Kraft's and McMillan's Inequalities

Phil Pützstück, 377247
Proseminar Informationstheorie

October 10, 2018

(1.0) Assumptions

- Basic Graph-theory (Trees, acyclic, directed, height, etc.).
- Familiarity with instantaneous Codes, as defined in [JJ00].

(1.1) Definition (r-ary Trees from r-ary Codes)**.**
Let $q, r \in \mathbb{N}, [0, r-1]$ be the Code-Alphabet for some r-ary Code $\mathcal{C}$ with word-lengths $l \in \mathbb{N}^q$. Set $h := \max\{l_i \mid i \in [1, q]\}$. Define $W := \bigcup_{i \in [0,h]} [0, r-1]^i$ to be the set of all words over T of maximum length h. Thus $\mathcal{C} \subseteq W$. We define a rooted r-ary tree as a directed graph:

$$V := \{v_w \mid w \in W\} \qquad E := \{(v_w, v_{w'}) \mid v_w, v_{w'} \in V \wedge w' = wx, x \in [0, r-1]\}$$

Which means we have a vertex for each word in W, and an edge from v_w to $v_{w'}$ iff w is a prefix of w' with $|w| = |w'| - 1$. We set $\mathcal{T}_r^h := (V, E)$ as the rooted r-ary tree of height h. The root $R(\mathcal{T}_r^h)$ is given by v_ε, since $\varepsilon \in T^0 \subseteq W$ and $\varepsilon \sqsubseteq w$ for all $w \in W$. We denote $V(\mathcal{T}_r^h) := V$ and $E(\mathcal{T}_r^h) := E$. We can easily define the height $\mathcal{H}_{\mathcal{T}_r^h}(v_w) := |w| - |\mathcal{W}(R(\mathcal{T}_r^h))|$, where $\mathcal{W}(v_w) = w$, which defines the height of a vertex $v_w \in V(\mathcal{T}_r^h)$ as the length of w minus the length of the word at the root of the tree, which is usually $|\varepsilon| = 0$, but can be different.

(1.2) Definition (Subtrees and Ordering)**.**
Let $h, r \in \mathbb{N}, v_w, v_{w'} \in V(\mathcal{T}_r^h)$. We say T is a rooted subtree of $\mathcal{T}_r^h$, written $T \leq \mathcal{T}_r^h$, iff $V(T) \subseteq V(\mathcal{T}_r^h), E(T) \subseteq E(\mathcal{T}_r^h)$ and T fullfills the standard criteria of a rooted (directed) tree.

We say T is a rooted r-ary subtree of $\mathcal{T}_r^h$, written $T \leq_r \mathcal{T}_r^h$ iff $T \leq \mathcal{T}_r^h$ and T is r-ary.

We write $v_w \leq v_{w'}$ iff $w \sqsubseteq w'$. Now let $T \leq \mathcal{T}_r^h$ be a rooted subtree and $v_w \in T \setminus \{R(T)\}$.

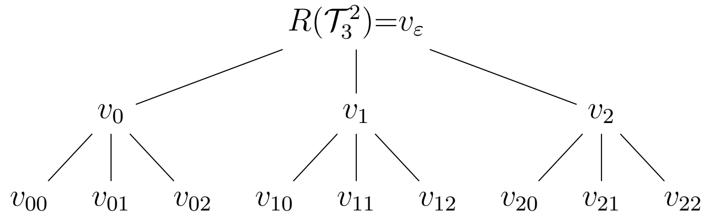
$$V := \{v \in V(T) \mid v_w \leq v\} = \{v_{w'} \in V(T) \mid w \sqsubseteq w'\} \quad E := \{(v, v') \in V(T) \mid v, v' \in V\}$$

If we have $(V, E) \leq_r \mathcal{T}_r^h$, meaning the Graph (V, E) is a rooted r-ary subtree of $\mathcal{T}_r^h$, then we define

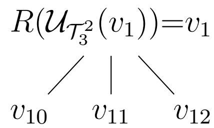
$$\mathcal{U}_T(v_w) := \mathcal{U}_{\mathcal{T}_r^h}(v_w) = (V, E) \quad \text{and} \quad T \setminus v_w := T \setminus \mathcal{U}_T(v_w) = (V(T) \setminus V, E(T) \setminus E)$$

