

# Kraft's and McMillan's Inequalities

Phil Pützstück, 377247

Proseminar Informationstheorie, Steffen van Bergerem

November 4, 2018

## Abstract

We study the existence of uniquely decodable or instantaneous  $r$ -ary codes for some given word-lengths. To do this, we prove and discuss the known Kraft and McMillan Inequalities by utilising graph theory. The approach is based on [JJ00].

## Introduction

As uniquely decodable  $r$ -ary codes and instantaneous codes are important concepts, we want to know under which constraints these exist. We specifically look at  $r$ -ary codes with given word-lengths, as the inequalities we will later prove relate these concepts. After introducing a certain rooted Tree we show the relation between it and  $r$ -ary codes, which we use in the proof for Kraft's Inequality. Following the proofs we discuss the implications of these inequalities and give an example. For a quick introduction to our graph-terminology:

Consider a tree, which is defined to be an acyclic, connected and undirected graph. A tree  $T = (V, E)$  has the set of vertices  $V$  and the set of edges  $E$ , which we denote by  $V(T) := V, E(T) := E$ . We call  $T$  a rooted tree, if we have a distinct vertex  $v \in V(T)$ , called the root of  $T$  and denoted as  $root(T) = v$ . Note that we then have a unique path from the root of  $T$  to each vertex of  $T$ . In this case each vertex  $v \in V(T)$  has a height, denoted  $height(v) = height_T(v)$ , defined as the length of that unique path from  $root(T)$  to  $v$ .

For the rest of this paper we will only consider rooted trees, in particular  $r$ -ary rooted trees of some height  $h$ , where  $r, h \in \mathbb{N}$ , which are rooted trees of finite height  $h$  where each vertex of height less than  $h$  has exactly  $r$  children. We introduce subtrees and an ordering of vertices:

### (1.1) Definition (Subtrees and Ordering).

Let  $T, T'$  be rooted Trees. We say  $T'$  is a rooted subtree of  $T$ , iff  $T'$  is a subgraph of  $T$  and write  $T' \leq T$ . We write  $T' \leq_r T$  iff  $r \in \mathbb{N}$  and  $T, T'$  are both  $r$ -ary.

Now let  $v, w \in V(T)$ . We write  $v \leq w$  iff the unique path from  $root(T)$  to  $w$  visits  $v$ .

Let  $v \in T \setminus \{root(T)\}$ . Set  $V_v := \{u \in V(T) \mid v \leq u\}$ ,  $E_v := \{(u, u') \in E(T) \mid u, u' \in V_v\}$  and let  $sub(v) = sub_T(v) := (V_v, E_v)$  be the rooted subtree of  $T$  with  $v$  as its root.

At last we define  $T \setminus v := T \setminus sub(v)$  by usual graph difference to be the rooted subtree of  $T$  where all vertices and edges in  $sub(v)$  (all vertices  $v' \in V(T)$  with  $v \leq v'$  and their edges) are deleted from  $T$ .

We will soon relate the height of a vertex to the length of a given word, and as we are searching for Codes with given word-lengths, it will be useful to state some reminders about counting vertices in trees.

**(1.2) Reminder** (Number of vertices of some height in rooted  $r$ -ary Trees).

Let  $h, r \in \mathbb{N}$  and  $T$  be a rooted  $r$ -ary tree of height  $h$ . Then  $T$  has exactly  $r^{h'}$  vertices of height  $h' \leq h$ .

**(1.3) Corollary** (Number of Leaves of  $T \setminus v$ ).

Let  $h, r \in \mathbb{N}$ ,  $T$  be a rooted  $r$ -ary tree of height  $h$ ,  $T' \leq T$  and  $v \in V(T') \setminus \{\text{root}(T')\}$  such that  $\text{sub}_{T'}(v) \leq_r T$ . If  $L$  is the number of leaves of  $T'$ , then  $T' \setminus v$  has  $L - r^{h - \text{height}_T(v)}$  leaves.

*Proof.* Since  $\text{sub}_{T'}(v)$  is  $r$ -ary and has height  $h - \text{height}_T(v)$ , we know  $\text{sub}_{T'}(v)$  has  $r^{h - \text{height}_T(v)}$  leaves by (1.2). Thus  $T' \setminus v = T' \setminus \text{sub}_{T'}(v)$  has  $L - r^{h - \text{height}_T(v)}$  leaves.  $\square$

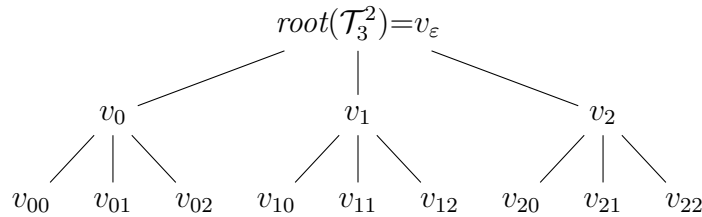
For us, intervals are over  $\mathbb{N}_0$ , so for  $m, n \in \mathbb{N}_0$  we have  $[m, n] := \{p \in \mathbb{N}_0 \mid m \leq p \leq n\}$ .

