HW 1: Extra problems

Instructors: Har-Peled

CS/ECE 374A: Intro. Algorithms & Models of Computation, Fall 2022 Version: **1.0**

Each homework assignment will include at least one solved problem, similar to the problems assigned in that homework, together with the grading rubric we would apply *if* this problem appeared on a homework or exam. These model solutions illustrate our recommendations for structure, presentation, and level of detail in your homework solutions. Of course, the actual *content* of your solutions won't match the model solutions, because your problems are different!

Solved Problems

1 Recall that the *reversal* w^R of a string w is defined recursively as follows:

$$w^R := \begin{cases} \epsilon & \text{if } w = \epsilon \\ x^R \bullet a & \text{if } w = a \cdot x \end{cases}$$

A *palindrome* is any string that is equal to its reversal, like *AMANAPLANACANALPANAMA*, *RACECAR*, *POOP*, *I*, and the empty string.

1. Give a recursive definition of a palindrome over the alphabet Σ.
2. Prove $w = w^R$ for every palindrome w (according to your recursive definition).
3. Prove that every string w such that $w = w^R$ is a palindrome (according to your recursive definition).

In parts (b) and (c), you may assume without proof that $(x \cdot y)^R = y^R \bullet x^R$ and $(x^R)^R = x$ for all strings x and y.

Rubric:[induction] For problems worth 10 points:

+ 1 for explicitly considering an *arbitrary* object
+ 2 for a valid **strong** induction hypothesis
  - **Deadly Sin!** Automatic zero for stating a weak induction hypothesis, unless the rest of the proof is *perfect*.
+ 2 for explicit exhaustive case analysis
  - No credit here if the case analysis omits an infinite number of objects. (For example: all odd-length palindromes.)
  - −1 if the case analysis omits an finite number of objects. (For example: the empty string.)
  - −1 for making the reader infer the case conditions. Spell them out!
  - No penalty if cases overlap (for example:
+ 1 for cases that do not invoke the inductive hypothesis ("base cases")
  - No credit here if one or more "base cases" are missing.
+ 2 for correctly applying the *stated* inductive hypothesis
  - No credit here for applying a *different* inductive hypothesis, even if that different inductive hypothesis would be valid.
+ 2 for other details in cases that invoke the inductive hypothesis ("inductive cases")

– No credit here if one or more “inductive cases” are missing.

2 (100 PTS.) Repeat that.

2.A. Let $x_1, \ldots, x_n$ be a sequence of integer numbers, such that $\alpha \leq x_i \leq \beta$, for all i, where α, β are some integer numbers. Prove that there are at least $\lceil n/(\beta - \alpha + 1) \rceil$ numbers in this sequence that are all equal.

2.B. Let $\mathsf{G} = (\mathsf{V}, \mathsf{E})$ be an undirected graph. Unless we say otherwise, a graph has no loops or parallel edges. Prove, using (A), that if $|\mathsf{V}| \geq 2$ there are two distinct nodes u and v such that degree of u is equal to degree of v. Recall that the degree of a node x is the number of edges incident to x.

2.C. Prove that if all vertices in G are of degree at least one, then there is a (simple) path between two distinct nodes u and v such that degree of u is equal to degree of v.

3 (100 PTS.) Mix this.

The *sort*, w^s, of a string $w \in \{0, 1\}^*$ is obtained from w by sorting its characters. For example, $010101^s = 000111$. The sort function is formally defined as follows:

$$w^s := \begin{cases} \epsilon & \text{if } w = \epsilon \\ 0x^s & \text{if } w = 0x \\ x^s 1 & \text{if } w = 1x \end{cases}$$

The ***merge*** function, evaluated in order from top to bottom, is

$$m(x, y) := \begin{cases} y & \text{if } x = \varepsilon \\ x & \text{if } y = \varepsilon \\ 0m(x', y) & \text{if } x = 0x' \\ 0m(x, y') & \text{if } y = 0y' \\ 1m(x', y) & \text{if } x = 1x' \end{cases}$$

For example, we have $m(10, 10) = 1010$, $m(10, 010) = 01010$, and $m(010, 0001100) = 0000101100$.

For a string $x \in \{0, 1\}^*$, let $\#_0(x)$ and $\#_1(y)$ be the number of 0s and 1s in x, respectively. For example, $\#_0(0101010) = 4$ and $\#_1(0101010) = 3$.

3.A. (**Not for submission.**) Prove by induction that for any string $w \in \{0, 1\}^*$ we have that $w^s \in 0^*1^*$.

3.B. Prove by induction that for any string $w \in \{0, 1\}^*$ we have that $\#_0(w) = \#_0(w^s)$. Conclude that $\#_1(w) = \#_1(w^s)$ and $|w| = \left|w^s\right|$, for any string w.

3.C. Prove by induction that for any two strings $x, y \in \{0, 1\}^*$ we have that

$$\#_0(m(x, y)) = \#_0(x) + \#_0(y).$$

(**Hint:** Do induction on $|x| + |y|$.)
Conclude that $\#_1(m(x, y)) = \#_1(x) + \#_1(y)$. and $\left|m(x, y)\right| = \left|x\right| + \left|y\right|$.
[This part is somewhat tedious if you write carefully all the details out explicitly. Avoid repetition by stating that you are (essentially) repeating an argument that was already seen in the proof.]

3.D. Prove by induction that for any two strings x, y of the form 0^*1^*, we have that $m(x, y)$ is of the form 0^*1^*.

3.E. Prove (using the above) that $(x \bullet y)^s = m(x^s, y^s)$ for all strings $x, y \in \{0, 1\}^*$.

