Assume we have the same database of documents as in slides, but the indexing vocabulary changes: K1= {garden, pond, spring, fall} K2 = {garden, pond, spring}

How the representation of documents changes?

K1= {garden, pond, spring, fall}
D1 = {1,0,1,1}
D2 = {1,0,1,0}
D3 = {1,0,0,1}
D4 = {1,0,0,0}
D5 = {1,0,0,1}

K2 = {garden, pond, spring}
D1 = {1,0,1}
D2 = {1,0,1}
D3 = {1,0,0}
D4 = {1,0,0}
D5 = {1,0,0}

Assuming the same query as in the slides How does the query representation change? How do the results of retrieval change?

The changed query for k1 will be
(G ∩S) ∪(G ∩~F)
((1,?,?,?) ∩ (?,?,1,?)) ∪ ((1,?,?,?)∩(?,?,?,0))
= (1,?,1,?) ∪(1,?,?,0)
= **(1,0,1,1)** ∪ (1,1,1,0) ∪ (1,1,1,1) ∪**(1,0,1,0)** ∪**(1,0,0,0)** ∪(1,1,0,0) ∪(1,1,1,0) ∪(1,0,1,0)
Document: D5, D1,D2

The changed query for k2 will be
(G ∩S) ∪(G ∩~F)
= ((1,?,?) ∩ (?,?,1)) ∪ ((1,?,?)∩ (?,?,?))
=(1,?,1) ∪ (1, ?, ?)
=(1,0,1) (1,1,1) **(1,0,0)** (1,1,0) (1,0,1) (1,1,1)
Document- D1, D2,D3,D4,D5

Justification:

The change (size, order, items) in indexing vocabulary(k1) and absence of document word in vocabulary (k2) changed the representation of the document, query and consequently it changes the result for both cases