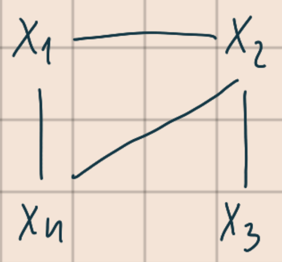
Exercise 1: Given stochastic variables X_1, X_2, X_3, X_4 and the following Markov Network with factors Φ_i



The joint probability distribution $P(X_1, X_2, X_3, X_4)$ can be factorized as:

a. $\frac{1}{Z} \Phi_1(X_1, X_2) \Phi_2(X_2, X_3) \phi_3(X_2, X_4)$

b. $\frac{1}{Z} \Phi_1(X_1, X_2) \phi_2(X_1, X_4) \phi_3(X_2, X_3)$

c. $\frac{1}{Z} \Phi_1(X_2, X_4) \phi_2(X_4, X_1) \Phi_3(X_1, X_2)$

d. $\frac{1}{Z} \phi_1(X_1, X_2, X_4) \Phi_2(X_2, X_3)$ ✓

e. $\frac{1}{Z} \phi_1(X_2, X_4, X_3) \phi_2(X_1, X_4) \Phi_3(X_2, X_4)$

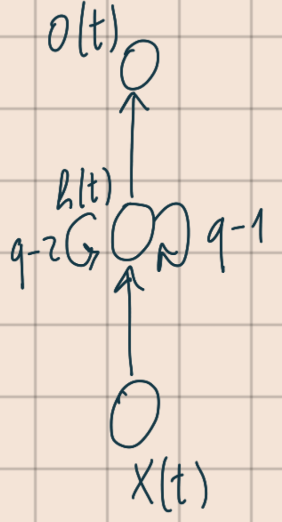
SOLUTION: Consider just MAXIMAL cliques, i.e., cliques from which you CANNOT obtain other cliques by adding one or more vertices in the graph.

In our case: $\{X_2, X_3\}$ is MAXIMAL.
$\{X_1, X_2\}$ is NOT MAXIMAL because if you add X_4 you obtain another clique. Same reasoning for $\{X_2, X_4\}$ and $\{X_1, X_4\}$.

Instead, $\{X_1, X_2, X_4\}$ is MAXIMAL.

$$\Rightarrow P(X_1, X_2, X_3, X_4) = \frac{1}{Z} \Phi_1(X_1, X_2, X_4) \cdot \Phi_2(X_2, X_3).$$

Exercise 2: Consider an RNN with the following structure:



Given that:

- THE INPUT SEQUENCE HAS 10 TERMS;
- THE TRANSITION FUNCTION IS IMPLEMENTED BY A FULLY CONNECTED FEEDFORWARD NETWORK WITH ONE HIDDEN LAYER OF SIZE 10 AND AN OUTPUT LAYER OF SIZE 5;
- THE OUTPUT FUNCTION IS IMPLEMENTED BY A FULLY CONNECTED FEEFORWARD NETWORK WITH ONE HIDDEN LAYER OF SIZE 20 AND AN OUTPUT LAYER OF SIZE 3.