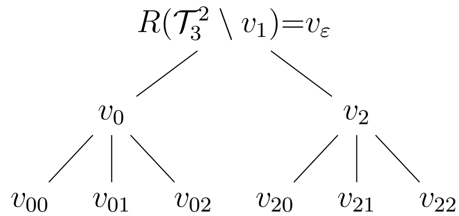
(1.3) Examples

$\mathcal{T}_3^2$ is given by



We have $\mathcal{H}_{\mathcal{T}_3^2}(v_\varepsilon) = 0$ and $\mathcal{H}_{\mathcal{T}_3^2}(v_{12}) = 2$. $v_0 \leq v_{02}$ holds, $v_0 \leq v_{10}$ does **not**.
The subtree $\mathcal{U}_{\mathcal{T}_3^2}(v_1)$ and $\mathcal{T}_3^2 \setminus v_1$ are given by:





Note $\mathcal{U}_{\mathcal{T}_r^h}(v_1)$ is a 3-ary rooted subtree of height 1, but $\mathcal{T}_3^2 \setminus v_1$ is only a rooted subtree of height 3, not r-ary for any $r \in \mathbb{N}$. We now have $\mathcal{H}_{\mathcal{U}_{\mathcal{T}_3^2}}(v_1) = 0, \mathcal{H}_{\mathcal{U}_{\mathcal{T}_3^2}}(v_{12}) = 1$, but still $\mathcal{H}_{\mathcal{T}_3^2 \setminus v_1}(v) = \mathcal{H}_{\mathcal{T}_3^2}(v)$ for $v \in V(\mathcal{T}_3^2 \setminus v_1)$.

(1.4) Proposition (Number of nodes in rooted r-ary subtrees)**.**
Let $h, r \in \mathbb{N}, T \leq_r \mathcal{T}_r^h$ be a rooted r-ary subtree of $\mathcal{T}_r^h$ with height $h' \leq h$. Then T has exactly r^l vertices of height l for $l \in [0, h']$.

Proof. Left as exercise for the reader. □

(1.5) Corollary (Number of leafs of $T \setminus v$)**.**
Let $h, r \in \mathbb{N}, T \leq \mathcal{T}_r^h, v_w \in V(T) \setminus \{R(\mathcal{T}_r^h)\}$ such that $\mathcal{U}_T(v_w)$ is well defined, in particular r-ary. Let $L \leq r^h$ be the number of leaves of T. Then $T \setminus v_w$ has $L - r^{h-|w|}$ leaves.

Proof. Since $\mathcal{U}_T(v_w)$ has height $h - \mathcal{H}_{\mathcal{T}_r^h}(v_w) = h - |w|$, we know $\mathcal{U}_{\mathcal{T}_r^h}(v)$ has $r^{h-|w|}$ leaves by (1.4). Thus $T \setminus v_w = T \setminus \mathcal{U}_T(v_w)$ has $L - r^{h-|w|}$ leaves. □

(1.6) Theorem (Kraft's Inequality)**.**
Let $q, r \in \mathbb{N}, l \in \mathbb{N}^q$. Then there is an instantaneous r-ary Code $\mathcal{C}$ with word-lengths l iff

$$\sum_{k=1}^{q} \frac{1}{r^{l_k}} \leq 1 \tag{1}$$

Proof. If $q = 1$, then we always have an instantaneous Code, and since $r \in \mathbb{N}$, (1) always holds as well. So assume WLOG that $q > 1$ and $\forall i \in [1, q-1] : 0 < l_i \leq l_{i+1}$. Furthermore we can assume WLOG that the Code-Alphabet of $\mathcal{C}$ is $[0, r-1]$, since any other Alphabet of length r is in bijection to this.

We first show that (1) implies the existence of an r-ary prefix-Code, which by [JJ00] is instantaneous. Set $h := l_q$ to be the maximum length of the supposed Code-words of $\mathcal{C}$. Thus we should have, like in (1.1), that $\mathcal{C} \subseteq \bigcup_{i\in[0,h]}[0, r-1]^i =: W$, where W is in bijection with $V(\mathcal{T}_r^h)$. So we construct the Code-words w_i of the prefix-Code $\mathcal{C}$, with $|w_i| = l_i$ for $i \in [1, q]$ via finite induction over i.

