COMS 331: Theory of Computing, Spring 2023
Homework Assignment 10

Neha Maddali

Due at 10:00PM, Wednesday, April 26, on Gradescope.

Problem 64.
Let $q, r \in Q$ be rational numbers with $0 \leq q \leq r \leq 1$. Let M_1 be a TM that on input $\langle s_k, \pi \rangle$ outputs $x0^k$. Then let M_2 be a TM that on input $\langle s_k, \pi' \rangle$ outputs z such that $z = U(\pi)$ up to $|U(\pi)| - k$ bits. Then, $U(\pi') = x0^k$ implies that $M_2(\langle s_k, \pi' \rangle) = x$. Using the Theorem from lecture $C(x) \leq C_M(x) + c_M$, for M_1 we have $C(x0^k) \leq C_{M_1}(x0^k) + c_{M_1}$, which is $C(x0^k) \leq |\langle s_k, \pi \rangle| + c_{M_1}$, which is $C(x0^k) \leq 2log(n+1) + C(x) + c_{M_1}$ where C(x) comes from $U(\pi) = x$. Likewise, for M_2, we have $C(x) \leq C_{M_2}(x) + c_{M_2}$, which is $C(x) \leq |\langle s_k, \pi' \rangle| + c_{M_2}$, which is $C(x) \leq 2log(n+1) + C(x0^k) + c_{M_2} + 1$, where $C(x0^k)$ comes from $U(\pi') = x0^k$. So $|C(x0^k) - C(x)| \leq 2log(n+1) + c$. Now, we define $J = \{n | 2(2log(n+1) + c + a) < (r-q)n\}$ where some a is from the statement $C(x) \leq |x| + a$. Then, from Corollary 6 in lecture, for all n there exists some $y \in \{0,1\}^n$ such that $C(y) \geq |y|$. Now, let $x = y0^k$ where $|x| = n$ (where $n \in J$), $C(y) \geq |y|$ and $|y| = m$. Then, there exists some m such that $qn < m - (2log(n+1) + c + a) < m + (2log(n+1) + c + a) < rn$. Now we have $|C(y0^k) - C(y)| \leq 2log(n+1) + c$, which is $|C(x) - C(y)| \leq 2log(n+1) + c$. Then, we have $C(x) \geq C(y) - (2log(n+1) + c)$ which is $C(x) \leq C(y) + (2log(n+1) + c)$, which after substitutions gives $C(x) < m + 2log(n+1) + c + a$. Then also, $C(x) > m - 2log(n+1) + c + a$. By showing C(x) is greater than and less than those expressions, it follows from a previous statement in the proof that $qn < C(x) < rn$. ■

Problem 65.
Let (f,g) be a lossless data compression scheme. Let us assume that $|f(x)| < |x|$ for all $x \in \{0,1\}^n$. The number of strings with length less than n is $\sum_{i=1}^{n-1} 2^i = 2^n - 1$. The number of strings $x \in \{0,1\}^n = 2^n$. For f to be lossless it must be one to one, but we can see that there are less f(x) than x, so f by definition cannot be one to one. Therefore, since f is not one to one, f is not lossless. This is a contradiction, so there must be a string $x \in \{0,1\}^n$ such that $|f(x)| \geq |x|$. ■

Problem 66.
Let (f,g) be a lossless data compression scheme. From a theorem discussed in lecture, we have $C(x) \leq C_M(x) + c_M$. Then we can construct a TM M that on input f(x) outputs g(f(x)). This is possible because f and g are computable functions. Then $C_M(x) = |f(x)|$ so we have $C(x) \leq |f(x)| + c_M$. Let $c_{(f,g)} = c_M$, and we have $C(x) \leq |f(x)| + c_{(f,g)}$. ■

Problem 67.
Let (f,g) be a lossless data compression scheme. For the sake of contradiction, assume that there only exist finitely many x such that $C(x) < |f(x)|$. So there are x where $C(x) \geq |f(x)|$. Then there must be some n such that $|x| = n$ and then all $x \in \{0,1\}^n$ have $C(x) \geq |f(x)|$. Then let f'(x) be 0 if

