Student Number: [redacted] Name: Bryan Hoang

5. (20 points)

(a) **Answer:**

The LZ78 parsing of the sequence $x^n = 11111111\dots1$ is

$$\boxed{1, 11, 111, 1111, \dots, \underbrace{1\dots1}_{l_n}},$$

where the length of the last phrase, l_n, is

$$l_n = \begin{cases} T_n & \text{if } n \text{ is a triangular number,} \\ \lfloor T_n \rfloor & \text{otherwise,} \end{cases}$$

and where

$$T_n = \frac{\sqrt{8n+1}-1}{2},$$

is the triangular root of n.

(b) **Answer:**

Proof. We know that

$$l(x^n) = c(x^n)(\log c(x^n) + O(1)),$$

where $c(x^n)$ denotes the number of phrases in the dictionary obtained by parsing x^n. In this case, $c(x^n) = T_n$ from part (a). We also know that

$$T_n = \frac{\sqrt{8n+1}-1}{2} \leq \sqrt{2n}. \tag{1}$$

Therefore,

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n} l(x^n) &= \lim_{n\to\infty} \frac{1}{n} T_n(\log T_n + O(1)) \\ &\leq \lim_{n\to\infty} \frac{1}{n} \sqrt{2n}(\log \sqrt{2n} + O(1)) && \text{by (1)} \\ &= \lim_{n\to\infty} \frac{\sqrt{2}(\log \sqrt{2n} + O(1))}{\sqrt{n}} \\ &= 0. \end{aligned}$$

□