1.

s_1 = The man saw a car in the park

s_2 = 3 saw the man park the car

	The	3	man	saw	a	car	park	in
s_1	2	0	1	1	1	1	1	1
s_2	2	1	1	1	0	1	1	0

a/ euclidean

$$\sqrt{\sum (s_{1i} - s_{2i})^2} = \sqrt{0^2 + (-1)^2 + 0 + 0 + 1 + 0 + 0 + 1} = \sqrt{3}$$

$$\text{sim}(s_1, s_2) = \frac{1}{1 + d(s_1, s_2)} = \frac{1}{1+\sqrt{3}} \simeq 0.366$$

b/ Vector cosine:

$$\text{sim}_{\cos}(s_1, s_2) = \frac{\sum s_{1i}\, s_{2i}}{\sqrt{\sum s_{1i}^2} \cdot \sqrt{\sum s_{2i}^2}} = \frac{4+0+1+1+0+1+1+0}{\sqrt{10} \cdot \sqrt{9}} = \frac{8}{3\sqrt{10}}$$

$$\simeq 0.843$$

c/ Jaccard

$$\text{sim}_{jac}(s_1, s_2) = \frac{a}{a+b+c}$$

Feature set:

	The	3	man	saw	a	car	park	in
s_1	1	0	1	1	1	1	1	1
s_2	1	1	1	1	0	1	1	0

$a = 5$

$b = 2$

$c = 1$

$$\text{sim}_{jac}(s_1, s_2) = \frac{5}{8} \simeq 0.625$$

d/ Overlap

$$\text{sim}_{overlap}(s_1, s_2) = \frac{a}{\min(a+b, a+c)} = \frac{5}{\min(7, 6)} = \frac{5}{6} \simeq 0.833$$