Now we will come to the relationship between  $r$ -ary codes and  $r$ -ary rooted trees.

**(1.4) Definition** ( $r$ -ary Trees from  $r$ -ary Codes).

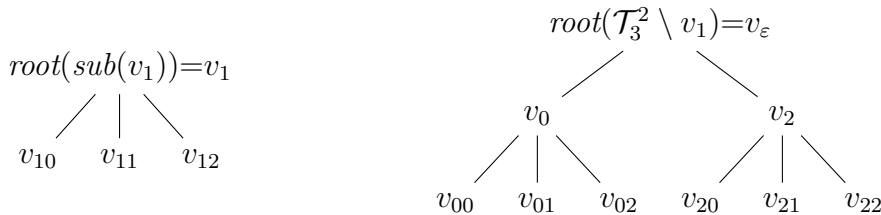
Let  $q, r \in \mathbb{N}$ ,  $A := [0, r - 1]$  be the code-alphabet for some  $r$ -ary code  $\mathcal{C}$  with word-lengths  $l \in \mathbb{N}^q$ . This choice of the code-alphabet can always be made since any other code-alphabet for  $\mathcal{C}$  would stand in bijection to  $A$ . Set  $h := \max\{l_i \mid i \in [1, q]\}$ . Define  $W := \bigcup_{i \in [0, h]} A^i$  to be the set of all words over  $A$  of maximum length  $h$ . Thus  $\mathcal{C} \subseteq W$ . We construct a rooted  $r$ -ary tree  $\mathcal{T}_r^h$  of height  $h$  indexed by  $W$  by setting  $\text{root}(\mathcal{T}_r^h) := v_\varepsilon$ ,  $V(\mathcal{T}_r^h) := \{v_w \mid w \in W\}$  and  $E(\mathcal{T}_r^h) := \{(v_w, v_{w'}) \mid w, w' \in W, wx = w', x \in A\}$ . Note that  $\mathcal{T}_r^h$  is uniquely determined by  $r, h$ .

**(1.5) Examples.**  $\mathcal{T}_3^2$  is given by



We have  $\text{height}(v_\varepsilon) = 0$  and  $\text{height}(v_{12}) = 2$ .  $v_0 \leq v_{02}$  holds,  $v_0 \leq v_{10}$  does **not**.

The subtrees  $\text{sub}(v_1)$  and  $\mathcal{T}_3^2 \setminus v_1$  are given by:



$\text{sub}(v_1)$  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  $\text{height}_{\text{sub}(v_1)}(v_1) = 0$ ,  $\text{height}_{\text{sub}(v_1)}(v_{12}) = 1$ , but still  $\text{height}_{\mathcal{T}_3^2 \setminus v_1}(v) = \text{height}_{\mathcal{T}_3^2}(v)$  for  $v \in V(\mathcal{T}_3^2 \setminus v_1)$ .

**(1.6) Remark.** One can now see the relation between the tree  $\mathcal{T}_r^h$  and its code-words; We have  $v_w \leq v_{w'} \iff w \sqsubseteq w'$  and  $\text{height}(v_w) = |w|$  for  $v_w, v_{w'} \in V(\mathcal{T}_r^h)$ . (Proof Omitted)

**(1.7) 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

$$\sum_{k=1}^q \frac{1}{r^{l_k}} \leq 1 \quad (1)$$

*Proof.* If  $q = 1$ , then we always have an instantaneous code, and since  $r \in \mathbb{N}$ , (1) always holds as well. So assume w.l.o.g. that  $q > 1, \forall i \in [1, q-1] : 0 < l_i \leq l_{i+1}$  and that the code-alphabet of  $\mathcal{C}$  is  $[0, r-1]$ .

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 (1.4), 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  $i$ . The idea is to remove the subtrees rooted at  $v_{w_i}$  and then chose  $v_{w_{i+1}}$  from the remaining vertices to uphold the prefix property of  $\mathcal{C}$ , since for all  $j \in [1, i]$  we already removed all  $v_w$  with  $w_j \sqsubseteq w$  before choosing  $v_{w_{i+1}}$ , so with  $l_j \leq l_{i+1}$  we then know the prefix property still holds.

