Chapter 7

Kraft's Inequality and its Converse

KRAFT'S INEQUALITY constrains entropy to increase at a certain rate. Its converse—sometimes known as Chaitin's lemma—shows that we can construct machines approaching arbitrarily close to this constraint.

7.1 Kraft's inequality

Recall that, for any Turing machine T, the set of strings

$$S = \{p \in \mathbb{S} : T(p) \text{ defined}\}$$

is *prefix-free*.

Theorem 7.1. *(Kraft's Inequality) If $S \subset \mathbb{S}$ is prefix-free then*

$$\sum_{s \in S} 2^{-|s|} \leq 1.$$

Proof ▶. To each string $s = b_1 b_2 \dots b_n$ we associate the binary number

$$B(s) = 0 \cdot b_1 b_2 \dots b_n \in [0, 1),$$

and the half-open interval

$$I(s) = [B(s), B(s) + 2^{-|s|}) \subset [0, 1).$$

Lemma 1. *The real numbers $B(s),\ s \in \mathbb{S}$ are dense in $[0, 1)$.*

Proof ▶. If

$$x = 0.b_1 b_2 \cdots \in [0, 1)$$

then

$$B(b_1),\ B(b_1 b_2),\ B(b_1 b_2 b_3) \to x.$$

Recall that we write $s \prec s'$ to mean that s is a prefix of s', eg

$$01101 \prec 0110110.$$

Lemma 2. *For any two strings $s, s' \in \mathbb{S}$,*

1. $B(s') \in I(s) \iff s \prec s'$;
2. $I(s') \subset I(s) \iff s \prec s'$;
3. $I(s), I(s')$ *are disjoint unless* $s \prec s'$ *or* $s' \prec s$

Proof ▶. 1. Let

$$s = b_1 \dots b_n.$$

Suppose $s \prec s'$, say

$$s' = b_1 \dots b_n b_{n+1} \dots b_{n+r}.$$

Then

$$B(s) \le B(s') = B(s) + 2^{-n} 0.b_{n+1} \dots b_{n+r} < B(s) + 2^{-n} = B(s) + 2^{-|s|}.$$

Conversely, supppose $s \not\prec s'$. Then either $s' \prec s$ (but $s' \neq s$); or else s, s' differ at some point, say

$$s = b_1 \dots b_{r-1} b_r b_{r+1} \dots b_n,\ s' = b_1 \dots b_{r-1} c_r c_{r+1} \dots c_m,$$

where $b_r \neq c_r$.

If $s' \prec s$ or $b_r = 1,\ c_r = 0$ then $B(s') < B(s)$.

If $b_r = 0,\ c_r = 1$ then

$$B(s') \ge 0.b_1 \dots b_{r-1} 1 > B(s) = 0.b_1 \dots b_{r-1} 0 b_{r+1} \dots b_n /$$

Thus

$$B(s) = \frac{a}{2^n},\ B(s') = \frac{b}{2^n},$$

with $a < b$. Hence

$$B(s') \ge B(s) + \frac{1}{2^n}.$$

2. Suppose $s \prec s'$. Then

$$B(s'') \in I(s') \Longrightarrow s' \prec s'' \Longrightarrow s \prec s'' \Longrightarrow B(s'') \in I(s).$$

It follows that

$$I(s') \subset I(s).$$

Conversely,

$$I(s') \subset I(s) \Longrightarrow B(s') \in I(s) \Longrightarrow s \prec s'.$$

3. If $I(s),\ I(s')$ are disjoint then we can find $s'' \in \mathbb{S}$ such that

$$B(s'') \in I(s) \cap I(s'),$$

so that

$$s \prec s'',\ s' \prec s''$$

which implies that

$$s \prec s' \text{ or } s' \prec s.$$

Conversely,

$$s \prec s' \Longrightarrow I(s') \subset I(s),\ s' \prec s \Longrightarrow I(s) \subset I(s');$$

and in neither case are $I(s),\ I(s')$ disjoint.

