



COMP90049 – Knowledge Technologies

Week4 – Workshop slides

Aug 2019

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	?					
e	-2						
a	-3						
d	-4						
e	-5						
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

$A[j][k] = \max (A[j-1][k] + i, //\text{Insert}$
 $A[j][k-1] + d, //\text{Deletion}$
 $A[j-1][k-1] + \text{equal}(f[k-1], t[j-1])); //\text{Replace or match}$

Max {

$-1 + (-1) = -2$	<i>insert</i>
$-1 + (-1) = -2$	<i>delete</i>
$0 + \underbrace{\text{equal}('l', 'd')}_{-1}$	<i>replace</i>

$\Rightarrow -1$

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-1	?				
e	-2						
a	-3						
d	-4						
e	-5						
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

Max {

$-2 + (-1) = -3$	insert	$\Rightarrow -2$
$-1 + (-1) = -2$	delete	
$-1 + \underbrace{\text{equal}('e', 'd')}_{-1} = -2$	replace	

delete
or
replace

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-1	-1	-2			
e	-2						
a	-3						
d	-4						
e	-5						
n	-6						

Annotations:

- A black arrow points from the value -1 at position $(d, 1)$ to the value -1 at position $(e, 1)$.
- A red arrow points from the value -1 at position $(e, 1)$ to the value -2 at position $(e, 2)$.
- A red arrow points from the value -2 at position $(e, 2)$ to the value -1 at position $(d, 2)$, which is crossed out with a red line.

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-1	-2	-3	?		
e	-2						
a	-3						
d	-4						
e	-5						
n	-6						

Annotations:

- Red arrows point from the bottom row (n) to the second column (-1), third column (-2), and fourth column (-3).
- A black arrow points from the first column (0) to the second column (-1).
- The value -3 is crossed out with a red line through it.
- A question mark is placed at the end of the sequence in the fourth column (-3).

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

Max

$$\left\{ \begin{array}{ll} -4 + (-1) = -5 & \text{insert} \\ -3 + (-1) = -4 & \text{delete} \\ -3 + \underbrace{\text{equal}('d', 'd')}_{+1} = -2 & \text{match} \end{array} \right\} \Rightarrow -2 \text{ match}$$

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-1	-2	-3	-2	?	
e	-2						
a	-3						
d	-4						
e	-5						
n	-6						

Annotations:

- Red arrows point from the bottom row (n) to the second column (1), the fourth column (e), and the fifth column (n).
- A black arrow points from the first column (ε) to the second column (1).
- The value m is written above the arrow pointing from the fourth column (e) to the fifth column (n).

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

Max

$$\left\{ \begin{array}{ll} -5 + (-1) = -6 & \text{insert} \\ -2 + (-1) = -3 & \text{delete} \\ -4 + \underbrace{\text{equal('d', 'e')}}_{-1} = -5 & \text{replace} \end{array} \right\} \Rightarrow -3 \quad \text{delete}$$

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-1	-2	-3	-2	-3	?
e	-2						
a	-3						
d	-4						
e	-5						
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

Max

$$\left\{ \begin{array}{l} -6 + (-1) = -7 \quad \text{insert} \\ -3 + (-1) = -4 \quad \text{delete} \\ -5 + \underbrace{\text{equal}('d', 'd')}_{+1} = -4 \quad \text{match} \end{array} \right\} \Rightarrow -4$$

delete
or
match

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-1	-2	-3	-2	-3	-4
e	-2	?					
a	-3						
d	-4						
e	-5						
n	-6						

The diagram illustrates a global edit distance matrix with handwritten annotations. The columns are labeled with symbols: ε , 1, e, n, d, e, d. The rows are labeled with symbols: ε , d, e, a, d, e, n. The matrix entries are numerical values: 0, -1, -2, -3, -4, -5, -6. Handwritten arrows indicate operations: a vertical arrow from ε to d is labeled 'd'; a horizontal arrow from ε to 1 is labeled '1'; a diagonal arrow from ε to e is labeled 'e'; a diagonal arrow from 1 to n is labeled 'n'; a diagonal arrow from n to d is labeled 'd'; a diagonal arrow from d to e is labeled 'e'; a diagonal arrow from e to d is labeled 'd'. A question mark '?' is placed under the 'e' row at the position of the value -2.

