Edit Distance
CSCI 532
Dynamic Programming

Dynamic Programming Process:
1. Characterize structure of optimal solution.
2. Recursively define value of optimal solution.
3. Compute value of optimal solution.
4. Construct optimal solution from computed information.
Rod Cutting

partition(n, p)

\[ r[0,...,n] = [0,...,0] \]

for \( m = 1 \) to \( n \)

\[ \text{max} = 0 \]

for \( \text{length } l \leq m \)

\[ \text{if } r[m - 1] + p_l > \text{max} \]

\[ \text{max} = r[m - 1] + p_l \]

\[ r[m] = \text{max} \]

return \( r[n] \)
Rod Cutting

\texttt{partition}(n, p)

\[ r[0,...,n] = [0,...,0] \]

\textbf{for} \( m = 1 \) to \( n \)

\hspace{1cm} \textbf{max} = 0

\hspace{1cm} \textbf{for} \ text{ length } l \leq m

\hspace{2cm} \textbf{if} \ r[m - 1] + \ r[l] > \text{ max}

\hspace{3cm} \text{ max } = r[m - 1] + \ r[l] \]

\hspace{2cm} \text{ } r[m] = \text{ max}

\textbf{return} \ r[n]

<table>
<thead>
<tr>
<th>length</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>...</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>profit</td>
<td>$1</td>
<td>$5</td>
<td>$8</td>
<td>$9</td>
<td>...</td>
<td>$?</td>
</tr>
</tbody>
</table>

Issues?
Rod Cutting

\[
\text{partition}(n, p) \\
\quad r[0,...,n] = [0,...,0] \\
\quad \text{for } m = 1 \text{ to } n \\
\quad \quad \max = 0 \\
\quad \quad \text{for length } l \leq m \\
\quad \quad \quad \text{if } r[m - 1] + r[l] > \max \\
\quad \quad \quad \quad \max = r[m - 1] + r[l] \\
\quad \quad r[m] = \max \\
\quad \text{return } r[n]
\]

<table>
<thead>
<tr>
<th>length</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>...</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>profit</td>
<td>$1</td>
<td>$5</td>
<td>$8</td>
<td>$9</td>
<td>...</td>
<td>$?</td>
</tr>
</tbody>
</table>

Issues?

• Need to initialize \( r[] \) to values from profit table.
• Extra details needed to reconstruct optimal solution.
Edit Distance

Given two strings, how many edits are needed to turn one string into another?

SNOWY vs SUNNY
Edit Distance

Need:
- Strings – Snowy, Sunny
- Cost function - character misalignment = +1

What are the costs of these two different alignments?

\[ S \rightarrow N O W Y \]
\[ S \rightarrow N N Y \]
\[ \text{cost} = ? \]

\[ - \rightarrow S N O W - Y \]
\[ S U N - - N Y \]
\[ \text{cost} = ? \]
Edit Distance

Need:

• Strings – Snowy, Sunny
• Cost function - character misalignment = +1

What are the costs of these two different alignments?

\[
\begin{align*}
\text{S} & \quad \text{N} \quad \text{O} \quad \text{W} \quad \text{Y} & \quad - \quad \text{S} \quad \text{N} \quad \text{O} \quad \text{W} \quad - \quad \text{Y} \\
\text{S} \quad \text{U} \quad \text{N} \quad \text{N} \quad - \quad \text{Y} & \quad - \quad \text{S} \quad \text{U} \quad \text{N} \quad - \quad - \quad \text{N} \quad \text{Y} \\
\text{cost} = 3 & \quad \text{cost} = 5
\end{align*}
\]

Edit distance = cheapest possible alignment.
Edit Distance

We want to align two strings, \( x = [x_1, \ldots, x_n] \) and \( y = [y_1, \ldots, y_m] \).
Edit Distance

We want to align two strings, $x = [x_1, \ldots, x_n]$ and $y = [y_1, \ldots, y_m]$.
Edit Distance

We want to align two strings, \( x = [x_1, \ldots, x_n] \) and \( y = [y_1, \ldots, y_m] \).

\[ E(i, j) = \text{optimal cost of aligning } [x_1, \ldots, x_i] \text{ and } [y_1, \ldots, y_j]. \]
Edit Distance

We want to align two strings, $x = [x_1, ..., x_n]$ and $y = [y_1, ..., y_m]$.

$E(i, j) =$ optimal cost of aligning $[x_1, ..., x_i]$ and $[y_1, ..., y_j]$.

