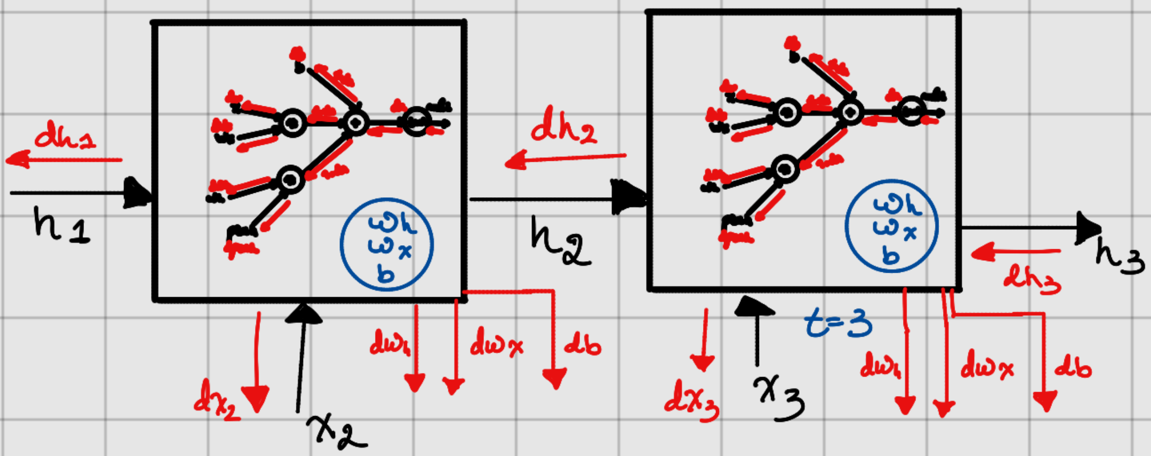

since we've already made a rnn_step_backwards, we can visualize each RNN cell as a blackbox with: - an internal state of $\underbrace{W_x, W_h, b}_{\text{shared between ALL cells}}$

- inputs: h_{t-1}, x_t
- outputs: h_t

. since the internal state is shared between all RNNs,

- $dW_x = \sum_{t=1}^{T} dW_x(t)$, where $dW_x(t) = x_{(t)}^T \cdot da_{(t)}$ (see eq 3 in rnn_step_backprop) transpose
- $dW_h = \sum_{t=1}^{T} dW_h(t)$, where $dW_h(t) = prev\_h_{(t)}^T da_{(t)}$, see eq 5
- $db = \sum_{t=1}^{T} db(t)$, $db(t) = \sum_{n=1}^{N} da_{(t)}$, eq 6

or, if we were to model this as a for loop,

$$\boxed{\begin{array}{l} \bullet\ dW_x(t) \mathrel{+}= dW_x(t-1) \\ \bullet\ dW_h(t) \mathrel{+}= dW_h(t-1) \\ \bullet\ db(t) \mathrel{+}= db(t-1) \end{array}}$$

- using same logic as before,
  - $dh_t = dh + prev\_h$, where $prev\_h(t) = da \cdot W_h^T$, eq 4

- $dx_{(t)}$ would simply be the same as that of rnn_step_back

  $\therefore dx_{(t)} = da \cdot W_x^T$, eq 2

all dimensions are the same as RNN step back

in summary, add all the outputs from rnn_step_back except for dx
(since $x_t \neq x_{t-1}$)