What is the total # of PARAMETERS? Considering bias terms for each neuron

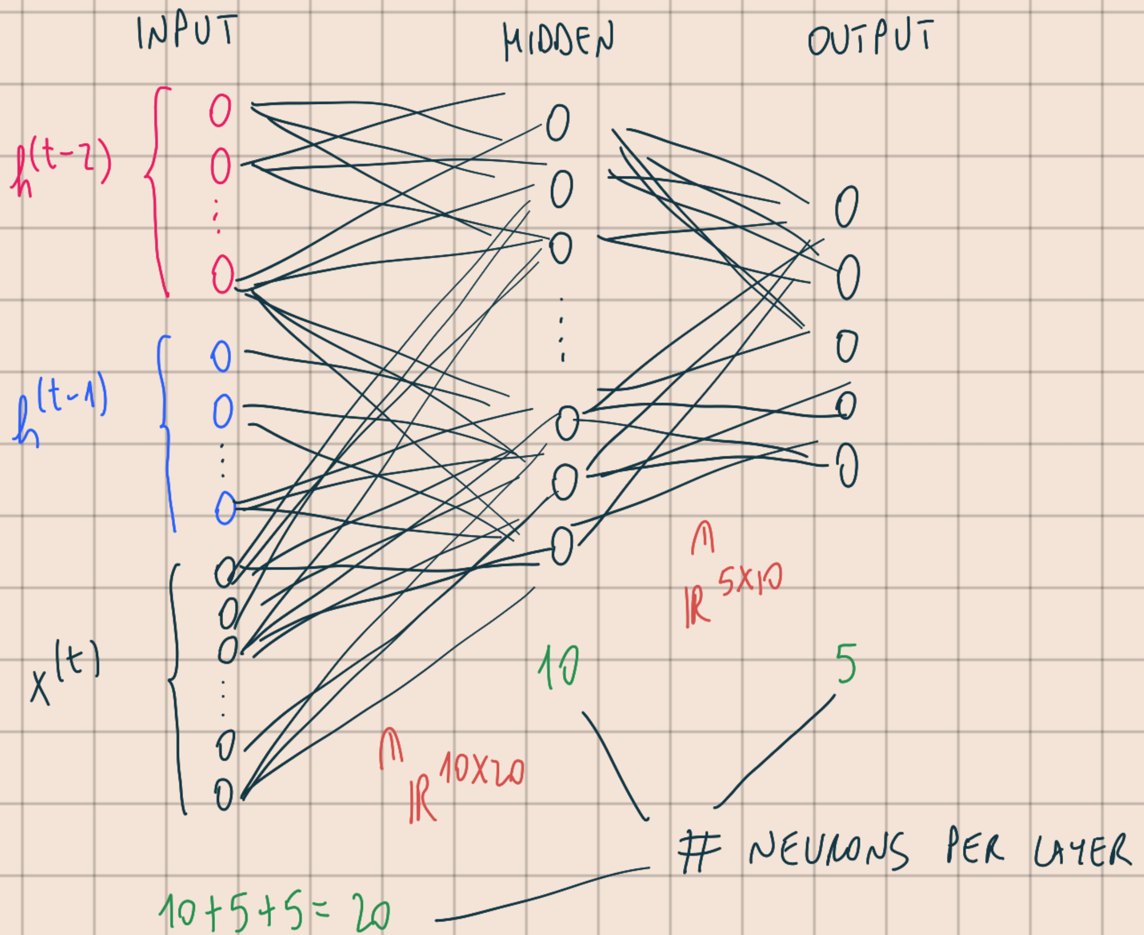
SOLUTION:

$x(t) \in \mathbb{R}^{10}$, $h(t) \in \mathbb{R}^5$, $o(t) \in \mathbb{R}^3$

The TRANSITION FUNCTION is:

$$h(t) = \underset{\substack{\uparrow \\ f \text{ is implemented by a FFN}}}{f}\,(x(t);\, h(t-1);\, h(t-2))$$

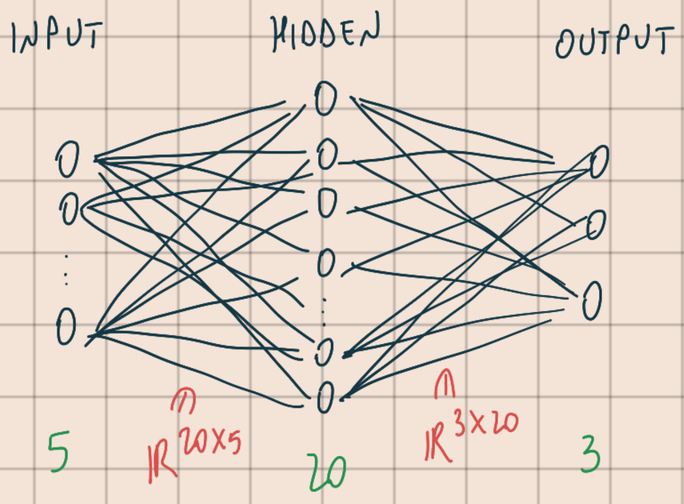
① NETWORK IMPLEMENTING THE TRANSITION



- \# PARAMS INPUT → HIDDEN: $10\times 20 = 200 + \underset{\substack{\uparrow \\ \text{BIAS}}}{10} = 210$
- \# PARAMS HIDDEN → OUTPUT: $5\times 10 = 50 + \underset{\substack{\uparrow \\ \text{BIAS}}}{5} = 55$

In total for the FIRST FFN we have $210+55=265$ PARAMS.