Can we say anything about optimal alignment of $[x_1, ..., x_i]$ and $[y_1, ..., y_j]$?
Edit Distance

We want to align two strings, \( x = [x_1, \ldots, x_n] \) and \( y = [y_1, \ldots, y_m] \).

\[ E(i, j) = \text{optimal cost of aligning } [x_1, \ldots, x_i] \text{ and } [y_1, \ldots, y_j]. \]

Can we say anything about optimal alignment of \([x_1, \ldots, x_i]\) and \([y_1, \ldots, y_j]\)?

Specifically, how must the optimal alignments end?
(three possibilities).
Edit Distance

We want to align two strings, \( x = [x_1, \ldots, x_n] \) and \( y = [y_1, \ldots, y_m] \).

\[ E(i, j) = \text{optimal cost of aligning} \ [x_1, \ldots, x_i] \text{ and } [y_1, \ldots, y_j]. \]

Can we say anything about optimal alignment of \([x_1, \ldots, x_i]\) and \([y_1, \ldots, y_j]\)?

Optimal alignments end in one of three ways:

\[
\begin{align*}
  x_i & \quad - \quad x_i \\
  - & \quad y_j \quad y_j
\end{align*}
\]

Cost: \(1 \quad 1 \quad 0,1\)
Edit Distance

We want to align two strings, $x = [x_1, \ldots, x_n]$ and $y = [y_1, \ldots, y_m]$.

$E(i, j) = \text{optimal cost of aligning } [x_1, \ldots, x_i] \text{ and } [y_1, \ldots, y_j]$.

Can we say anything about optimal alignment of $[x_1, \ldots, x_i]$ and $[y_1, \ldots, y_j]$?

Optimal alignments end in one of three ways:

\[
\begin{align*}
&x_i \quad - \quad x_i \\
&- \quad y_j \quad y_j
\end{align*}
\]

Need to align $[x_1, \ldots, x_{i-1}]$ with $[y_1, \ldots, y_{j-1}]$: $E(i-1, j-1)$
Edit Distance

We want to align two strings, \( x = [x_1, \ldots, x_n] \) and \( y = [y_1, \ldots, y_m] \).

\[ E(i, j) = \text{optimal cost of aligning} \ [x_1, \ldots, x_i] \text{ and } [y_1, \ldots, y_j]. \]

Can we say anything about optimal alignment of \([x_1, \ldots, x_i]\) and \([y_1, \ldots, y_j]\)?

Optimal alignments end in one of three ways:

\[ x_i \quad - \quad x_i \quad \text{Need to align } [x_1, \ldots, x_i] \text{ with } [y_1, \ldots, y_{j-1}]: E(i, j - 1) \]

\[ - \quad y_j \quad y_j \]
Edit Distance

We want to align two strings, \( x = [x_1, \ldots, x_n] \) and \( y = [y_1, \ldots, y_m] \).

\[ E(i, j) = \text{optimal cost of aligning } [x_1, \ldots, x_i] \text{ and } [y_1, \ldots, y_j]. \]

Can we say anything about optimal alignment of \([x_1, \ldots, x_i]\) and \([y_1, \ldots, y_j]\)?

Optimal alignments end in one of three ways:

\[ x_i \quad \quad \quad \quad \quad x_i \]
\[ \quad \quad - \quad \quad \quad - \]
\[ y_j \quad \quad y_j \]

Need to align \([x_1, \ldots, x_{i-1}]\) with \([y_1, \ldots, y_j]\): \( E(i-1, j) \)
Edit Distance

\[ E(i, j) = \min \left\{ \right\} \]
Edit Distance

\[ E(i,j) = \min \left\{ \begin{array}{c}
\end{array} \right\} \]

Need to align \([x_1, ..., x_{i-1}]\) with \([y_1, ..., y_j]\): \(E(i-1,j)\)
Edit Distance

\[
E(i, j) = \min \begin{cases} 
  E(i - 1, j) + 1, & \\
  E(i, j - 1) + 1, & \\
  E(i - 1, j - 1) + \text{diff}(i, j), & \\
\end{cases}
\]

where \( \text{diff}(i, j) = \begin{cases} 
  0, & x_i = y_j \\
  1, & x_i \neq y_j \\
\end{cases} \)
Edit Distance

\[
E(i, j) = \min \left\{ \begin{array}{l}
E(i - 1, j) + 1 \\
E(i, j - 1) + 1 \\
E(i - 1, j - 1) + \text{diff}(i, j)
\end{array} \right.
\]

where \( \text{diff}(i, j) = \begin{cases} 
0, & x_i = y_j \\
1, & x_i \neq y_j
\end{cases} \)

How should we find \( E(n, m) \)?
Edit Distance

\[ E(i, j) = \min \left\{ \begin{array}{ll}
E(i - 1, j) + 1 \\
E(i, j - 1) + 1 \\
E(i - 1, j - 1) + \text{diff}(i, j) \\
\end{array} \right. \]

where \( \text{diff}(i, j) = \begin{cases} 
0, & x_i = y_j \\
1, & x_i \neq y_j 
\end{cases} \)