It follows from the last result that if the set of strings $S \subset \mathbb{S}$ is prefix-free then the half-intervals

$$\{I(s) : s \in S\}$$

are disjoint; and so, since they are all contained in $[0, 1)$,

$$\sum_{s \in S} |I(s)| = \sum_{s \in S} 2^{-|s|} \leq 1.$$

7.1.1 Consequences of Kraft's inequality

Proposition 7.1. *For each Turing machine T,*

$$\sum_{s \in \mathbb{S}} 2^{-H_T(s)} \leq 1.$$

Proof ▶. We know that

$$\sum_{p:T(p) \text{ is defined}} 2^{-|p|} \leq 1.$$

But each s for which $T(s)$ is defined arises from a unique minimal input

$$p = \mu_T(s);$$

while if $T(s)$ is not defined that

$$H_T(s) = \infty \Longrightarrow 2^{-H_T(s)} = 0.$$

It follows that the entropy of strings must increase sufficiently fast to ensure that

$$\sum_{s \in \mathbb{S}} 2^{-H(s)} \leq 1.$$

Thus there cannot be more than 2 strings of entropy 2, or more than 16 strings of entropy 4; if there is one string of entropy 2 there cannot be more than 2 of entropy 3; and so on.

7.2 The converse of Kraft's inequality

Theorem 7.2. *Suppose $\{h_i\}$ is a set of integers such that*

$$\sum 2^{-h_i} \leq 1.$$

Then we can find a prefix-free set $\{p_i\} \subset \mathbb{S}$ of strings such that

$$|p_i| = h_i.$$

Moreover this can be achieved by the following strategy: The strings $p_0, p_1, \ldots$ are chosen successively, taking p_i to be the first string (in lexicographical order) of length h_i such that the set

$$\{p_0, p_1, \ldots, p_i\}$$

is prefix-free.

Recall that the lexicographical order of $\mathbb{S}$ is

$$\square < 0 < 1 < 00 < 01 < 10 < 11 < 000 < \cdots,$$

where $\square$ denotes the empty string.

Proof ▶. Suppose the strings $p_0, p_1, \ldots, p_{i-1}$ have been chosen in accordance with the above specification. The remaining space (the 'gaps' in $[0,1)$)

$$G = [0,1) \setminus (I(p_0) \cup I(p_1) \cup \cdots \cup I(p_{i-1}))$$

is expressible as a finite union of disjoint half-open intervals $I(s)$, say

$$C = I(s_0) \cup I(s_1) \cup \cdots \cup I(s_j)$$

where

$$B(s_0) < B(s_1) < \cdots < B(s_j).$$

(This expression is unique if we agree to amalgamate any adjoining 'twin' intervals of the form

$$B(b_1, \ldots, b_r, 0),\ B(b_1, \ldots, b_r, 1)$$

to form the single interval

$$B(b_1, \ldots, b_r)$$

of twice the length.)

Lemma 3. *The intervals $I(s_0), \ldots, I(s_j)$ are strictly increasing in length, ie*

$$|s_0| > |s_1| > \cdots > |s_j|;$$

and

$$h_i \leq |s_j|,$$

so that it is possible to add another string p_i of length h_i.

Proof ▶. We prove the result by induction on i. Suppose it is true for the prefix-free set $\{p_0, \ldots, p_{i-1}\}$.

Since the intervals $I(s_k)$ are strictly increasing in size, each $I(s_k)$ is at most half as large as its successor $I(s_{k+1})$:

$$|I(s_k)| \leq \frac{1}{2}|I(s_{k+1}|.$$

It follows that the total space remaining is

$$< |I(s_j)| \left(1 + \frac{1}{2} + \frac{1}{4} + \cdots\right) = 2|I(s_j)|.$$

The next interval we are to add is to have length h_i. By hypothesis

$$2^{-h_0} + \cdots + 2^{h_{i-1}} + 2^{h_i} \leq 1.$$

Thus

$$\begin{aligned} 2^{-h_i} &\leq 1 - 2^{h_0} - \cdots - 2^{h_{i-1}} \\ &= |[0,1) \setminus (I(p_0) \cup I(p_1) \cup \cdots \cup I(p_{i-1}))| \\ &= |I(s_0) \cup I(s_1) \cup \cdots \cup I(s_j)| \\ &< 2|I(s_j)|. \end{aligned}$$

It follows that

$$2^{-h_i} \leq |I(s_j)|,$$

or

$$h_i \geq |s_j|.$$

So we can certainly fit an interval $I(p)$ of length 2^{-h_i} into one of our 'gap' intervals $I(s_k)$.