Let $i = 1$. Choose a Code-word $w_1 \in [1, r]^{l_1}$ of length l_1. Since $w_1 \in W$ and $l_1 > 0$ we have $v_{w_1} \in V(\mathcal{T}_r^h) \setminus \{R(\mathcal{T}_r^h)\}$. Define $\mathcal{T}_1 := \mathcal{T}_r^h \setminus v_{w_1}$. We know from (1.5) that $\mathcal{T}_1$ has

$$r^h - r^{h-l_1} = r^h \left(1 - \sum_{k=1}^{1} \frac{1}{r^{l_k}}\right) \overset{q>1}{>} r^h \left(1 - \sum_{k=1}^{q} \frac{1}{r^{l_k}}\right) \overset{(1)}{\geq} 0$$

leaves. Now let $i \in [1, q-1]$ such that $\mathcal{C} := \{w_j \mid j \in [1, i]\}$ is a prefix-Code with $|w_j| = l_j$ for $j \in [1, i]$ and such that $\mathcal{T}_i$ is a rooted subtree of $\mathcal{T}_r^h$ and has $r^h(1 - \sum_{k=1}^{i} \frac{1}{r^{l_k}}) > 0$ leaves. Then since $l_{i+1} \leq l_q = h$ we know that there must also be at least one vertex $v_w \in V(\mathcal{T}_i)$ with $\mathcal{H}_{\mathcal{T}_i}(v_w) = \mathcal{H}_{\mathcal{T}_r^h}(v_w) = l_{i+1} \implies |w| = l_{i+1}$ (since trees are connected). So set $w_{i+1} := w$. If we had $w_j \sqsubseteq w_{i+1}$ for some $j \in [1, i]$, then it would follow that $v_{w_j} \leq v_{w_{i+1}}$, but then we would have $v_{w_{i+1}} \notin V(\mathcal{T}_j) \subseteq V(\mathcal{T}_i)$, a contradiction. Thus $\mathcal{C} := \{w_j \mid j \in [1, i+1]\}$ is still a prefix-Code. If $i + 1 = q$ we are done, as we have constructed the desired prefix-Code. Otherwise, we set $\mathcal{T}_{i+1} := \mathcal{T}_i \setminus w_{i+1}$ and we get for the number of leaves:

$$r^h \left(1 - \sum_{k=1}^{i} \frac{1}{r^{l_k}}\right) - r^{h-l_{i+1}} = r^h \left(1 - \sum_{k=1}^{i+1} \frac{1}{r^{l_k}}\right) > r^h \left(1 - \sum_{k=1}^{q} \frac{1}{r^{l_k}}\right) \overset{(1)}{\geq} 0$$

Thus we constructed the desired prefix-Code $\mathcal{C}$ by finite induction.

Now we show the existence of a instantaneous r-ary Code $\mathcal{C}$ with word-lengths l implies (1). We know from [JJ00] that $\mathcal{C}$ is a prefix-Code. Let

$$L_i := \{v_w \in V(\mathcal{T}_r^h) \mid w_i \sqsubseteq w \wedge |w| = h\} = \{v_w \in \mathcal{U}_{\mathcal{T}_r^h}(v_{w_i}) \mid \mathcal{H}_{\mathcal{T}_r^h}(v_w) = h - |w_i|\}$$

be the set of leaves in $\mathcal{U}_{\mathcal{T}_r^h}(v_{w_i})$, where $w_i \in \mathcal{C}$ with $|w_i| = l_i$ for $i \in [1, q]$. We know from (1.4) that $|L_i| = r^{h-l_i}$ for $i \in [1, q]$, as we have $\mathcal{H}_{\mathcal{U}_{\mathcal{T}_r^h}(v_{w_i})}(v_w) = h - l_i$ for $v_w \in L_i$. Furthermore we know that for each $i \neq j \in [1, q]$ $L_i \cap L_j = \varnothing$:
Assume $i, j \in [1, q]$ and WLOG $i < j$. Let $v_w \in L_i \cap L_j$. Thus we get

