15-455: UCT	**K. Sutner**
Assignment 9	Due: **April 4, 2024, 24:00**.

1. Kraft's Inequality (20)

Background

Kraft's inequality is a combinatorial lemma that is often used in the theory of prefix codes; it is also crucial for the construction of Chaitin's Ω.

Kraft's Lemma:
Let $S \subseteq \mathbf{2}^\star$ be a prefix set of binary words. Then

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

On the other hand, call a non-decreasing sequence of natural numbers $\ell = (\ell_k)_{k<N}$, where $N \in \mathbb{N} \cup \{\omega\}$, admissible if

$$\sum_{k<N} 2^{-\ell_k} \leq 1$$

Then for every admissible sequence ℓ there exists a prefix set $S = \{ s_k \mid k < N \}$ of cardinality N such that $\ell_k = |s_k|$, in which case S is said to realize ℓ. For example, $\ell_k = k + 1$ is admissible and can be realized by $s_k = 0^k 1$.

Task

A. Prove the lemma for any finite prefix set S.

B. Find an algorithm that constructs a realizing set S from a finite admissible list ℓ.
Prove your algorithm correct and analyze its running time.

C. Conclude that the lemma holds for arbitrary prefix sets.

D. Characterize the finite prefix sets for which $\sum_{x \in S} 2^{-|x|} = 1$.

Extra Credit: Extend the characterization to infinite sets.

2. Prefix Encoding (30)

Background

Recall our project of finding a good prefix encoding that works for any binary string. In the following, by an encoding or code we mean any injective function $\mathbf{2}^\star \to \mathbf{2}^\star$. An encoding f is prefix if $f(\mathbf{2}^\star)$ is prefix. To avoid pesky edge cases, we will simply ignore words of length 0 or 1 (extra credit: figure out how to fix this).

$$\begin{aligned}
E(x_1 \dots x_n) &= x_1 0\, x_2 0\, \dots\, x_{n-1} 0\, x_n 1 \\
E_0(x) &= E(x) \\
E_{i+1}(x) &= E_i(\mathsf{blen}\, x)\, x \\
\widehat{E}(x) &= E_k(x) & k = \mathsf{blen}^\star\, x \\
E_\infty(x) &= E(k)\, \widehat{E}(x) & k = \mathsf{blen}^\star\, x
\end{aligned}$$

Using a slightly different approach, here is another attempt at a prefix code

$$G(x) = \mathsf{blen}^k(x)\, 0\; \mathsf{blen}^{k-1}(x)\, 0\, \dots\, |x|\, 0\, x\, 1$$

For example, G turns any string x of length 20000 into

$$G(x) = 11\ \underline{0}\ 100\ \underline{0}\ 1111\ \underline{0}\ 100111000100000\ \underline{0}\ x\ \underline{1}$$

where the extra spaces and underlining are added for visually clarity, they are missing in the actual code.

Task

A. Show that E_k is a prefix encoding for all fixed $k \geq 0$.

B. Show that $\widehat{E}$ is an encoding but not prefix.

C. Show that E_∞ is a prefix encoding.

D. Show that G is a prefix encoding.

3. Kolmogorov versus Primes (30)

Background

One can abuse Kolmogorov-Chaitin complexity to show that there are infinitely many primes, though many would argue that the original argument is far superior. But, with a little bit of extra effort, one can push this argument to get a fairly good estimate for the density of primes (which results are important for algorithms trying to produce large primes). Write $\pi(n)$ for the number of primes up to n. The celebrated and difficult prime number theorem (PNT) says that $\pi(n) \approx n/\log n$. We will settle for a weaker claim: $\pi(n) \geq cn/\log^2 n$

Write $p_1, p_2, \ldots$ for the sequence of primes, so that for any number n we have a unique decomposition $n = \prod_{i \leq m} p_i^{e_i}$ where $0 \leq e_i$ and $e_m \neq 0$.

Task

A. Use Kolmogorov-Chaitin complexity to show that there are infinitely many primes.

B. Use Kolmogorov-Chaitin complexity to prove $\pi(n) \geq cn/\log^2 n$, for some constant c and infinitely many n.

Comment Use the fact that a number n can be decomposed into its largest prime factor p and n/p; our prefix coding functions also come in handy. For the second part, it is easier to use prefix complexity, but the argument does not depend on it.