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

Max

$$\left\{ \begin{array}{l} -1 + (-1) = -2 \text{ insert} \\ -2 + (-1) = -3 \text{ delete} \\ -1 + \underbrace{\text{equal}('l', 'e')}_{-1} = -2 \text{ match} \end{array} \right\} \Rightarrow -2$$

insert
or
match

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-1	-2	-3	-2	-3	-4
e	-2	-2	?				
a	-3						
d	-4						
e	-5						
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

Max

$$\left\{ \begin{array}{l} -2 + (-1) = -3 \quad \text{insert} \\ -2 + (-1) = -3 \quad \text{delete} \\ -1 + \underbrace{\text{equal}('e', 'e')}_{+1} = 0 \quad \text{match} \end{array} \right\} \Rightarrow 0 \quad \text{match}$$

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-1	-2	-3	-2	-3	-4
e	-2	-2	0				
a	-3						
d	-4						
e	-5						
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

Max

$$\left\{ \begin{array}{l} -3 + (-1) = -4 \quad \text{insert} \\ 0 + (-1) = -1 \quad \text{delete} \\ -2 + \underbrace{\text{equal}('n', 'e')}_{-1} = -3 \quad \text{replace} \end{array} \right\} \Rightarrow -1 \quad \text{delete}$$

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-1	-2	-3	-2	-3	-4
e	-2	-2	0	-1			
a	-3						
d	-4						
e	-5						
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-1	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2		
a	-3						
d	-4						
e	-5						
n	-6						

The diagram illustrates the edit operations between two sequences. The top row represents the target sequence: $\varepsilon, 1, e, n, d, e, d$. The left column represents the source sequence: $\varepsilon, d, e, a, d, e, n$. The grid shows the cost of each operation:

- Match:** Diagonal arrows from $\varepsilon \rightarrow \varepsilon$, $1 \rightarrow d$, $e \rightarrow e$, $n \rightarrow n$, $d \rightarrow d$, $e \rightarrow e$.
- Substitution:** Diagonal arrows from $1 \rightarrow e$, $d \rightarrow n$, $e \rightarrow d$.
- Insertion:** Upward arrows from $d \rightarrow \varepsilon$, $e \rightarrow 1$, $a \rightarrow e$, $d \rightarrow n$, $e \rightarrow d$.
- Deletion:** Downward arrows from $\varepsilon \rightarrow d$, $d \rightarrow e$, $e \rightarrow a$, $d \rightarrow d$, $e \rightarrow n$.

The values in the grid represent the edit distance at each position, starting from 0 at the top-left corner.

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-1	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2	m	-1
a	-3						
d	-4						
e	-5						
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-1	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2	-1	-2
a	-3						
d	-4						
e	-5						
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-1	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2	-1	-2
a	-3	-3					
d	-4						
e	-5						
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-1	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2	-1	-2
a	-3	-3	-1	i			
d	-4						
e	-5						
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-1	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2	-1	-2
a	-3	-3	-1	-1	-2	-1	-2
d	-4						
e	-5						
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-1	-1	-2
a	-3	-3	-1	-1	-2	-2	
d	-4						
e	-5						
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2	-1	-2
a	-3	-3	-1	-1	-2	-2	
d	-4						
e	-5						
n	-6						

The diagram illustrates the edit operations between the sequences '1' and 'd'. The columns represent the sequence '1' and the rows represent the sequence 'd'. The values in the grid represent the cost of each operation. The arrows indicate the type of edit:

- Upward arrow: Insertion of ε (cost 0).
- Downward arrow: Deletion of ε (cost 0).
- Rightward arrow: Substitution/match (cost 0).
- Diagonal up-right arrow: Substitution/match (cost 0).
- Diagonal down-right arrow: Deletion of ε (cost 0).
- Horizontal rightward arrow: Insertion of ε (cost 1).
- Horizontal downward arrow: Deletion of ε (cost 1).
- Vertical upward arrow: Insertion of ε (cost 1).
- Vertical downward arrow: Deletion of ε (cost 1).