How should we find \( E(n, m) \)?
- Find all the other \( E(i, j) \)’s for \( i < n, j < m \).
Edit Distance

\[ E(i, j) = \min \left\{ \begin{array}{ll} E(i-1, j) + 1 \\
E(i, j-1) + 1 \\
E(i-1, j-1) + \text{diff}(i, j) \end{array} \right. \]

where \( \text{diff}(i, j) = \begin{cases} 0, & x_i = y_j \\ 1, & x_i \neq y_j \end{cases} \)

How should we find \( E(n, m) \)?

- Find all the other \( E(i, j) \)'s for \( i < n, j < m \).
- Store intermediate \( E(i, j) \) values in a 2d array.
### Edit Distance

$$E(i, j) = \min \left\{ \begin{array}{l}
E(i - 1, j) + 1 \\
E(i, j - 1) + 1 \\
E(i - 1, j - 1) + \text{diff}(i, j)
\end{array} \right\}$$

$$\text{diff}(i, j) = \begin{cases} 
0, & x[i] = y[j] \\
1, & x[i] \neq y[j]
\end{cases}$$

Where can we start?
Edit Distance

$E(i, j) = \min \begin{cases} 
E(i - 1, j) + 1 \\
E(i, j - 1) + 1 \\
E(i - 1, j - 1) + \text{diff}(i, j)
\end{cases}
$

\text{diff}(i, j) = \begin{cases} 
0, & x[i] = y[j] \\
1, & x[i] \neq y[j]
\end{cases}

Where can we start? 
$E(0, 1)$ or $E(1, 0)$
Edit Distance

\[ E(i, j) = \min \begin{cases} 
E(i - 1, j) + 1 \\
E(i, j - 1) + 1 \\
E(i - 1, j - 1) + \text{diff}(i, j)
\end{cases} \]

\[ \text{diff}(i, j) = \begin{cases} 
0, & x[i] = y[j] \\
1, & \text{otherwise}
\end{cases} \]

\[ E(0, 1) = \min \begin{cases} 
E(-1, 1) + 1 \\
E(0, 0) + 1 \\
E(-1, 0) + 1
\end{cases} = ? \]
Edit Distance

\[ E(i, j) = \min \begin{cases} 
E(i-1, j) + 1 \\
E(i, j-1) + 1 \\
E(i-1, j-1) + \text{diff}(i, j) 
\end{cases} \]

\[ \text{diff}(i, j) = \begin{cases} 
0, & x[i] = y[j] \\
1, & \text{otherwise} 
\end{cases} \]

\[ E(0, 1) = \min \begin{cases} 
E(1, 1) + 1 \\
E(0, 0) + 1 \\
E(1, 0) + 1 
\end{cases} = ? \]
Edit Distance

\[
E(i, j) = \min \begin{cases} 
E(i - 1, j) + 1 \\
E(i, j - 1) + 1 \\
E(i - 1, j - 1) + \text{diff}(i, j)
\end{cases}
\]

\[
\text{diff}(i, j) = \begin{cases} 
0, & x[i] = y[j] \\
1, & \text{otherwise}
\end{cases}
\]

\[
E(0, 1) = \min \begin{cases} 
E(1, 1) + 1 \\
E(0, 0) + 1 \\
E(1, 0) + 1
\end{cases} = 1
\]
Edit Distance

\[ E(i, j) = \min \begin{cases} 
E(i - 1, j) + 1, \\
E(i, j - 1) + 1, \\
E(i - 1, j - 1) + \text{diff}(i, j) 
\end{cases} \]

\[ \text{diff}(i, j) = \begin{cases} 
0, & x[i] = y[j] \\
1, & \text{otherwise} 
\end{cases} \]

\[ E(1, 1) = \min \begin{cases} 
E(0, 1) + 1, \\
E(1, 0) + 1, \\
E(0, 0) + 0 
\end{cases} \]
Edit Distance

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>O</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>W</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
E(i, j) = \min \left\{ \begin{array}{l}
E(i - 1, j) + 1 \\
E(i, j - 1) + 1 \\
E(i - 1, j - 1) + \text{diff}(i, j)
\end{array} \right.
\]

\[
\text{diff}(i, j) = \begin{cases}
0, & x[i] = y[j] \\
1, & \text{otherwise}
\end{cases}
\]

\[
E(1, 1) = \min \left\{ \begin{array}{l}
E(0, 1) + 1 \\
E(1, 0) + 1 \\
E(0, 0) + 0
\end{array} \right. = ?
\]

