

## WST540: Quiz 2

**Problem 1.** Suppose that we have the following document collection:

document ID	words
1	a b a a c a
2	b b c c
3	a d c a d b
4	e

Give the document-level inverted lists of all the words a, b, c, d, and e. Each entry of an inverted list should have the format (doc id, term frequency).

**Solution.**

term $w$	inverted list for $w$
a	(1, 4), (3, 2)
b	(1, 1), (2, 2), (3, 1)
c	(1, 1), (2, 2), (3, 1)
d	(3, 2)
e	(4, 1)

**Problem 2.** Give the Elias' gamma and delta codes of 37.

**Solution.**  $\text{gamma}(37) = 11111000101$ , and  $\text{delta}(37) = 1101000101$ .

**Problem 3.** Decompress the Elias' gamma code 111111111100000001101 into an integer.

**Solution.**  $1024 + 13 = 1037$ .

**Problem 4.** Decompress the Elias' delta code 111000100000110 into an integer.

**Solution.** The gamma code 1110001 equals 9. Hence, the answer is  $2^8 + 6 = 256 + 6 = 262$ .

**Problem 5.** Let  $S = \{\text{aabb}\perp, \text{aab}\perp, \text{abb}\perp, \text{abbba}\perp\}$ , where the strings have ids 1, 2, ..., 4 (from left to right), respectively. Give the Patricia trie for  $S$ . For each internal node, indicate (i) its positional index, and (ii) the id of its representative. For each edge, indicate the character that it carries.

**Solution.**

