} \right)^r = \sum_{w_1, \dots, w_r \in S} \prod_{i=1}^{r} 2^{-|w_i|} = \sum_{w_1, \dots, w_r \in S} 2^{-\sum_{i=1}^{r} |w_i|} = \sum_{w_1, \dots, w_r \in S} 2^{-|w_1 \cdots w_r|}.$$

Unique decodability implies that the mapping $(w_1, \dots, w_r) \mapsto w_1 \cdots w_r$ is an injective mapping from $S^r$ to $\{0,1\}^*$. In fact, since $|w_1 \cdots w_r| = \sum_{i=1}^{r} |w_i| \leq \sum_{i=1}^{r} \ell = r\ell$ and similarly $|w_1 \cdots w_r| \geq r$, it is an injective mapping from $S^r$ to $\bigcup_{s=r}^{r\ell} \{0,1\}^s$. It follows that

$$C^r = \sum_{w_1, \dots, w_r \in S} 2^{-|w_1 \cdots w_r|} \leq \sum_{s=r}^{r\ell} \sum_{x \in \{0,1\}^s} 2^{-|x|} = \sum_{s=r}^{r\ell} 2^s 2^{-s} = r\ell - r + 1 \overset{r \geq 1}{\leq} r\ell.$$

Taking an $r$'th root, we have

$$C \leq (r\ell)^{1/r},$$

and so

$$\log_2 C \leq \frac{\log_2 r + \log \ell}{r}.$$

Let $r = 2^{2t}$. Then

$$\log_2 r = 2t \overset{\text{Cantor}}{<} 2 \cdot 2^t = 2^{t+1}.$$

Therefore if $2^{t+1} \geq \log_2 \ell$ then

$$\log_2 C \leq \frac{2t + \log_2 \ell}{2^{2t}} < \frac{2^{t+1} + 2^{t+1}}{2^{2t}} = \frac{2^{t+2}}{2^{2t}} = \frac{1}{2^{t-2}}.$$

We claim that this implies that $\log_2 C \leq 0$, and so $C \leq 1$. Indeed, if not, then we can find $t$ such that $2^{t+1} \geq \log_2 \ell$ and $2^{t-2} > 1/\log_2 C$, and reach a contradiction. We can simply take

$$t = 2 + \log_2(1 + \max(\log_2 \ell, 1/\log_2 C)).$$

We leave the routine verification to the reader. □

The remarks mentioned for Kraft's inequality are relevant in this more general setting as well.

## 3.4 Unique decodability of prefix-free codes

Prefix-free codes are uniquely decodable (the converse doesn't hold: for example, $\{0, 00, 01\}$ is uniquely decodable since it is the reverse of a prefix-free code).

**Theorem 3.4.** *If a finite set $S$ of words is prefix-free and $\epsilon \notin S$ then it is uniquely decodable.*

*Proof.* We prove by strong induction on $|x|$ that every word $x \in \{0,1\}^*$ can be written in at most one way as $x = w_1 \dots w_r$ (for any $r$), where $w_1, \dots, w_r \in S$.

The base case is when $|x| = 0$. In this case $x = \epsilon$. If $x = w_1 \dots w_r$ then $w_1 = \dots w_r = \epsilon$. Since $\epsilon \notin S$, we must have $r = 0$. Therefore the representation is unique.

Now suppose that $|x| > 0$, and consider any two representations $x = w_1 \dots w_r = z_1 \dots z_s$. Since $|x| > 0$, we must have $r, s \geq 1$.

Suppose without loss of generality that $|w_1| \leq |z_1|$. We claim that $w_1$ is a prefix of $z_1$. Indeed, $z_1 = w_1 q$, where $q = x_{|z_1|+1} \dots x_{|w_1|}$. Since $S$ is prefix-free, $w_1 = z_1$. This implies that $w_2 \dots w_r = z_2 \dots z_s$. Denote the common value by $y$.

Since $|x| = |z_1| + |y|$ and $z_1 \neq \epsilon$ (since $z_1 \in S$ and $\epsilon \notin S$), we have $|y| < |x|$, and so the induction hypothesis implies that $r = s$ and $w_i = z_i$ for all $i > 1$. Therefore $w_i = z_i$ for all $i$, completing the inductive proof. □

**Corollary 3.2.** *If a finite set $S$ of words is prefix-free and $|S| \geq 2$ then it is uniquely decodable.*

*Proof.* It suffices to show that $\epsilon \notin S$. Indeed, suppose that $\epsilon \in S$. Since $|S| \geq 2$, there exists $w \in S$ such that $w \neq \epsilon$. Since $\epsilon$ is a prefix of $w$, we reach a contradiction. □