

Info 4300: Language and Information - Lecture 2  
Minimum edit distance worksheet.

Name:

Netid:

Target

	#	K	A	R	D	A	S	H	I	A	N	S
#												
D												
A												
L												
M												
A												
T												
I												
A												
N												
S												

Source

Name:  
Netid:

Sketch of the Wagner Fisher algorithm for detecting minimum edit distance; we consider a cost of 1 for insertions, 1 for deletions, and a cost of 2 for substitutions.

$D(i,j)$ ="The cost of transforming the first  $i$  letters of the Source in the first  $j$  letters of the Target"

*Initializations:*

$$D(0,0)$$

for  $i=0, \text{len}(\text{Source})$ :

$$D(i,0)=i$$

for  $i=0, \text{len}(\text{Target})$

$$D(0,j)=j$$

for  $i=0, \text{len}(\text{Source})$ :

for  $j=0, \text{len}(\text{Target})$ :

$$D(i,j) = \min \begin{cases} D(i-1,j) + 1 \\ D(i,j-1) + 1 \\ D(i-1,j-1) + \begin{cases} 2; & \text{if } \text{Target}(i) \neq \text{Source}(j) \\ 0; & \text{if } \text{Target}(i) = \text{Source}(j) \end{cases} \end{cases}$$

**Return**  $D(\text{len}(\text{Source}), \text{len}(\text{Target}))$