Not calculated yet!
## Edit Distance

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>S</td>
<td>U</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Equation:**

\[
E(i, j) = \min \left\{ \begin{array}{ll}
E(i-1, j) + 1 \\
E(i, j-1) + 1 \\
E(i-1, j-1) + \text{diff}(i, j)
\end{array} \right.
\]

**Definition:**

\[
\text{diff}(i, j) = \begin{cases} 
0, & x[i] = y[j] \\
1, & \text{otherwise}
\end{cases}
\]

**Note:**

Need upper left hand corner filled out before we can progress.
# Edit Distance

\[
E(i, j) = \min \left\{ \begin{array}{ll}
E(i-1, j) + 1 \\
E(i, j-1) + 1 \\
E(i-1, j-1) + \text{diff}(i, j)
\end{array} \right.
\]

\[
\text{diff}(i, j) = \begin{cases} 
0, & x[i] = y[j] \\
1, & \text{otherwise}
\end{cases}
\]

\[
E(0, 2) = \min \left\{ \begin{array}{ll}
E(-1, 2) + 1 \\
E(0, 1) + 1 \\
E(-1, 1) + 1
\end{array} \right. = 2
\]

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>O</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>W</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Edit Distance

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>O</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>W</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The Edit Distance between two strings $X = X_1 X_2 \ldots X_n$ and $Y = Y_1 Y_2 \ldots Y_m$ is defined as the minimum number of operations required to transform the string $X$ into the string $Y$. The operations are:

- **Insert**: Insert a new character to the string $X$.
- **Delete**: Delete a character from the string $X$.
- **Substitute**: Substitute a character in the string $X$.

### Dynamic Programming Formula

$$E(i, j) = \min\left( \begin{array}{c} E(i-1, j) + 1 \\ E(i, j-1) + 1 \\ E(i-1, j-1) + \text{diff}(i, j) \end{array} \right)$$

### Example

In the table:

- $E(0, 0) = 0$
- $E(1, 0) = \min\left( \begin{array}{c} E(0, 0) + 1 \\ E(1, -1) + 1 = 1 \end{array} \right)$
- $\text{diff}(i, j) = \begin{cases} 0, & x[i] = y[j] \\ 1, & \text{otherwise} \end{cases}$
Edit Distance

\[
E(i, j) = \min \left\{ \begin{array}{ll}
E(i - 1, j) + 1 \\
E(i, j - 1) + 1 \\
E(i - 1, j - 1) + \text{diff}(i, j)
\end{array} \right.
\]

\[
\text{diff}(i, j) = \begin{cases}
0, & x[i] = y[j] \\
1, & \text{otherwise}
\end{cases}
\]

\[
E(1, 1) = \min \left\{ \begin{array}{ll}
E(0, 1) + 1 \\
E(1, 0) + 1 = 0 \\
E(0, 0) + 0
\end{array} \right.
\]
## Edit Distance


<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>S</td>
<td>U</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

\[
E(i, j) = \min\left\{ E(i-1, j) + 1, \quad E(i, j-1) + 1, \quad E(i-1, j-1) + \text{diff}(i, j) \right\}
\]

\[
\text{diff}(i, j) = \begin{cases} 
0, & x[i] = y[j] \\
1, & \text{otherwise}
\end{cases}
\]

Running Time?
### Edit Distance

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>S</td>
<td>U</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>S</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>O</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>W</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Y</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

$E(i, j) = \min \left\{ \begin{array}{c} E(i - 1, j) + 1 \\ E(i, j - 1) + 1 \\ E(i - 1, j - 1) + \text{diff}(i, j) \end{array} \right.$

$\text{diff}(i, j) = \begin{cases} 0, & x[i] = y[j] \\ 1, & \text{otherwise} \end{cases}$

Running Time?

Fill out $n \times m$ table with constant operations: $O(nm)$
Edit Distance

<table>
<thead>
<tr>
<th>i</th>
<th>S</th>
<th>U</th>
<th>N</th>
<th>N</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Edit distance = 3.

How can we recreate the actual alignments?

Backtracking.

Ask the question: “How did we get here?”
Edit Distance

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>S</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>O</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>W</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Y</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

How did we get to $E(5,5)$?
How did we get to $E(5,5)$?
From $E(5,4)$?
How did we get to $E(5,5)$?
## Edit Distance

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>S</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>O</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>W</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Y</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

How did we get to $E(5,5)$? 
From $E(4,5)$?
# Edit Distance

<table>
<thead>
<tr>
<th>i</th>
<th>S</th>
<th>U</th>
<th>N</th>
<th>N</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>S</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>O</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>W</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Y</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

How did we get to $E(5,5)$?