4 (100 PTS.) OLD Homework problem (not for submission):

Walk on the grid.

Let $p_0 = (x_0, y_0)$ be a point on the positive integer grid (i.e., x_0, y_0 are positive integer numbers). A point (x, y) is ***good*** if $x = y$ or $x = 0$ or $y = 0$. For a point $p = (x, y)$ its ***successor*** is defined to be

$$\alpha(p) = \begin{cases} (x, y - x - 1) & y > x \qquad \text{(vertical move)} \\ (x - y - 1, y) & x > y \qquad \text{(horizontal move).} \end{cases}$$

Consider the following sequence $W(p_0) = p_0, p_1, \ldots$ computed for p_0. In the ith stage of computing the sequence, if p_{i-1} is good then the sequence is done as we arrived to a good location. Otherwise, we set $p_i = \alpha(p_{i-1})$.

4.A. Prove, by induction, that starting with any point p on the positive integer grid, the sequence $W(p)$ is finite (i.e., the algorithm performs a finite number of steps before stopping).

4.B. (Harder.) Given such a sequence, every step between two consecutive points is either a vertical or a horizontal move. A ***run*** is a maximal sequence of steps in the walk that are the same (all vertical or all horizontal). Prove that starting with a point $p = (x, y)$ there are at most $O(\log x + \log y)$ runs in the sequence $W(p)$.

5 (100 PTS.) Balanced or not.

Let $\Sigma = \{a, b\}$. Consider a string $s \in \Sigma^*$ of length n. The ***depth*** of a string s is $d(s) = \#_a(s) - \#_b(s)$, where $\#_c(s)$ is the number of times the character c appears in s. The maximum depth of a string s is $d_{\max}(s) = \max_{p \text{ is any prefix of } s} d(p)$.

A string $t \in \{a, b\}^*$ is ***weakly balanced*** if $d(t) = 0$. The string t is ***balanced*** if it is weakly balanced, and for any prefix substring p of t, we have that $\#_a(p) \geq \#_b(p)$.

In the following, you can assume that $\forall x, y \in \Sigma^*$, we have $d(xy) = d(x) + d(y)$.

5.A. (20 PTS.) Let $s = s_1 s_2 \ldots s_n$ be the given string. For any i, let $s_{\leq i}$ be the prefix of s formed by the first i characters of s, where $0 \leq i \leq n$. For any i, let $f(i) = d(s_{\leq i})$. Prove that if there are indices i and j, such that $i < j$ and $f(i) = f(j)$, then $s_{i+1} s_{i+2} \ldots s_j$ is a weakly balanced string.

5.B. (40 PTS.) Prove (but not by induction please) that if a string $s \in \Sigma^*$ is balanced, then either:

(i) $s = \epsilon$,

(ii) $s = xy$ where x and y are non-empty balanced strings, or

(iii) $s = axb$, where x is a balanced string.

5.C. (40 PTS.) Prove that for any string $s \in \{a, b\}^*$ of length n, that is balanced, at least one of the following must happen:

(i) The maximum depth of s is $\geq \sqrt{n}$, or

(ii) s can be broken into m non-empty substrings $s = t_1|t_2|\cdots|t_m$, such that $t_2, t_3, \ldots t_{m-1}$ are weakly balanced strings, and $m \geq \sqrt{n} - 1$.

For example, the string *abaababaabaabbbaaaabbbbb* can be broken into substrings

$$a|ba|ab|ab|aabaabbb|aaaabbbb|b$$

Hint: Let $f(i) = d(s_{\leq i})$. Analyze the sequence $f(0), f(1), \ldots, f(n)$, and what happens if the same value repeats in this sequence many times.

5.D. (Harder + not for submission.) Prove that for any string $s \in \Sigma^*$ of length n, that is balanced, with maximum depth $< \sqrt{n}/2$, it must be that s can be broken into $2m+1$ substrings as follows $s = t_1 t_2 t_3 \ldots t_{2m+1}$, such that the m substrings $t_2, t_4, t_6, \ldots t_{2m}$ are non-empty and balanced. Here m has to be at least $\sqrt{n}/2 - 2$.

6 (100 PTS.) How the first mega tribe was created.

According to an old African myth, in the beginning there were only $n > 1$ persons in the world, and each person formed their own tribe. There were all living in the same forest. Every once in a while two tribes would meet. These meeting tribes would always fight each other to decide which tribe is better, and after a short war, invariably, the tribe with the fewer people (that always lost) would merge into the bigger tribe (if the two tribes were of equal size, one of the tribes would be the losing side). Every person in the tribe that just lost, had to sacrifice a lamb to the forest god, for reasons that remain mysterious, as the lambs did nothing wrong. In the end, only one tribe remained.

Prove, that during this process, at most $n \log_2 n$ lambs got sacrificed. (You can safely assume that no new people were born during this period.)

7 (100 PTS.) A few recurrences.

7.A. Consider the recurrence

$$T(n) = 2n + T(\lfloor n/4 \rfloor) + T(\lfloor (3/4)n \rfloor),$$

where $T(n) = 1$ if $n < 10$. Prove by induction that $T(n) = O(n \log n)$.

7.B. Consider the recurrence

$$T(n) = \begin{cases} T(\lfloor n/2 \rfloor) + T(\lfloor n/6 \rfloor) + T(\lfloor n/7 \rfloor) + n & n \geq 24 \\ 1 & n < 24. \end{cases}$$

Prove by induction that $T(n) = O(n)$.

(An easier proof follows from using the techniques described in section 3 of these notes on recurrences.)