1. Are the following statements true or false?

1. Stemming increases the size of the vocabulary.

2. Stemming should be invoked at indexing time but not while processing a query.

3. Skip pointers are not useful for queries of the form \(x \text{ OR } y\)

2. Phrase search

Given is a part of a positional index in the form ‘\text{term: docId1: posList1; docId2: posList2; ...}’: 

- angels: 2: (36,174,252,651); 4: (12,22,102,432); 7: (17); 
- fools: 2: (1,17,74,222); 4: (8,78,108,458); 7: (3,13,23,193); 
- fear: 2: (87,704,722,901); 4: (13,43,113,433); 7: (18,328,528); 
- in: 2: (3,37,76,444,851); 4: (10,20,110,470,500); 7: (5,15,25,195); 
- rush: 2: (47,86,234,999); 4: (14,24,774,944); 7: (199,319,599,709); 
- to: 2: (47,86,234,999); 4: (14,24,774,944); 7: (20,320); 
- tread: 2: (57,94,333); 4: (15,35,155); 7: (20,320); 

1. Which documents (if any) are relevant to the following phrase queries?
   a. ‘fools rush in’
   b. ‘fools rush in’ AND ‘angels fear to tread’

2. Create a part of the bi-word index that is required to compute search results for the second query.

3. Will using a bi-word index to compute results for a phrase query return the same documents as a positional index? How do both compare in terms of space requirements?

4. Assume the terms ‘in’ and ‘to’ are filtered out as stop words. Find a way to combine a positional index with stop word filtering in a search system. Which documents are returned for the above queries?

3. Skip-Lists

Given is the query ‘\text{ice AND age}’ together with the posting lists of both terms:

- ice: (4, 6, 10, 12, 14, 16, 18, 20, 22, 32, 47, 81, 120, 122, 157, 180) 
- age: (47)
How many posting elements and/or pointers need to be compared in the posting list intersection for answering the query, using:
1. Traditional posting lists
2. Posting lists with skip pointers
?

Explain your answer!