② NETWORK IMPLEMENTING THE OUTPUT FUNCTION



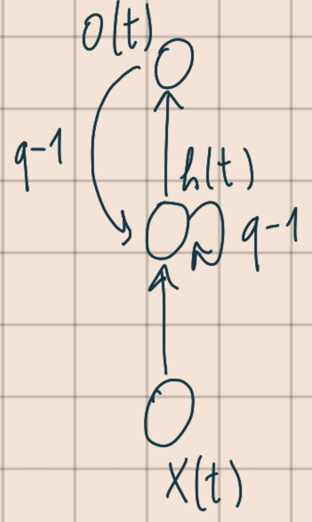
- # PARAMS INPUT → HIDDEN: $5 \times 20 = 100 + \underset{\substack{\uparrow \\ \text{BIAS}}}{20} = 120$

- # PARAMS HIDDEN → OUTPUT: $3 \times 20 = 60 + \underset{\substack{\uparrow \\ \text{BIAS}}}{3} = 63$

In total for the 2^{nd} FFN we have $120 + 63 = 183$ PARAMS.

So, in total, we have $265 + 183 = 448$ PARAMS.

Exercise 3: Consider an RNN with the following structure:



Given that:

- THE INPUT SEQUENCE HAS 10 TERMS;
- THE TRANSITION FUNCTION IS IMPLEMENTED BY A FULLY CONNECTED FEEDFORWARD NETWORK WITH ONE HIDDEN LAYER OF SIZE 15 AND AN OUTPUT LAYER OF SIZE 5;
- THE OUTPUT FUNCTION IS IMPLEMENTED BY A FULLY CONNECTED FEEFORWARD NETWORK WITH ONE HIDDEN LAYER OF SIZE 10 AND AN OUTPUT LAYER OF SIZE 2.

What is the total # of PARAMETERS? Considering bias terms for each neuron

SOLUTION:

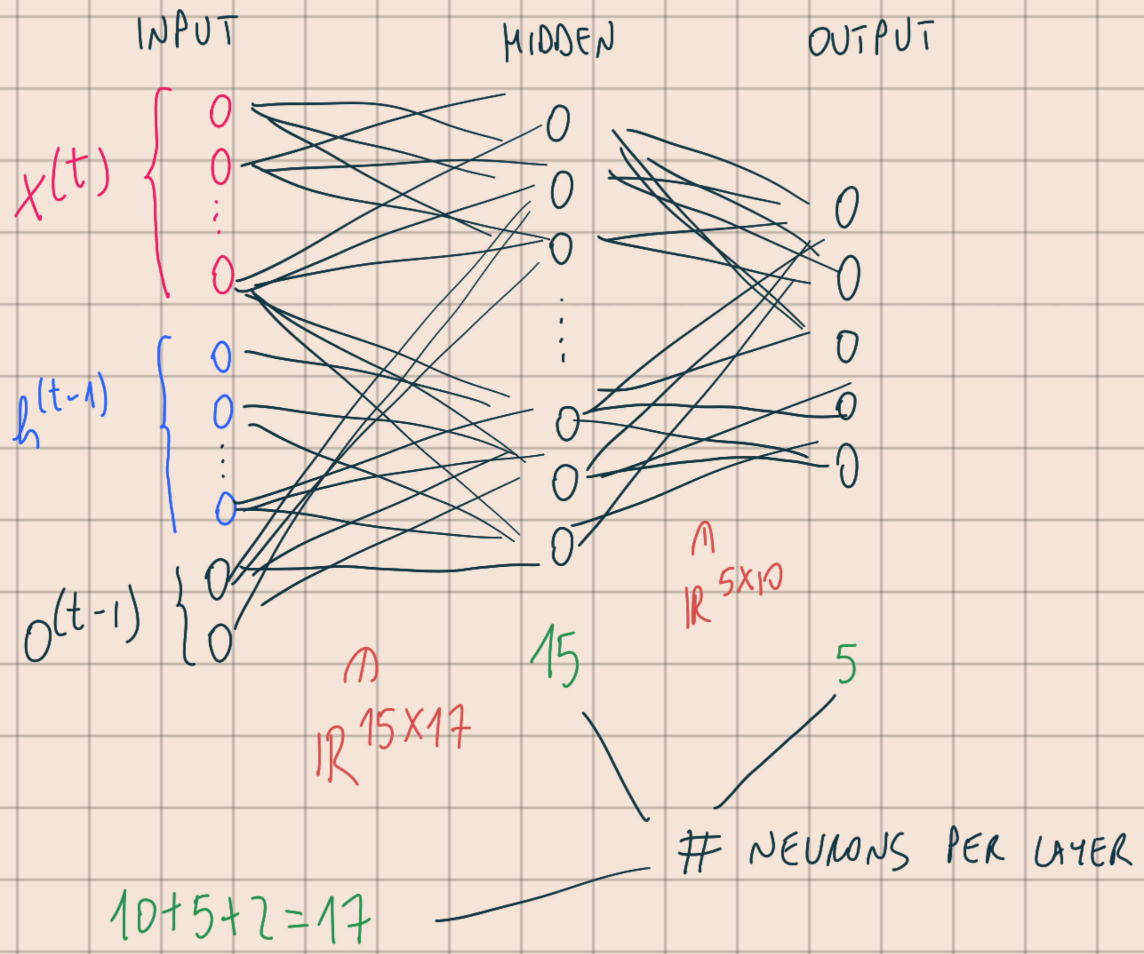
$x(t) \in \mathbb{R}^{10}$, $h(t) \in \mathbb{R}^{5}$, $o(t) \in \mathbb{R}^{2}$

