

We say that position  $i$  in  $k$ -mers  $p_1 \dots p_k$  and  $q_1 \dots q_k$  is a **mismatch** if  $p_i \neq q_i$ . example, CGAAT and CGGAC have two mismatches. Our observation that a *DnaA* box may appear with s variations leads to the following generalization of the Pattern Matching Problem:

**Approximate Pattern Matching Problem:** Find all approximate occurrences of a pattern in a string.

**Input:** Two strings *Pattern* and *Text* along with an integer  $d$ .

**Output:** All positions where *Pattern* appears in *Text* with at most  $d$  mismatches.

**CODE CHALLENGE:** Solve the Approximate Pattern Matching Problem

**Sample Input:**

ATTCTGGA

CGCCCGAATCCAGAACGCATTCCCATATTTCGGGACCACTGGCCTCCACGGTACGGACGTCAATCAAAT

3

**Sample Output:**

6 7 26 27