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2	-1	-2
a	-3	-3	-1	-1	-2	-2	-2
d	-4						
e	-5						
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2	-1	-2
a	-3	-3	-1	-1	-2	-2	-2
r	-4	-4					
i							
d							
e							
n							

The grid illustrates the edit operations between the sequences "rider" and "distance". The columns represent "rider" and the rows represent "distance". The cost values are as follows:

- Row 1 (ε):** ε, 1, e, n, d, e, d. Values: 0, -1, -2, -3, -4, -5, -6.
- Row 2 (d):** ε, -1, -2, -2, -3, -2, -3, -4. Values: -1, -2, -2, -3, -2, -3, -4.
- Row 3 (e):** ε, -2, -2, 0, -1, -2, -1, -2. Values: -2, -2, 0, -1, -2, -1, -2.
- Row 4 (a):** ε, -3, -3, -1, -1, -2, -2, -2. Values: -3, -3, -1, -1, -2, -2, -2.
- Row 5 (r):** ε, -4, -4, -4, -4, -4, -4, -4. Values: -4, -4, -4, -4, -4, -4, -4.
- Row 6 (i):** ε, -4, -4, -4, -4, -4, -4, -4. Values: -4, -4, -4, -4, -4, -4, -4.
- Row 7 (d):** ε, -4, -4, -4, -4, -4, -4, -4. Values: -4, -4, -4, -4, -4, -4, -4.
- Row 8 (e):** ε, -5, -5, -5, -5, -5, -5, -5. Values: -5, -5, -5, -5, -5, -5, -5.
- Row 9 (n):** ε, -6, -6, -6, -6, -6, -6, -6. Values: -6, -6, -6, -6, -6, -6, -6.

Arrows indicate the operations: up (deletion), down (insertion), diagonal (match), and horizontal (substitution). The final cost values are: ε, 1, e, n, d, e, d. The value 1 is highlighted in blue.

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2	-1	-2
a	-3	-3	-1	-1	-2	-2	-2
d	-4	-4	-2	-1	-1	-1	-1
e	-5						
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2	-1	-2
a	-3	-3	-1	-1	-2	-2	-2
d	-4	-4	-2	-2	-2	-1	
e	-5						
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2	-1	-2
a	-3	-3	-1	-1	-2	-2	-2
d	-4	-4	-2	-2	0		
e	-5						
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2	-1	-2
a	-3	-3	-1	-1	-2	-2	-2
d	-4	-4	-2	-2	0	-1	
e	-5						
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2	-1	-2
a	-3	-3	-1	-1	-2	-2	-2
d	-4	-4	-2	-2	0	-1	-1
e	-5						
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2	-1	-2
a	-3	-3	-1	-1	-2	-2	-2
d	-4	-4	-2	-2	0	-1	-1
e	-5	-5	-3				
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2	-3	-4
a	-3	-3	-1	-1	-2	-1	-2
d	-4	-4	-2	-2	-2	-1	-1
e	-5	-5	-3	-3	0	-1	-2
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2	-3	-4
a	-3	-3	-1	-1	-2	-1	-2
d	-4	-4	-2	-2	-2	-2	-1
e	-5	-5	-3	-3	-1	-1	-1
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-3	-4	-5	-6	
e	-2	-1	0	-1	-2	-3	-4
a	-3	-2	-1	0	-1	-2	-3
d	-4	-3	-2	-1	0	-1	-2
e	-5	-4	-3	-2	-1	0	-1
n	-6						

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-3	-4	-5	-6	-7
e	-2	-1	0	-1	-2	-3	-4
a	-3	-2	-1	0	-1	-2	-3
d	-4	-3	-2	-1	0	-1	-2
e	-5	-4	-3	-2	-1	0	-1
n	-6						0

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2	-3	-4
a	-3	-3	-1	-1	-2	-1	-2
d	-4	-4	-2	-2	-2	-2	-1
e	-5	-5	-3	-2	-1	-1	-1
n	-6	-6	i	-3	-2	-1	0

The diagram illustrates the edit operations between the sequences '1' and 'd'. The columns represent the sequence '1' and the rows represent the sequence 'd'. The values in the grid represent the edit distance at each position. Arrows indicate the operations: up (insertion), down (deletion), and right (match/substitution). The path starts at (0,0) and ends at (7,7).