Let  $i = 1$ . Choose a code-word  $w_1 \in [0, r-1]^{l_1}$  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  $\mathcal{T}_0 := \mathcal{T}_r^h, \mathcal{T}_1 := \mathcal{T}_0 \setminus v_{w_1}$ . We know from (1.3) that  $\mathcal{T}_1$  has

$$r^h - r^{h-\text{height}(v_{w_1})} = r^h - r^{h-l_1} = r^h \left(1 - \sum_{k=1}^1 \frac{1}{r^{l_k}}\right) > r^h \left(1 - \sum_{k=1}^q \frac{1}{r^{l_k}}\right) \stackrel{(1)}{\geq} 0$$

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  $\text{height}(v_w) = l_{i+1}$  since trees are connected and we have a leaf. So set  $w_{i+1} := w$ . If we had  $w_j \sqsubseteq w_{i+1}$  for some  $j \in [1, i]$ , then also  $v_{w_j} \leq v_{w_{i+1}}$ , but then  $v_{w_{i+1}} \notin V(\mathcal{T}_{j-1} \setminus v_{w_j}) = V(\mathcal{T}_j) \supseteq V(\mathcal{T}_i)$ , 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  $\mathcal{T}_{i+1} := \mathcal{T}_i \setminus v_{w_{i+1}}$  and we get for the number of leaves:

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

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

Now we show that the existence of an instantaneous  $r$ -ary code  $\mathcal{C}$  with word-lengths  $l$  implies (1). We know from [JJ00] that  $\mathcal{C}$  is a prefix-code. Let  $i \in [1, q], w_i \in \mathcal{C}, |w_i| = l_i$  and set

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

to be the set of leaves in  $\text{sub}(v_{w_i})$ . We know from (1.3) that  $|L_i| = r^{h-l_i}$  for  $i \in [1, q]$ .

Furthermore we know that for each  $i \neq j \in [1, q]$   $L_i \cap L_j = \emptyset$ :

Assume  $i, j \in [1, q]$  and w.l.o.g.  $i < j$ . Let  $v_w \in L_i \cap L_j$ . Thus we get

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

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 get what we wanted to show:

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

□

One could assume, that because being instantaneous implies being uniquely decodable, the constraints for being the latter are weaker. Suprisingly, this is not the case:

**(1.8) 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

$$\sum_{i=1}^q \frac{1}{r^{l_i}} \leq 1 \quad (1)$$

*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 uniquely decodable  $r$ -ary code with word-lengths  $l$ .

Let  $K := \sum_{i=1}^q \frac{1}{r^{l_i}}$  and  $n \in \mathbb{N}$ . So we want to show  $K \leq 1$ . We have

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

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

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  $i \in [1, q]^n$  (3). We define for  $j \in [mn, Mn]$ :

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

So  $t \in N_j$  is a code-sequence of length  $j$ , consisting of  $n$  code-words in  $\mathcal{C}$ .

But since  $\mathcal{C}$  is uniquely decodable, we know that  $\forall t \in N_j : \exists! i \in [1, q]^n : t = w_{i_1} \cdots w_{i_n}$ , meaning there is one and only one way to construct  $t \in N_j$  from  $n$  code-words of  $\mathcal{C}$ .

This implies that

$$|\{i \in [1, q]^n \mid \sum_{k=1}^n l_{i_k} = j\}| = |\{i \in [1, q]^n \mid \sum_{k=1}^n |w_{i_k}| = j\}| = |N_j| \quad (4)$$

Furthermore, since  $N_j \subseteq [0, r-1]^j$ , we have  $|N_j| \leq r^j$ . Thus, from (2), (3), (4) we get

$$K^n = \sum_{j=mn}^{Mn} \frac{|N_j|}{r^j} \leq \sum_{j=mn}^{Mn} 1 = (M-m)n + 1 \implies \frac{K^n}{n} \leq (M-m) + \frac{1}{n}$$

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

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

□

**(1.9) Corollary.**

Let  $r, q \in \mathbb{N}, l \in \mathbb{N}^q$ . Then by the above inequalities we get that there exists an instantaneous  $r$ -ary code with word-lengths  $l$  iff there exists a uniquely decodable  $r$ -ary code with word-lengths  $l$ .

