

Word Pair Analysis: Sunday → Saturday

1. Minimum Edit Distance

Strings: - S1 = Sunday - S2 = Saturday

Model A: Substitution = 1, Insertion = 1, Deletion = 1

Model B: Substitution = 2, Insertion = 1, Deletion = 1

Model A

One possible edit sequence: 1. Substitute U → A (1) 2. Insert T after A (1) 3. Insert R after T (1) 4. Insert D after R (1) 5. Insert A before Y (1)

Total cost = 5

Model B

Substitution is expensive; using mostly insertions: 1. Insert A after S (1) 2. Insert T (1) 3. Insert R (1) 4. Insert D (1) 5. Insert A (1)

Total cost = 5

Both models give the same minimum distance (5).

2. Example Edit Sequence (Model A)

Sunday → Saturday

1. Sunday
2. Substitute U → A : Sanday
3. Insert T after A : Satanday
4. Insert R after T : Satranday
5. Insert D after R : Satrdanday
6. Insert A before Y : Saturday

3. Levenshtein DP Table (Sub=2, Ins=1, Del=1)

Strings: - Rows: S1 = SUNDAY - Columns: S2 = SATURDAY

DP Formula:

```
D[i,j] = min(  
    D[i-1,j] + 1,          // deletion
```

```

    D[i,j-1] + 1,          // insertion
    D[i-1,j-1] + cost     // substitution (0 if match, 2 if mismatch)
)

```

First 4 Rows (i=0..3) and All Columns (j=0..8)

	""	S	A	T	U	R	D	A	Y
""	0	1	2	3	4	5	6	7	8
S	1	0	1	2	3	4	5	6	7
U	2	1	2	3	2	3	4	5	6
N	3	2	3	4	3	4	5	6	7
D	4								
A	5								
Y	6								

Value of D(3,4) = 3

(Corresponds to transforming "SUN" → "SATU")

4. Reflection

- Both models resulted in the **same distance of 5**.
- **Insertions** were the most useful operations, especially for Model B.
- Deletions played a minor role since most letters in Sunday remain.
- Substitution was cheaper in Model A, but in Model B it was avoided due to higher cost.
- Model choice affects applications: spell check may prefer substitutions, while DNA alignment may prefer insertions/deletions to preserve sequence information.