Parameters: \( \mathbf{W}_y, \mathbf{W}_{xh}, \mathbf{W}_{hh} \)

1. Complete the following:

\[
\begin{align*}
\mathbf{h}_1 &= \quad \mathbf{h}_2 &= \\
\hat{\mathbf{y}}_1 &= \quad \hat{\mathbf{y}}_2 &=
\end{align*}
\]

2. Write down the training set using one-hot encoding if the input sequence is \( \text{a b b b c} \).
3. Complete the following:

\[ P(X_3 = "b") = \hat{y}_3 \]

4. The cost associated with \( \hat{y}_2 \) is

(N.B. the training sequence is abbbcb)

5. The total cost is the sum of the costs associated with \( \hat{y}_0, \hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n \). Write it down