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-3	-4	-5	-6	-7
e	-2	-3	-4	-5	-6	-7	-8
a	-3	-4	-5	-6	-7	-8	-9
d	-4	-5	-6	-7	-8	-9	-10
e	-5	-6	-7	-8	-9	-10	-11
n	-6	-7	-8	-9	-10	-11	-12

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-3	-4	-5	-6	-7
e	-2	-3	-4	-5	-6	-7	-8
a	-3	-4	-5	-6	-7	-8	-9
d	-4	-5	-6	-7	-8	-9	-10
e	-5	-6	-7	-8	-9	-10	-11
n	-6	-7	-8	-9	-10	-11	-12

The grid shows the state transitions for the Global Edit Distance algorithm. The columns represent the string "1" and the rows represent the string "dene". The values in the grid are the edit distances at each step, and arrows indicate the operations: up (insertion), left (deletion), down-left (substitution), and diagonal (match/mismatch).

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-3	-4	-5	-6	-7
e	-2	-3	-4	-5	-6	-7	-8
a	-3	-4	-5	-6	-7	-8	-9
d	-4	-5	-6	-7	-8	-9	-10
e	-5	-6	-7	-8	-9	-10	-11
n	-6	-7	-8	-9	-10	-11	-12

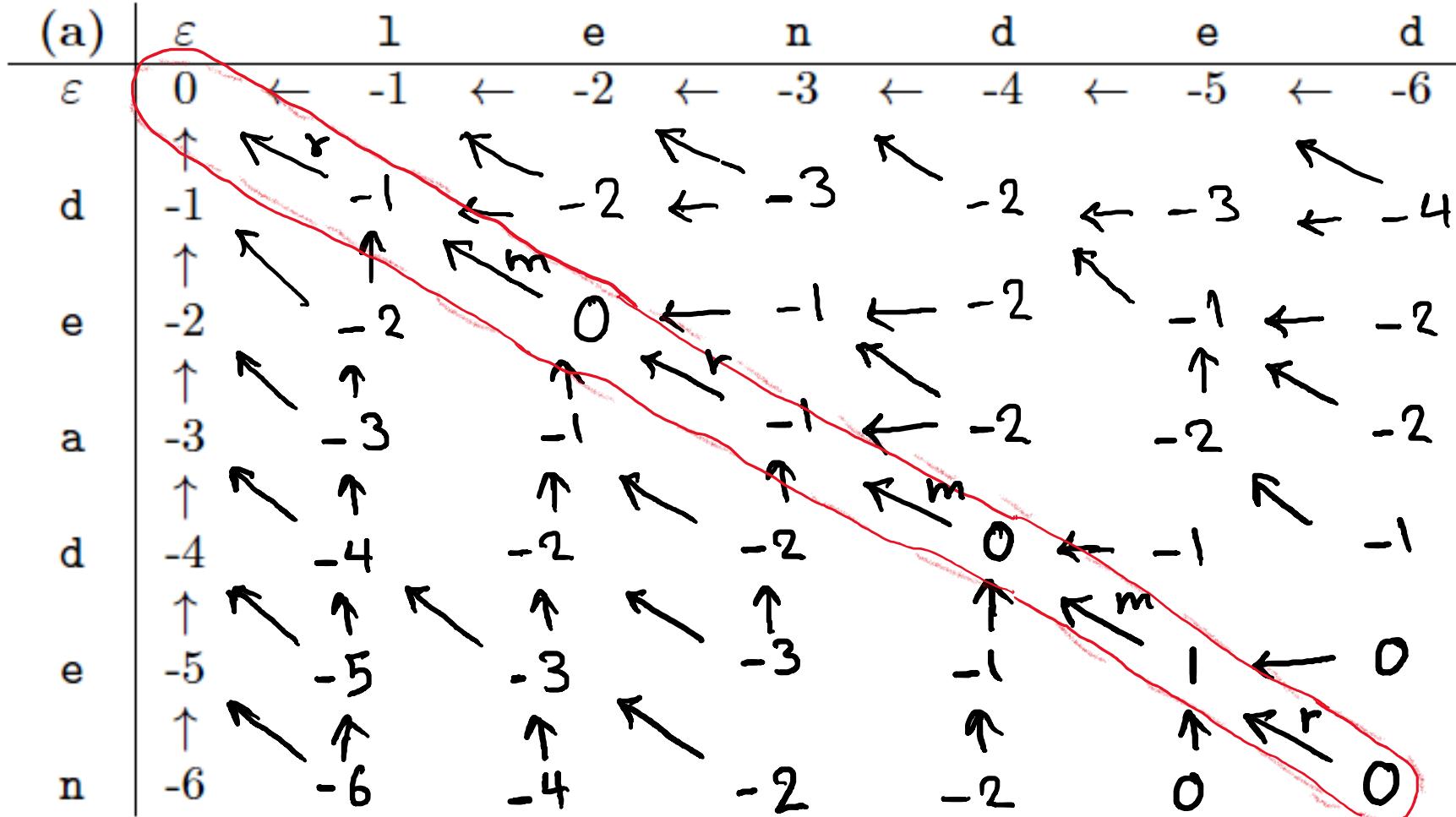
Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-3	-4	-5	-6	-7
e	-2	-3	-4	-5	-6	-7	-8
a	-3	-4	-5	-6	-7	-8	-9
d	-4	-5	-6	-7	-8	-9	-10
e	-5	-6	-7	-8	-9	-10	-11
n	-6	-7	-8	-9	-10	-11	-12

