CSE 425

Quiz 02

Nowshin Tasnim

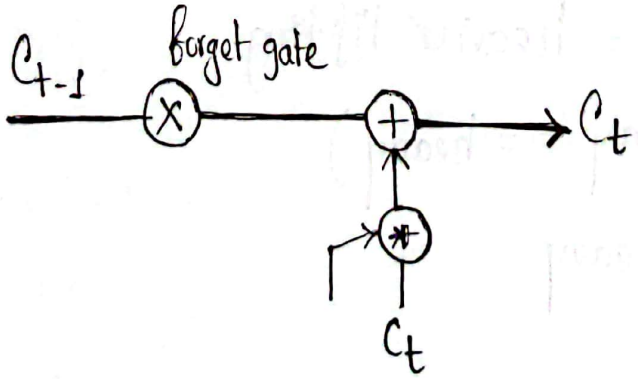
19101655

Sec: 02

Question 1

Vanishing Gradient Problem:

The term vanishing gradient problem is encountered when training artificial neural networks with gradient-based learning methods and backpropagation. The problem is that in some cases, the gradient will be vanishingly small, effectively preventing the weight from changing its value. Here the further you go the lower the gradient is and the harder it is to train the weights which has a domino effect on all the further weights throughout the network. That was the main roadblock to using recurrent neural networks It occurs when large data sequence appear. Also, when there is too old memories weather to remember it or not. So to solve the issue here we use sigmoid function



In the figure forget gate controlls how much data to be remembered. The sigmoid layer output number between 0 to 1. Which determines how much end component should go through. If is 0 then all going to be forgotten. If it is 1 all data will pass.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

where C_t handles vanishing gradient problem.

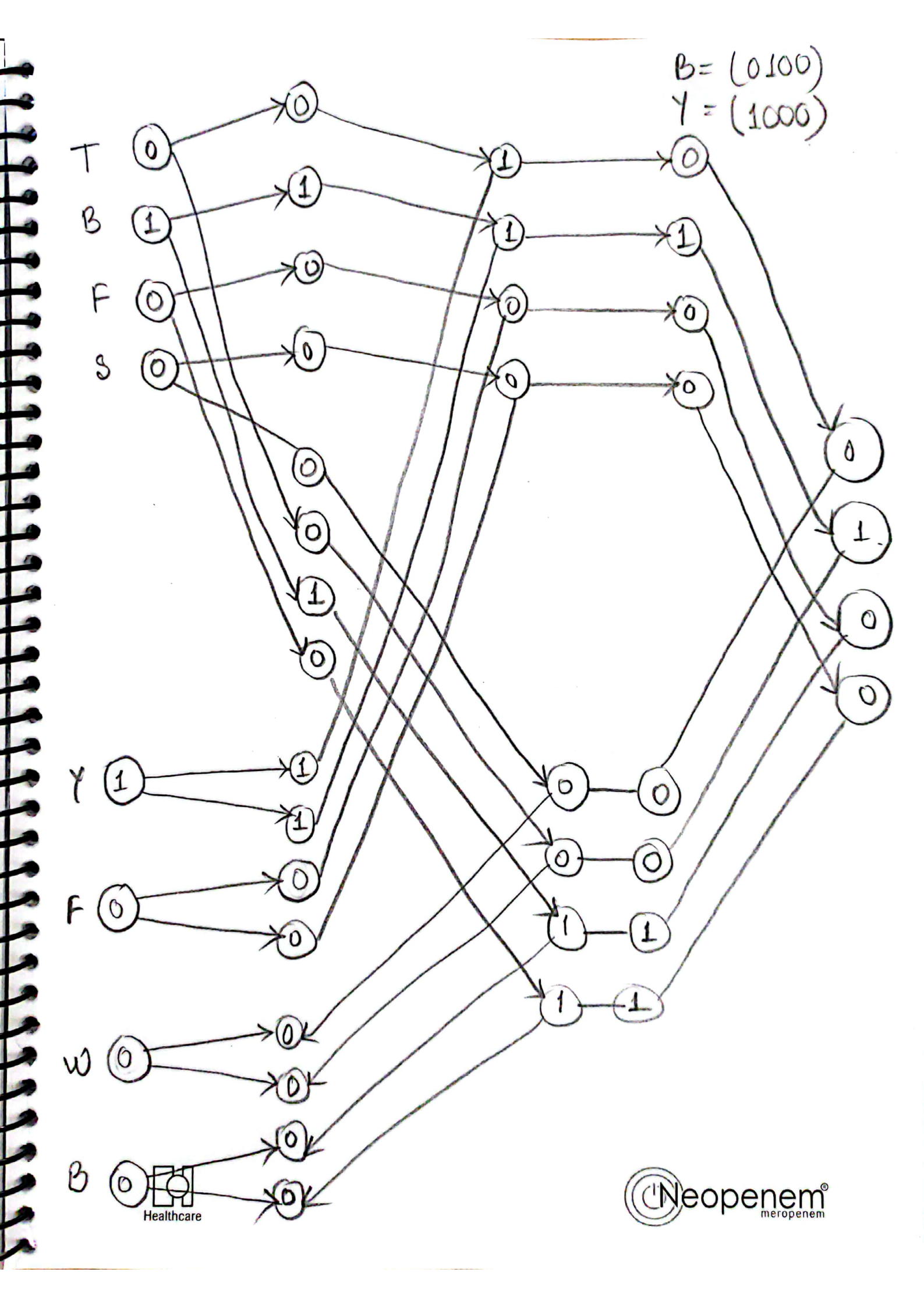
Question 2

Equipment = (T, B, F, S)
most heavy — least heavy

Color = (Y, R, W, B)
heavy

brighter dress = heavier lifting

if (yesterday == heavy)
next day ≠ heavy


B= (0100)
Y = (1000)
T
B
F
S
Y
F
W
B
Healthcare
Neopenem
meropenem