By prescription, we must take the 'first' position available for this new interval. Let us determine where 2^{-h_i} first fits into the sequence of strictly increasing gaps $I(s_0), I(s_1), \ldots$. Suppose

$$|I(s_{k-1})| < 2^{-h_i} \leq |I(s_k)|.$$

Then $I(s_k)$ is the first 'gap' into which we can fit an interval $I(p)$ of length 2^{-h_i}.

If in fact

$$2^{-h_i} = |I(s_k)|$$

then we set

$$p_i = s_k.$$

In this case, the gap is completely filled, and we continue with one fewer gap, the remaining gaps evidently satisfying the conditions of the lemma.

If however

$$2^{-h_i} < |I(s_k)|$$

then our strategy prescribes that $I(p_i)$ is to come at the 'beginning' of $I(s_k)$, ie

$$p_i = s_k \overbrace{0 \ldots 0}^{e \text{ 0's}},$$

where

$$e = h_i - |s_k|.$$

We note that

$$I(s_k) \setminus I(p_i) = I(t_0) \cup I(t_1) \cup \cdots \cup I(t_{e-1}),$$

where

$$t_0 = s_k \overbrace{0 \ldots 0}^{e-1 \text{ 0's}} 1,\ t_1 = s_k \overbrace{0 \ldots 0}^{e-2 \text{ 0's}} 1,\ \ldots, t_{e-2} = s_k 01,\ t_{e-1} = s_k 1.$$

Thus after the addition of the new interval $I(p_k)$ the complement

$$[0,1) \setminus (I(p_0) \cup \cdots \cup I(p_i)) = \\ I(s_0) \cup \cdots \cup I(s_{k-1}) \cup I(t_0) \cup \cdots I(t_r) \cup I(s_{k+1}) \cup \cdots \cup I(s_j)$$

retains the property described in the lemma. It therefore follows by induction that this property always holds.

It follows that the strategy can be continued indefinitely, creating a prefix-free set of strings with the required properties.

7.3 Chaitin's lemma

We would like to construct a machine T so that specified strings $s_0, s_1, \ldots$ have specified entropies $h_0, h_1, \ldots$:

$$H_T(s_i) = h_i.$$

By Kraft's Inequality this is certainly not possible unless

$$\sum_i 2^{-h_i} \leq 1.$$

But suppose that is so. The converse to Kraft's inequality encourages us to believe that we should be able to construct such a machine.

But one question remains. What exactly do we mean by saying that the entropies h_i are 'specified'? *How* are they specified?

If the machine T is to 'understand' the specification, it must be in 'machine-readable' form. In other words, we must have *another* machine M outputting the numbers h_i.

Theorem 7.3. *Suppose*

$$S \subset \mathbb{S} \times \mathbb{N}$$

is a set of pairs (s, h_s) *such that*

1. *The integers h_s satisfy Kraft's condition:*

$$\sum_{(s,h_s)\in S} 2^{-h_s} \leq 1.$$

2. *The set S is recursively enumerable.*

Then there exists a Turing machine T such that

$$H_T(s) \leq h_s$$

for all $(s, h_s) \in S$.

Proof ▶. By definition, there exists a machine M which generates the set S, say

$$M(n) = (s_n, h_n) \in S.$$

Suppose we are given an input string p. We have to determine $T(p)$. Our machine does this by successively building up a prefix-free set

$$P = \{p_0, p_1, p_2, \dots\},$$

where $|p_n| = h_n$, according to the prescription above. As each p_i is created, it is compared with the given string p; and if $p_n = p$ then T outputs the string s_n and halts.

If p never occurs in the prefix-free set P then $T(p)$ is undefined.

More fully, T functions in stages $0, 1, 2, \dots$. At stage n, T emulating each of $M(0), M(1), \dots, M(n)$ for n steps.

If $M(r)$ halts after $m \leq N$ steps, with

$$M(r) = \langle s_r \rangle \langle h_r \rangle.$$

Then T adds a further string p_i with $|p_i| = h_r$ to the prefix-free set

$$P = \{p_0, p_1, \dots, p_{i-1}\}$$

which it is building up, by following the 'Kraft prescription'.

Summary

We have constructed a machine T with specified entropies $H_T(s_i)$ for specified the string s_i, provided these entropies satisfy Kraft's inequality, and can be recursively generated.