1 (a)

S = database$

BWT(S) = etbda$saa

1 (b)

For this Qn, the initial values of the alphabets are less important. After MTF, the last two a's should be 2, 0.

i.e.,

101, 116, …, **2, 0**

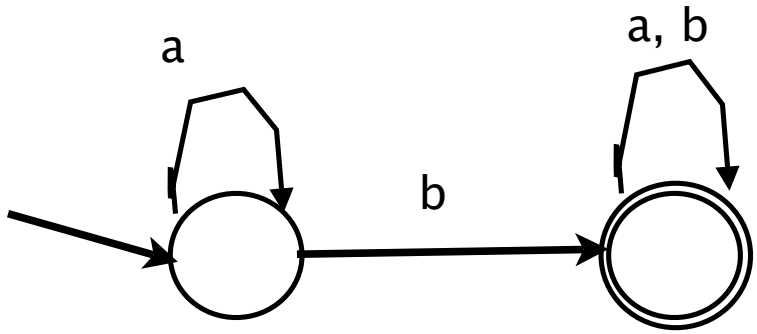
1 (c)

You can work it out using either F & L column, or C[] and Occ.

Final answer is **barbara$**

2 (a) & (b)

(a? bab)*



3

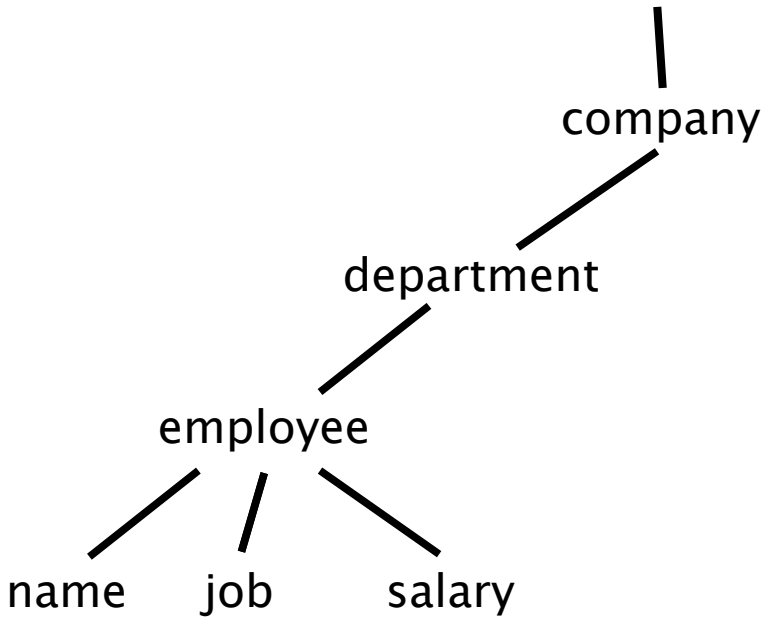
- /bib//book[year > 2000]/author always produces a subset of the result from /bib[*//year > 2000]//book/author
- As the latter can match more general cases, for example /bib//last-updated/year may contain the last update date of the entire bib database - which includes all the books.

4(a)

The idea:

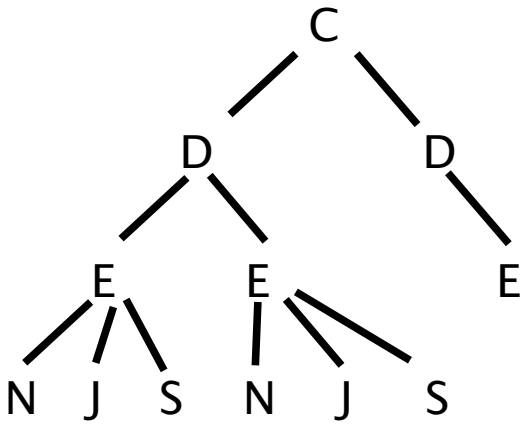
Your answer should consider the encoding of the XML tree into the balanced parentheses / bits:
((((())(())(()))((()))(())(()))(()))
You will then need to assume these parentheses are stored in blocks and explain how multi-tiered architecture as shown in ISX can efficiently compute some navigational operations.

4(b)



5

Use an example to illustrate how to match a sub-path on the following XBW (for simplicity, the text values are not shown here):

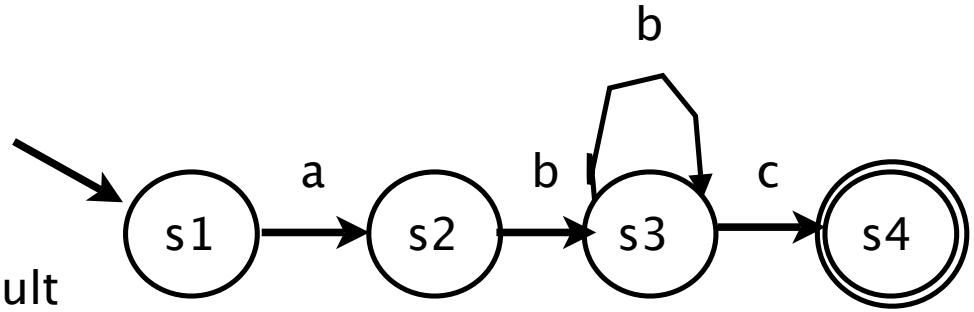


1 C	1 C	1 C
0 D C	0 D C	0 D
0 E DC	1 D C	1 D
0 N EDC	0 E DC	0 E
0 J EDC	1 E DC	1 E
1 S EDC →	1 E DC →	1 E
1 E DC	0 N EDC	0 N
0 N EDC	0 J EDC	0 J
0 J EDC	1 S EDC	1 S
1 S EDC	0 N EDC	0 N
1 D C	0 J EDC	0 J
1 E DC	1 S EDC	1 S

XBW

6

Start	Stop
Site1	
(x1,s1)	(y1,s3)
(x4,s3)	(y3,s4)
Site2	
(y1,s2)	(z2,s3)
(y1,s3)	(z2,s3)
(y3,s2)	(z2,s3)
(y3,s3)	(z2,s3)
Site3	
(z2,s2)	(x4,s3)
(z2,s3)	(x4,s3)



Start	Result
Site1	
(x1,s4)	x1
(x4,s4)	x4
(x4,s3)	x3
Site2	
(y1,s4)	y1
(y1,s2)	y3
(y1,s3)	y3
(y3,s4)	y3
Site3	
(z2,s1)	z3
(z2,s4)	z2
(z2,s2)	z2
(z2,s3)	z2

By tracing these two tables from (x1, s1), we obtain the final results:
y3, z2, x3, y3