Problem 1: Reversals and Breakpoints
Show a reversal $\rho$ that reduces the number of breakpoints by two, i.e. $b_\beta(\alpha \rho) = b_\beta(\alpha) - 2$, but does not make progress, i.e. $d_\beta(\alpha \rho) > d_\beta(\alpha) - 1$. You can assume $\beta$ is the identity.

Problem 2: Reversals and Cycles
Show a reversal $\rho$ that increases the number of cycles in the reality and desire graph, i.e. $c_\beta(\alpha \rho) = c_\beta(\alpha) + 1$, but does not make progress, i.e. $d_\beta(\alpha \rho) > d_\beta(\alpha) - 1$. You can assume $\beta$ is the identity.

Problem 3: Two Lower Bounds
We said in class that the lower bound $n + 1 - c_\beta(\alpha)$ based on the number of cycles in $RD_\beta(\alpha)$ is tighter than the lower bound $b_\beta(\alpha)/2$ based on the number of breakpoints. Prove it, i.e. prove that $n + 1 - c_\beta(\alpha) \geq b_\beta(\alpha)/2$. Hint: a non-breakpoint corresponds to a cycle of length 2.

Problem 4: Suffix trees
Describe how you can efficiently (in linear time) determine the following using a suffix tree data structure.

(a) a longest match between $x$ and $y$
(b) a longest unique match between $x$ and $y$ if it exists
(c) a longest repeat in $x$
(d) a longest non-overlapping repeat in $x$

Problem 5: RNA secondary structure
In this problem we assume that the RNA does not form knots.

(a) Given RNA $r = AUGGCAUCCGU$, find a secondary structure with the maximum number of base pairs.
(b) Given RNA $r = r_1...r_n$, show that if $r_1$ and $r_n$ are complementary bases, then there is a secondary structure with maximum number of base pairs that contains the pair $(r_1, r_n)$.
(c) Assume that the RNA structure cannot fold sharply, so every hairpin loop has a length of at least $t$ unpaired bases. Modify the Nussinov algorithm we saw in class to obtain a minimum energy secondary structure with no sharp folds.
(d) Describe how you can use Nussinov algorithm (with and without the variation in (c)) to find a minimum energy secondary structure for a circular RNA in $O(n^3)$ time.