From $E(4,5)$? – No. Need +1 to move that direction.

$$E(i, j) = \min \left\{ \begin{array}{l}
E(i-1, j) + 1 \\
E(i, j-1) + 1 \\
E(i-1, j-1) + \text{diff}(i, j)
\end{array} \right.$$
Edit Distance

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>S</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>O</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>W</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Y</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

How did we get to \( E(5,5) \)?
From \( E(5,4) \)? – No. Can never go down in cost.
From \( E(4,5) \)? – No. Need +1 to move that direction.
From \( E(4,4) \)?

\[
E(i,j) = \min \begin{cases} 
E(i-1,j) + 1 \\
E(i,j-1) + 1 \\
E(i-1,j-1) + \text{diff}(i,j)
\end{cases}
\]
Edit Distance

How did we get to $E(5,5)$?


From $E(4,5)$? – No. Need +1 to move that direction.

From $E(4,4)$? – Yes. Match Y’s.

$$E(i,j) = \min \begin{cases} 
E(i-1,j) + 1 \\
E(i,j-1) + 1 \\
E(i-1,j-1) + \text{diff}(i,j)
\end{cases}$$
Edit Distance

Continuing the process yields all of the optimal solutions.

Diagonal move indicates \(?\)

Vertical move indicates \(?\)

Horizontal move indicates \(?\)

\[
E(i, j) = \min \begin{cases} 
E(i - 1, j) + 1 \\
E(i, j - 1) + 1 \\
E(i - 1, j - 1) + \text{diff}(i, j) 
\end{cases}
\]
Edit Distance

Continuing the process yields all of the optimal solutions.

Diagonal move indicates match.

Vertical move indicates complement.

Horizontal move indicates substitution.

$$E(i,j) = \min \left\{ \begin{array}{ll} E(i-1,j) + 1 \\ E(i,j-1) + 1 \\ E(i-1,j-1) + \text{diff}(i,j) \end{array} \right\}$$
Edit Distance

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>S</td>
<td>U</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

Continuing the process yields all of the optimal solutions.

Diagonal move indicates match.

Vertical move indicates space inserted in \( j \).

Horizontal move indicates ?

\[
E(i,j) = \min \left\{ \begin{array}{ll}
E(i-1,j) + 1 \\
E(i,j-1) + 1 \\
E(i-1,j-1) + \text{diff}(i,j)
\end{array} \right.
\]
Edit Distance

Continuing the process yields all of the optimal solutions.

Diagonal move indicates match.

Vertical move indicates space inserted in $j$.

Horizontal move indicates space inserted in $i$.

$E(i, j) = \min \left\{ \begin{array}{c}
E(i-1, j) + 1 \\
E(i, j-1) + 1 \\
E(i-1, j-1) + \text{diff}(i, j)
\end{array} \right. $
Edit Distance

Diagonal move indicates match.

Vertical move indicates space inserted in \( j \).

Horizontal move indicates space inserted in \( i \).

\[
\begin{array}{ccccccc}
  & 0 & 1 & 2 & 3 & 4 & 5 \\
\hline
0 & 0 & 1 & 2 & 3 & 4 & 5 \\
S & 1 & 0 & 1 & 2 & 3 & 4 \\
U & 2 & 1 & 1 & 2 & 3 & 3 \\
N & 3 & 2 & 2 & 2 & 3 & 3 \\
Y & 4 & 3 & 3 & 3 & 3 & 3 \\
\end{array}
\]
Edit Distance

Diagonal move indicates match.

Vertical move indicates space inserted in $j$.

Horizontal move indicates space inserted in $i$.

Alignment?
## Edit Distance

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>S</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>O</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>W</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Y</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Diagonal move indicates match.

Vertical move indicates space inserted in $j$.

Horizontal move indicates space inserted in $i$.

$S$ – $N$ $O$ $W$ $Y$

$S$ $U$ $N$ – $N$ $Y$
Edit Distance

Diagonal move indicates match.

Vertical move indicates space inserted in \( j \).