$|x| = n$ and $|f(x)|$ otherwise. Then we have $f'(x) \leq C(x)$. By theorem 11 in lecture, if f' is a lower bound of C, then f' is bounded. This implies that there exists some $m \in N$ such that $f'(x) \leq m$ for all $x \in \{0,1\}^*$. Then let $a = max\{n, m+1\}$. For all $x \in \{0,1\}^a, |f(x)| = f'(x) \leq m < |x|$, which implies that $|f(x)| < |x|$. This is a contradiction to problem 65, and if $|f(x)| < |x|$ were true, then f is not lossless. Since f is lossless, there exist infinitely many strings $x \in \{0,1\}^*$ such that $C(x) < |f(x)|$. ■

Problem 68.
We know that $diam(G) = max\{d_G(s,t)|s,t \in V\}$. When constructing the TM we will use the idea that we have points s, t, v_i where v_i is a point between s and t. Then for $v \in \{00, 01, 10\}^n$ the first bit represents an edge between s and v_i and the second bit represents an edge between v_i and t. A 1 means there is an edge, and a 0 means there is no edge. Then, we can construct a TM M to hav input $\langle s_n, s_s, s_t, x_H, \pi_v \rangle$ where $n = |V|$, s and t are points, x_H is the sub graph without points s or t and $U(\pi_v = v$. With this input, we can output the graph x_G. From a theorem discussed in lecture, $C(x_G) \leq C_M(x_G) + c_M$, which is $C(x_G) \leq |\langle s_n, s_s, s_t, x_H, \pi_v \rangle| + c_M$ which by definition of their encodings gives $C(x_G) \leq |0^{|s_n|}10^{|s_s|}10^{|s_t|}1s_n s_s s_t x_H \pi_v| + c_M$. This is equivalent to $C(x_G) \leq 2|s_n| + 2|s_s| + 2|s_t| + 3 + |x_H| + |\pi_v| + c_M$. Replacing the lengths with their values we have $C(x_G) \leq 6log(n+1) + 3 + \binom{n-2}{2} + n(log3) + 2log(n+1) + c_b + c_M$, and when we rearrange the terms and let $c = c_b + c_M$ we get $C(x_G) \leq \binom{n-2}{2} + 3 + n(log3) + 8log(n+1) + c, where \binom{n-2}{2} + 3 + n(log3) + 8log(n+1) + c < \binom{n}{2} - (2 - log3)n + 8log(n+1) + c_c$, so it follows that $C(x_G) \leq \binom{n}{2} - (2 - log3)n + 8log(n+1) + c$. ■

Problem 69.
If $diam(G) = 1$, then $C(x_G) \leq log(n+1) + c_a$. From problem 68, we know that if $diam(G) > 2$, then $C(x_G) \leq \binom{n-2}{2} - (2 - log3)n + 8log(n+1) + c_c$. We can see that $\binom{n}{2} - \frac{2}{5}n > log(n+1) + c_a$ because $log(n+1) + c_a$ will converge while $\binom{n}{2} - \frac{2}{5}n$ will grow at a faster rate, so we can conclude that $diam(G) \neq 1$ and $diam(G) \geq 2$. Then we see that $\binom{n}{2} - \frac{2}{5}n > \binom{n}{2} - (2 - log3)n + 8log(n+1) + c_c$ where we can cancel out the $\binom{n}{2}$ and rearrange the terms to get $(2 - log3)n - \frac{2}{5}n > 8log(n+1) + c_c$, which by doing some math gives $\frac{n}{100} > 8log(n+1) + c_c$. This implies that $diam(G) \leq 2$. Since $diam(G) \geq 2$ and $diam(G) \leq 2$, we can conclude that $diam(G) = 2$. ■

Problem 70.
Theorem 5 from lecture states that Prob$[C(x) \geq n - r]$> $1 - 2^{-r}$. From problem 69, we know $diam(G) = 2$ which implies that $C(x_G) \geq \binom{n}{2} - \frac{2}{5}n$ so we have Prob[diam(2)]=Prob$[C(x_G) \geq \binom{n}{2} - \frac{2}{2}n]$> $1 - 2^{-\frac{2}{5}n}$ where $n = \binom{n}{2}$ and $r = \frac{2}{5}n$. ■