Assignment-7

Dynamic Programming: Edit Distance with Gap Penalty

Edit Distance with Gap Penalty

- Consider two strings **A** and **B** comprising of English uppercase letters, of length **M** and **N** respectively. The distance from **A** to **B** is the total minimum penalty incurred in converting **A** to **B** using the following operations, each of which incurs a penalty as described below.
- Replacing a letter of **A** with another letter incurs a penalty of 2
- Inserting a letter in **A** incurs a penalty of 1 for each letter inserted. In addition, there is also a fixed penalty of **2** for every sequence of letters inserted (one-time penalty for starting the sequence, not per-letter). For example, inserting a one letter sequence (say "A") will incur a penalty of 3 (= **x** + 1), a 2-letter sequence (say "AB") will incur a penalty of 4 (= **x** + 2), and inserting any n-letter sequence will all incur a penalty of 2 + y.
- Deleting a letter from A incurs a penalty of 1 for each letter deleted. In addition, there is also a fixed penalty of **2** for every sequence of letters deleted (one-time penalty for starting the sequence, not per-letter).
- Write a C/C++ program to find the minimum penalty for converting **A** into **B**?

Edit Distance with Gap Penalty

- Example 1: A = “ABCDEFGH” B = “AB” Output: 8
- Example 2: A = “AB” B = “ABCDEFGH” Output: 8
- Example 3: A = “ABCDEFGH” B = “ABAB” Output: 10

- Hints: Create 4 two dimensional arrays as follows.
  - int Edit[M+1][N+1]
  - int Replace[M+1][N+1]
  - int Insert[M+1][N+1]
  - int Delete[M+1][N+1]
  - Why? Figure out ☺

Submission

- Last date: 20-OCT-2024 (till 11:59 P.M.) (Sunday)
- Programming language: C/C++
- Single File: 24CS06001_A7.c/.cpp or 24AI06001_A7.c/.cpp
- Subject Line: 24CS06001_A7 or 24AI06001_A7
- Email to: pds2016autumn@gmail.com