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-2	-3	-4	-5	-6	-7
e	-2	-3	-4	-5	-6	-7	-8
a	-3	-4	-5	-6	-7	-8	-9
d	-4	-5	-6	-7	-8	-9	-10
e	-5	-6	-7	-8	-9	-10	-11
n	-6	-7	-8	-9	-10	-11	-12

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$



Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	1	e	n	d	e	d
ε	0	-1	-2	-3	-4	-5	-6
d	-1	-1	-2	-3	-2	-3	-4
e	-2	-2	0	-1	-2	-1	-2
a	-3	-3	-1	-1	-2	-2	-2
d	-4	-4	-2	-2	0	-1	-1
e	-5	-5	-3	-3	-1	1	0
n	-6	-6	-4	-2	-2	0	0

GEO

Global Edit Distance, $[m, i, d, r] = [+1, -1, -1, -1]$

(a)	ε	ε	1	e	n	d	e	d
ε	0	\leftarrow	-1	\leftarrow	-2	\leftarrow	-3	\leftarrow
d	-1	\uparrow	-1	\leftarrow	-2	\leftarrow	-3	\leftarrow
e	-2	\uparrow	-2	\uparrow	0	\leftarrow	-1	\leftarrow
a	-3	\uparrow	-3	\uparrow	-1	\leftarrow	-2	\leftarrow
d	-4	\uparrow	-4	\uparrow	-2	\uparrow	0	\leftarrow
e	-5	\uparrow	-5	\uparrow	-3	\uparrow	-1	\leftarrow
n	-6	\uparrow	-6	\uparrow	-4	\uparrow	0	r