Horizontal move indicates space inserted in \( i \).

\[
\begin{array}{ccccccc}
  & 0 & 1 & 2 & 3 & 4 & 5 \\
 0 & 0 & 1 & 2 & 3 & 4 & 5 \\
 1 & 1 & 0 & 1 & 2 & 3 & 4 \\
 2 & 2 & 1 & 1 & 1 & 2 & 3 \\
 3 & 3 & 2 & 2 & 2 & 2 & 3 \\
 4 & 4 & 3 & 3 & 3 & 3 & 3 \\
 5 & 5 & 4 & 4 & 4 & 4 & 3 \\
\end{array}
\]

\[
\text{S N O W Y} \\
\text{S U N N Y}
\]
## Edit Distance

$E(i, j) = \min \left\{ \begin{array}{l} E(i - 1, j) + 1 \\ E(i, j - 1) + 1 \\ E(i - 1, j - 1) + \text{diff}(i, j) \end{array} \right.$

\[
\text{diff}(i, j) = \begin{cases} 0, & x[i] = y[j] \\ 1, & \text{otherwise} \end{cases}
\]

<table>
<thead>
<tr>
<th>$j$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i$</td>
<td>S</td>
<td>U</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

Space Requirements?
## Edit Distance

$$E(i, j) = \min\begin{cases} E(i - 1, j) + 1 \\ E(i, j - 1) + 1 \\ E(i - 1, j - 1) + \text{diff}(i, j) \end{cases}$$

$$\text{diff}(i, j) = \begin{cases} 0, & x[i] = y[j] \\ 1, & \text{otherwise} \end{cases}$$

Space Requirements?

$n \times m$ table: $O(nm)$
## Edit Distance

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>i</strong></td>
<td><strong>S</strong></td>
<td><strong>U</strong></td>
<td><strong>N</strong></td>
<td><strong>N</strong></td>
<td><strong>Y</strong></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

\[
E(i, j) = \min \begin{cases} 
E(i - 1, j) + 1 \\
E(i, j - 1) + 1 \\
E(i - 1, j - 1) + \text{diff}(i, j)
\end{cases}
\]

\[
\text{diff}(i, j) = \begin{cases} 
0, & x[i] = y[j] \\
1, & \text{otherwise}
\end{cases}
\]

Space Requirements?

\[n \times m \text{ table: } O(nm)\]

Can we do it in \(O(\min(n, m))\) space?
**Edit Distance**

Edit distance is a measure of the dissimilarity between two sequences. It is defined as the minimum number of operations required to transform one sequence into another. In the context of the edit distance, we consider three basic operations:

- Insertion: Adding an element to the sequence.
- Deletion: Removing an element from the sequence.
- Substitution: Replacing an element in the sequence with another.

Given two sequences, let's denote them as $S$ and $U$. We need to calculate the edit distance between them, where $S$ is the source sequence and $U$ is the target sequence.

### Table Representation

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>$i$</strong></td>
<td>S</td>
<td>U</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>S</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>O</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>W</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Y</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

### Edit Distance Formula

The edit distance $E(i, j)$ between the prefixes $S[1..i]$ and $U[1..j]$ is defined as:

$$ E(i, j) = \min \left( \begin{array}{l} E(i-1, j) + 1 \\ E(i, j-1) + 1 \\ E(i-1, j-1) + \text{diff}(i, j) \end{array} \right) $$

where $	ext{diff}(i, j)$ is the cost of changing the $i$th element of $S$ to the $j$th element of $U$:

$$ \text{diff}(i, j) = \begin{cases} 0, & x[i] = y[j] \\ 1, & \text{otherwise} \end{cases} $$

### Space Requirements

The standard algorithm requires an $n \times m$ table, which has a space complexity of $O(nm)$. However, is it possible to compute the edit distance in $O(\min(n, m))$ space?

**Can we do it in $O(\min(n, m))$ space?**
## Edit Distance

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>i</strong></td>
<td><strong>S</strong></td>
<td><strong>U</strong></td>
<td><strong>N</strong></td>
<td><strong>N</strong></td>
<td><strong>Y</strong></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

\[
E(i, j) = \min \begin{cases} E(i - 1, j) + 1 \\ E(i, j - 1) + 1 \\ E(i - 1, j - 1) + \text{diff}(i, j) \end{cases}
\]

\[
\text{diff}(i, j) = \begin{cases} 0, & x[i] = y[j] \\ 1, & \text{otherwise} \end{cases}
\]

**Space Requirements?**

- \(n \times m\) table: \(O(nm)\)
- Can we do it in \(O(\min(n, m))\) space?
Edit Distance

\[
E(i, j) = \min \begin{cases}
E(i - 1, j) + 1 \\
E(i, j - 1) + 1 \\
E(i - 1, j - 1) + \text{diff}(i, j)
\end{cases}
\]

\[
\text{diff}(i, j) = \begin{cases}
0, & x[i] = y[j] \\
1, & \text{otherwise}
\end{cases}
\]

Space Requirements?

A \( n \times m \) table: \( O(nm) \)

Can we do it in \( O(\min(n, m)) \) space?
# Edit Distance

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>i</strong></td>
<td>SUNN</td>
<td>SUNN</td>
<td>SUNN</td>
<td>SUNN</td>
<td>SUNN</td>
<td>SUNN</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>S</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>O</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>W</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Y</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

$$E(i, j) = \min \left\{ E(i - 1, j) + 1, E(i, j - 1) + 1, E(i - 1, j - 1) + \text{diff}(i, j) \right\}$$

$$\text{diff}(i, j) = \begin{cases} 
0, & x[i] = y[j] \\
1, & \text{otherwise}
\end{cases}$$

Space Requirements?

