

# IR (Chapter 03) Classwork Solution

Dell Zhang  
Birkbeck, University of London

1. (a) *steel* and *seal* have a Levenshtein distance of 2:

	""	s	e	a	l
""	0	1	2	3	4
s	1	0	1	2	3
t	2	1	1	2	3
e	3	2	1	2	3
e	4	3	2	2	3
l	5	4	3	3	<b>2</b>

- (b)  $A :=$  the set of 2-grams for *steel*: { \$s, st, te, ee, el, l\$ }  
 $B :=$  the set of 2-grams for *seal*: { \$s, se, ea, al, l\$ }

$$A \cap B = \{ \$s, l\$ \} \Rightarrow |A \cap B| = 2$$

$$A \cup B = \{ \$s, st, te, ee, el, l$, se, ea, al \} \Rightarrow |A \cup B| = 9$$

$$\text{Jaccard coefficient} = \frac{2}{9}$$