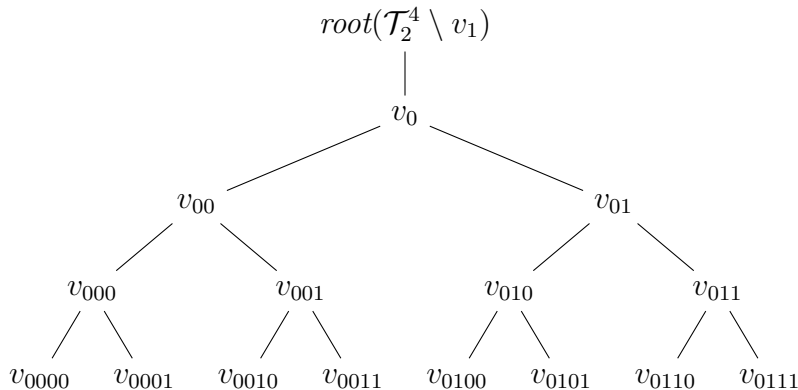
So now we have necessary conditions for when instantaneous  $r$ -ary Codes of some given length exist for some  $r \in \mathbb{N}$ . When searching / constructing a code, one usually wants a it to be instantaneous and have its word lengths and code-alphabet as small as possible. These properties are related through the Inequalities we proved. In particular it is not possible to construct an instantaneous or uniquely decodable  $r$ -ary Code with arbitrarily small word-lengths for some fixed  $r$ , neither for an arbitrarily small  $r$ , given fixed word-lengths;

There exists a lower bound given by these Inequalities.

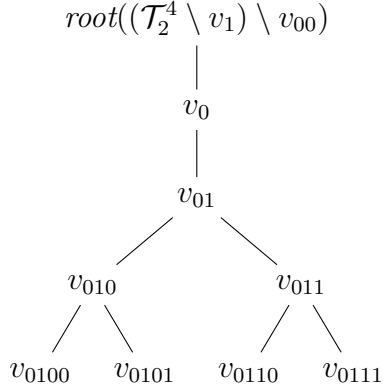
**(1.10) Remark.** Note that we know that if  $q, r \in \mathbb{N}, l \in \mathbb{N}^q$  satisfy Kraft's Inequality, there **exists** an instantaneous  $r$ -ary Code. This does in no way imply that every code with code-words of these lengths is instantaneous. Consider for example  $r = 2, q = 3, l = (1, 2, 3)$ . Then we have  $\sum_{k=1}^q \frac{1}{r^{l_k}} = \frac{7}{8} \leq 1$ , but the 2-ary code  $\{0, 01, 011\}$  is obviously not a prefix-code  $\mathcal{C}$  and thus not instantaneous. Similarly, by (1.9) we know that if we have some uniquely decodable code, there **exists** an instantaneous code with the same word lengths, not that  $\mathcal{C}$  is instantaneous. For this, consider the code  $\{0, 01, 11\}$ , which is uniquely decodable, but not instantaneous since  $0 \sqsubseteq 01$ .

As the proof for Kraft's Inequality is constructive, we conclude with an example of constructing an instantaneous code for given constraints satisfying Kraft's Inequality:

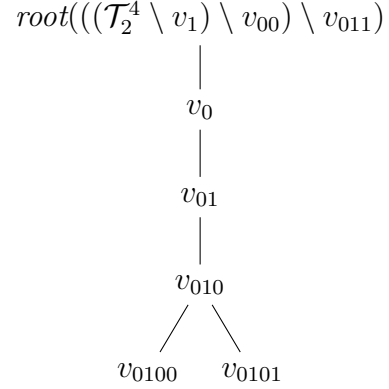
**(1.11) Example.** Let  $r = 2, q = 4, l = (1, 2, 3, 4)$ , which satisfy the Kraft Inequality. We may chose  $w_1 \in [0, r - 1]^{l_1} = [0, 1]$  so set  $w_1 := 1$ . Now consider  $\mathcal{T}_2^{\max l} \setminus v_{w_1} = \mathcal{T}_2^4 \setminus v_1$ :



Now we chose one of the vertices at the height  $l_2 = 2$ , lets say  $v_{00}$ , and thus set  $w_2 := 00$ . Again consider  $(\mathcal{T}_2^4 \setminus v_1) \setminus v_{00}$ :



For  $l_3 = 3$  we chose  $v_{011}$ ,  $w := 011$  and see  $((\mathcal{T}_2^4 \setminus v_1) \setminus v_{00}) \setminus v_{011}$  is given by:



For  $l_4 = 4$  we have 2 choices, the leaves of  $((\mathcal{T}_2^4 \setminus v_1) \setminus v_{00}) \setminus v_{011}$ , left and set  $w_4 := 0100$ . Now we have constructed the  $r$ -ary prefix-code  $\mathcal{C} := \{1, 00, 011, 0100\}$  with word-lengths  $l$ , which we know is instantaneous.

## References

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