The TRANSITION FUNCTION in this case is:

$$h(t) = \underset{\uparrow}{f}\,(x(t); h(t-1); o(t-1))$$

f is implemented by a FFN

① NETWORK IMPLEMENTING THE TRANSITION

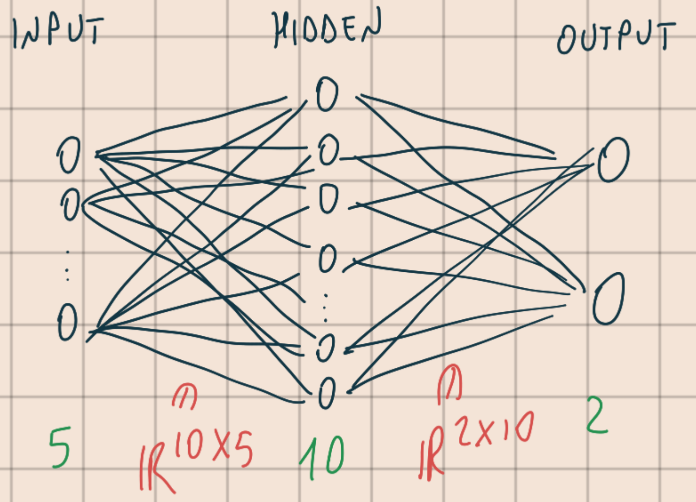


- # PARAMS INPUT → HIDDEN: $15 \times 17 = 255 + \underset{\substack{\uparrow \\ \text{BIAS}}}{15} = 270$

- # PARAMS HIDDEN → OUTPUT: $5 \times 15 = 75 + \underset{\substack{\uparrow \\ \text{BIAS}}}{5} = 80$

In total for the FIRST FFN we have $270 + 80 = 450$ PARAMS.

② NETWORK IMPLEMENTING THE OUTPUT FUNCTION



- # PARAMS INPUT → HIDDEN : $10 \times 5 = 50 + \underset{\substack{\uparrow \\ \text{BIAS}}}{10} = 60$

- # PARAMS HIDDEN → OUTPUT : $2 \times 10 = 20 + \underset{\substack{\uparrow \\ \text{BIAS}}}{2} = 22$

In total for the 2^{nd} FFN we have $60+22=82$ PARAMS.

So, in total, we have $450+82=532$ PARAMS.