$n \times m$ table: $O(nm)$

Can we do it in $O(\min(n, m))$ space?
Edit Distance

\[ E(i, j) = \min \begin{cases} 
E(i - 1, j) + 1 \\
E(i, j - 1) + 1 \\
E(i - 1, j - 1) + \text{diff}(i, j)
\end{cases} \]

\[ \text{diff}(i, j) = \begin{cases} 
0, & x[i] = y[j] \\
1, & \text{otherwise}
\end{cases} \]

Space Requirements?

\( n \times m \) table: \( O(nm) \)

Can we do it in \( O(\min(n, m)) \) space?
Edit Distance

\[ E(i, j) = \min \left\{ \begin{array}{l}
E(i-1, j) + 1 \\
E(i, j-1) + 1 \\
E(i-1, j-1) + \text{diff}(i, j)
\end{array} \right. \]

\[ \text{diff}(i, j) = \begin{cases} 
0, & x[i] = y[j] \\
1, & \text{otherwise}
\end{cases} \]

<table>
<thead>
<tr>
<th>( j )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( i )</td>
<td>S</td>
<td>U</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>S</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>O</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>W</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Y</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Space Requirements?

\[ n \times m \text{ table: } O(nm) \]

Can we do it in \( O(\min(n, m)) \) space?
Longest Common Subsequence

**Subsequence**: Sequence that can be derived from another by deleting elements without changing order.
Longest Common Subsequence

Subsequence: Sequence that can be derived from another by deleting elements without changing order.

Common Subsequence: A sequence that is a subsequence of two sequences

```
SUNNY     SNOWY
     SNY
      SY
```
Longest Common Subsequence

**Subsequence**: Sequence that can be derived from another by deleting elements without changing order.

**Common Subsequence**: A sequence that is a subsequence of two sequences

**Longest Common Subsequence**: A maximum length common subsequence.
Longest Common Subsequence

\[ X = \langle x_1, \ldots, x_m \rangle, \quad Y = \langle y_1, \ldots, y_n \rangle - \text{Sequences} \]
\[ Z = \langle z_1, \ldots, z_k \rangle - \text{Any LCS of } X \text{ and } Y. \]

Optimal Substructure?
Longest Common Subsequence

\[ X = \langle x_1, ..., x_m \rangle, \ Y = \langle y_1, ..., y_n \rangle - \text{Sequences} \]
\[ Z = \langle z_1, ..., z_k \rangle - \text{Any LCS of } X \text{ and } Y. \]

Claims:

1 – 

2 – 

3 – 
Longest Common Subsequence

\[ X = \langle x_1, \ldots, x_m \rangle, \quad Y = \langle y_1, \ldots, y_n \rangle \] - Sequences

\[ Z = \langle z_1, \ldots, z_k \rangle \] - Any LCS of \( X \) and \( Y \).

Claims:

1 – If \( x_m = y_n \), then \( z_k = x_m = y_n \) and \( Z_{k-1} \) is an LCS of \( X_{m-1} \) and \( Y_{n-1} \).

2 –

3 –
Longest Common Subsequence

\[ X = \langle x_1, \ldots, x_m \rangle, \ Y = \langle y_1, \ldots, y_n \rangle \] - Sequences
\[ Z = \langle z_1, \ldots, z_k \rangle \] - Any LCS of \( X \) and \( Y \).

Claims:

1 – If \( x_m = y_n \), then \( z_k = x_m = y_n \) and \( Z_{k-1} \) is an LCS of \( X_{m-1} \) and \( Y_{n-1} \).

2 – If \( x_m \neq y_n \), then \( z_k \neq x_m \) implies \( Z \) is an LCS of \( X_{m-1} \) and \( Y \).

3 –
Longest Common Subsequence

\[ X = \langle x_1, \ldots, x_m \rangle, \quad Y = \langle y_1, \ldots, y_n \rangle \quad \text{- Sequences} \]

\[ Z = \langle z_1, \ldots, z_k \rangle \quad \text{- Any LCS of } X \text{ and } Y. \]

Claims:

1 – If \( x_m = y_n \), then \( z_k = x_m = y_n \) and \( Z_{k-1} \) is an LCS of \( X_{m-1} \) and \( Y_{n-1} \).

2 – If \( x_m \neq y_n \), then \( z_k \neq x_m \) implies \( Z \) is an LCS of \( X_{m-1} \) and \( Y \).