$$v_{w_i} \leq v_w \wedge v_{w_j} \leq v_w \implies w_i \sqsubseteq w \wedge w_j \sqsubseteq w \implies w_i \sqsubseteq w_j$$

which is a contradiction to the fact that $\mathcal{C}$ is a prefix-Code. So now, since $\mathcal{T}_r^h$ only has r^h leafs, we have

$$r^h \geq |\bigcup_{i\in[1,q]} L_i| = \sum_{i=1}^{q} |L_i| = \sum_{i=1}^{q} r^{h-l_i} = r^h \sum_{i=1}^{q} \frac{1}{r^{l_i}} \iff \sum_{i=1}^{q} \frac{1}{r^{l_i}} \leq 1$$

$\square$

(1.7) Theorem (McMillan's Inequality)**.**
Let $q, r \in \mathbb{N}, l \in \mathbb{N}^q$. Then there is an uniquely decodable r-ary Code $\mathcal{C}$ iff

$$\sum_{i=1}^{q} \frac{1}{r^{l_i}} \leq 1 \tag{1}$$

Proof. If we assume (1), then by Kraft's Inequality we know that $\mathcal{C}$ is instantaneous, which by [JJ00] implies unique decodability.

Now assume that $\mathcal{C}$ is a unique decodable r-ary Code with word-lengths l.
Let $K := \sum_{i=1}^{q} \frac{1}{r^{l_i}}$ and $n \in \mathbb{N}$. Then we have

$$K^n = \left(\sum_{i=1}^{q} \frac{1}{r^{l_i}} \right)^n = \sum_{i \in [1,q]^n} \prod_{k=1}^{n} \frac{1}{r^{l_{i_k}}} = \sum_{i \in [1,q]^n} r^{-\sum_{k=1}^{n} l_{i_k}} \tag{2}$$

where the $i \in [1,q]^n$ represents n choices of q possible summands (with repitition).

Now there are many different $i \in [1,q]^n$ which have the same sum $\sum_{k=1}^{n} l_{i_k}$ (consider permutations for example). Set $M := \max\{l_k \mid k \in [1,q]\}, m := \min\{l_k \mid k \in [1,q]\}$. Then we get $mn \leq \sum_{k=1}^{n} l_{i_k} \leq Mn$ for all $i \in [1,q]^n$. We define for $p \in \mathbb{N}, j \in [mn, Mn]$:

$$N_{j,p} := \{w_{i_1} w_{i_2} \cdots w_{i_p} \mid i \in [1,q]^n \wedge |w_{i_1} \cdots w_{i_n}| = j\}$$

So $t \in N_{p,j}$ is a Code-sequence of length j, consisting of p Code-words in $\mathcal{C}$.
But since $\mathcal{C}$ is uniquely decodable, we know that

$$\forall t \in N_{j,p} : \exists!\ i \in [1,q]^n : t = w_{i_1} \cdots w_{i_n}$$

$$\Longrightarrow |\{i \in [1,q]^n \mid \sum_{k=1}^{n} l_{i_k} = j\}| = |\{i \in [1,q]^n \mid \sum_{k=1}^{n} |w_{i_k}| = j\}| = |N_{j,p}| \leq r^j$$

Hence, continuing in (2), we now have

$$K^n = \sum_{j=mn}^{Mn} \frac{|N_{j,n}|}{r^j} \leq \sum_{j=mn}^{Mn} 1 = (l-m)n + 1 \Longrightarrow \frac{K^n}{n} \leq (M-m) + \frac{1}{n}$$

Now M, m, K are fixed, while n may be arbitrarily large. From Analysis we know that as $n \to \infty$, the only way that $\frac{K^n}{n}$ stays bounded is if $K \leq 1$. Thus we get the desired result:

$$\sum_{i=1}^{q} \frac{1}{r^{l_i}} = K \leq 1$$

$\square$

References

[JJ00] Gareth A. Jones and J. Mary Jones. *Information and Coding Theory*. 2000.