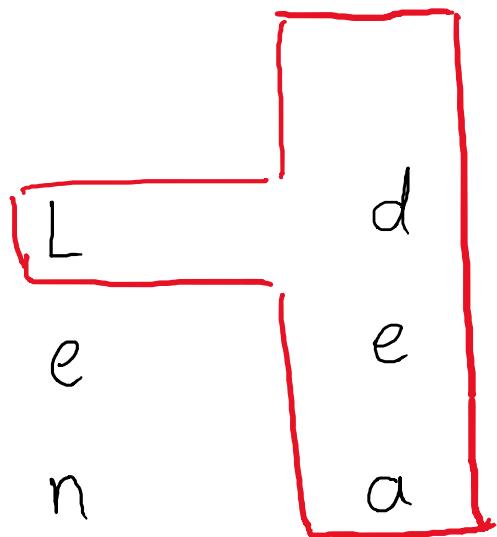
Jaro-Winkler

L	d
e	e
n	a
d	d
e	e
d	n

$$\left\lfloor \frac{\max(|a|, |b|)}{2} - 1 \right\rfloor$$

$$= \left\lfloor \frac{6}{2} - 1 \right\rfloor = 2$$

Jaro-Winkler

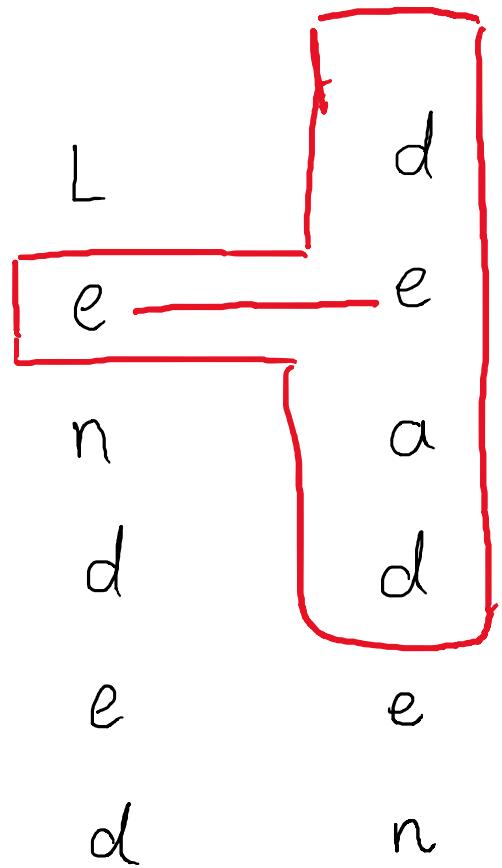


d d

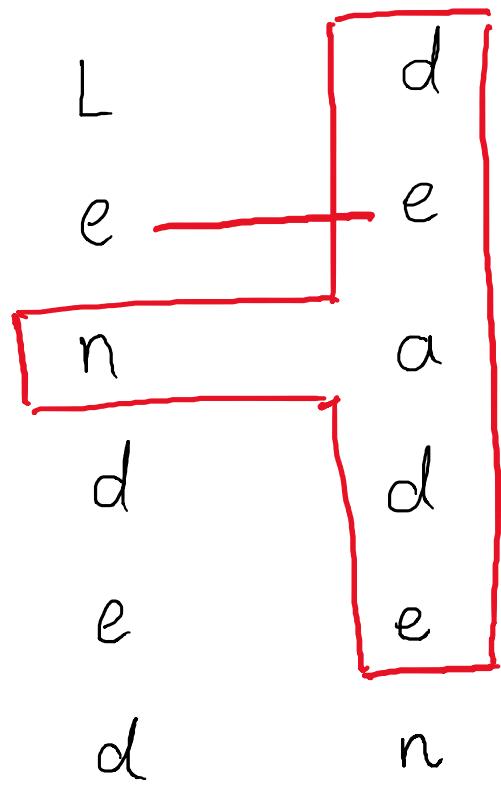
e e

d n

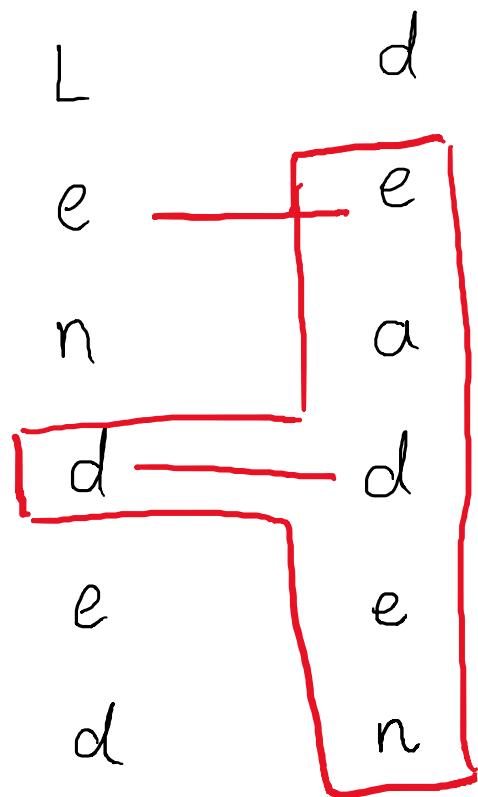
Jaro-Winkler



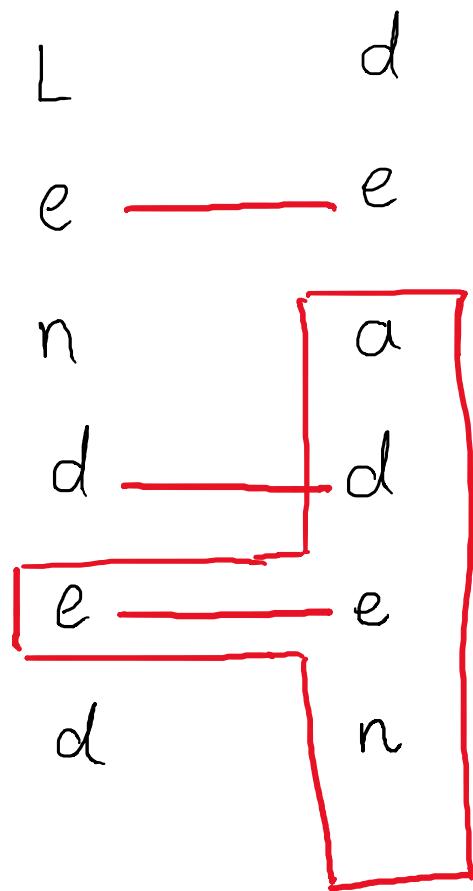
Jaro-Winkler



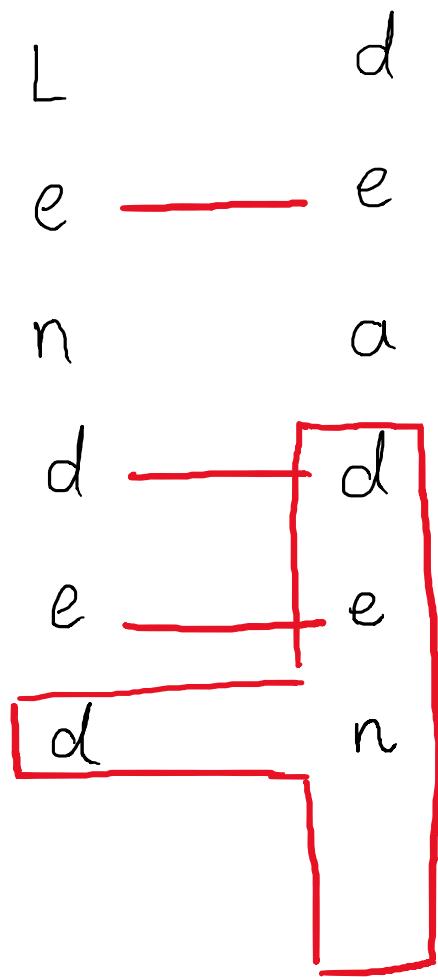
Jaro-Winkler



Jaro-Winkler



Jaro-Winkler



Jaro-Winkler

L d

e — e

n a

d — d

e — e

d n

m = 3

t = 0

Jaro-Winkler

$$Sim_j(a, b) = \frac{1}{3} \left(\frac{m}{|a|} + \frac{m}{|b|} + \frac{m - t}{m} \right)$$

$$Sim_j('deaden', 'lended') = \frac{1}{3} \left(\frac{3}{6} + \frac{3}{6} + \frac{3 - 0}{3} \right) = \frac{2}{3} \cong 0.67$$

Jaro-Winkler

$$Sim_w(a, b) = Sim_j(a, b) + \ell \times p \times (1 - Sim_j(a, b))$$