3 – If \( x_m \) is not the last character of \( Z \), then \( Z \) is the LCS of all of \( Y \) and everything up to \( x_m \).
Longest Common Subsequence

\[ X = \langle x_1, ..., x_m \rangle, \quad Y = \langle y_1, ..., y_n \rangle - \text{Sequences} \]
\[ Z = \langle z_1, ..., z_k \rangle - \text{Any LCS of } X \text{ and } Y. \]

Claims:

1. If \( x_m = y_n \), then \( z_k = x_m = y_n \) and \( Z_{k-1} \) is an LCS of \( X_{m-1} \) and \( Y_{n-1} \).

2. If \( x_m \neq y_n \), then \( z_k \neq x_m \) implies \( Z \) is an LCS of \( X_{m-1} \) and \( Y \).

3. If \( x_m \neq y_n \), then \( z_k \neq y_n \) implies \( Z \) is an LCS of \( X \) and \( Y_{n-1} \).
Longest Common Subsequence

1 – If \( x_m = y_n \), then \( z_k = x_m = y_n \) and \( Z_{k-1} \) is an LCS of \( X_{m-1} \) and \( Y_{n-1} \).

2 – If \( x_m \neq y_n \), then \( z_k \neq x_m \) implies \( Z \) is an LCS of \( X_{m-1} \) and \( Y \).

3 – If \( x_m \neq y_n \), then \( z_k \neq y_n \) implies \( Z \) is an LCS of \( X \) and \( Y_{n-1} \).
Longest Common Subsequence

1 – If $x_m = y_n$, then $z_k = x_m = y_n$ and $Z_{k-1}$ is an LCS of $X_{m-1}$ and $Y_{n-1}$.

2 – If $x_m \neq y_n$, then $z_k \neq x_m$ implies $Z$ is an LCS of $X_{m-1}$ and $Y$.

3 – If $x_m \neq y_n$, then $z_k \neq y_n$ implies $Z$ is an LCS of $X$ and $Y_{n-1}$.

Let $c(i, j) = \text{length of LCS of sequences } X_i \text{ and } Y_j$. 
Longest Common Subsequence

1 – If $x_m = y_n$, then $z_k = x_m = y_n$ and $Z_{k-1}$ is an LCS of $X_{m-1}$ and $Y_{n-1}$.

2 – If $x_m \neq y_n$, then $z_k \neq x_m$ implies $Z$ is an LCS of $X_{m-1}$ and $Y$.

3 – If $x_m \neq y_n$, then $z_k \neq y_n$ implies $Z$ is an LCS of $X$ and $Y_{n-1}$.

Let $c(i, j) =$ length of LCS of sequences $X_i$ and $Y_j$.

$c(i, j) = \ ?$
Longest Common Subsequence

1. If \( x_m = y_n \), then \( z_k = x_m = y_n \) and \( Z_{k-1} \) is an LCS of \( X_{m-1} \) and \( Y_{n-1} \).

2. If \( x_m \neq y_n \), then \( z_k \neq x_m \) implies \( Z \) is an LCS of \( X_{m-1} \) and \( Y \).

3. If \( x_m \neq y_n \), then \( z_k \neq y_n \) implies \( Z \) is an LCS of \( X \) and \( Y_{n-1} \).

Let \( c(i, j) = \) length of LCS of sequences \( X_i \) and \( Y_j \).

\[
c(i, j) = \begin{cases} 
0 & i = 0 \text{ or } j = 0 \\
(c(i - 1, j - 1) + 1) & i, j > 0 \text{ and } x_i = y_j \\
\max(c(i, j - 1), c(i - 1, j)) & i, j > 0 \text{ and } x_i \neq y_j 
\end{cases}
\]
Longest Common Subsequence

\[
c(i, j) = \begin{cases} 
0 & \text{if } i = 0 \text{ or } j = 0 \\
c(i - 1, j - 1) + 1 & \text{if } i, j > 0 \text{ and } x_i = y_j \\
\max(c(i, j - 1), c(i - 1, j)) & \text{if } i, j > 0 \text{ and } x_i \neq y_j 
\end{cases}
\]
Longest Common Subsequence

\[
c(i, j) = \begin{cases} 
0 & i = 0 \text{ or } j = 0 \\
\ c(i-1, j-1) + 1 & i, j > 0 \text{ and } x_i = y_j \\
\ \max(c(i, j-1), c(i-1, j)) & i, j > 0 \text{ and } x_i \neq y_j
\end{cases}
\]

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>S</td>
<td>U</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>S</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>O</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>W</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>Y</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

SUNNY  SNOWY  

SUNNY  SNO\_Y  

SNY