$$Sim_w('deaden', 'lended') = 0.67 + 0 \times 0.1 \times (1 - 0.67) = 0.67$$

Jaro-Winkler

L L

e e
n n
d t

e

d

$$\left\lfloor \frac{\max(|a|, |b|)}{2} - 1 \right\rfloor$$

$$= \left\lfloor \frac{6}{2} - 1 \right\rfloor = 2$$

Jaro-Winkler

L — L

$$m = 3$$

e — e

$$t = 0$$

n — n

d t

e

d

Jaro-Winkler

L — L

e — e

n — n

d t

e

d

$$\text{sim}_j(a, b) = \frac{1}{3} \left(\frac{3}{6} + \frac{3}{4} + \frac{3}{3} \right)$$

$$= 0.75$$

Jaro-Winkler

L — L

e — e

n — n

d +

e

d

$$\text{sim}_j(a, b) = \frac{1}{3} \left(\frac{3}{6} + \frac{3}{4} + \frac{3}{3} \right) = 0.75$$

$$l = 3 \quad p = 0.1$$

Jaro-Winkler

L — L

e — e

n — n

d +

e

d

$$\text{Sim}_J(a, b) = \frac{1}{3} \left(\frac{3}{6} + \frac{3}{4} + \frac{3}{3} \right) = 0.75$$

$$l = 3 \quad p = 0.1$$

$$\text{Sim}_\omega(a, b) = 0.75 + 3 \times 0.1 \times (1 - 0.75)$$

$$= 0.825$$