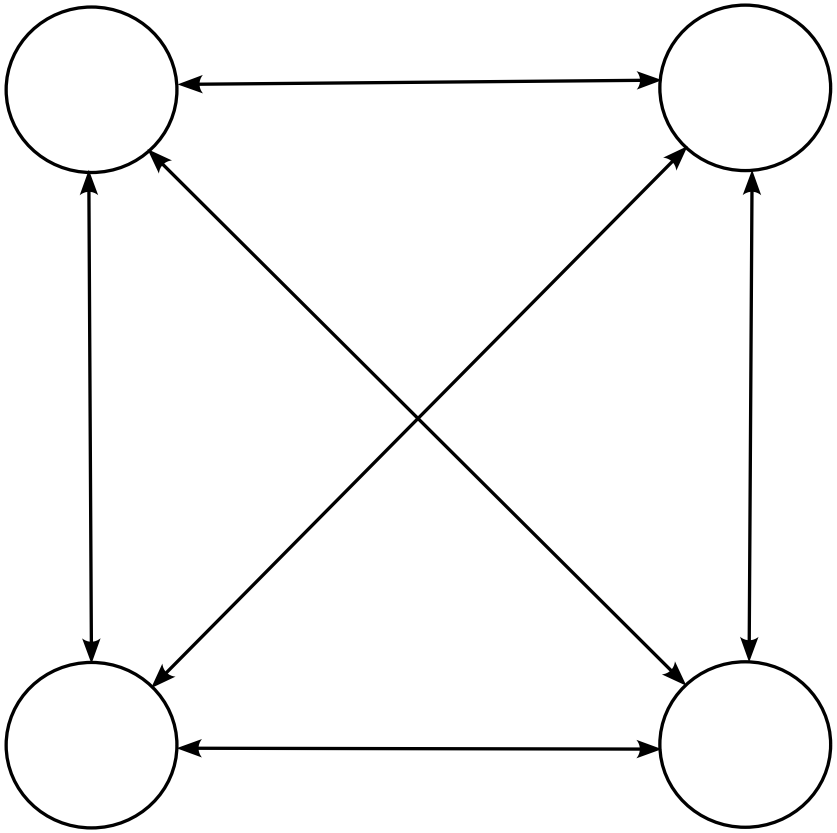
5. Sequence modeling with recurrent neural networks

5.1. Structure of the Elman Recursive Neural Network

Manel Martínez-Ramón
Meenu Ajith
Aswathy Rajendra Kurup

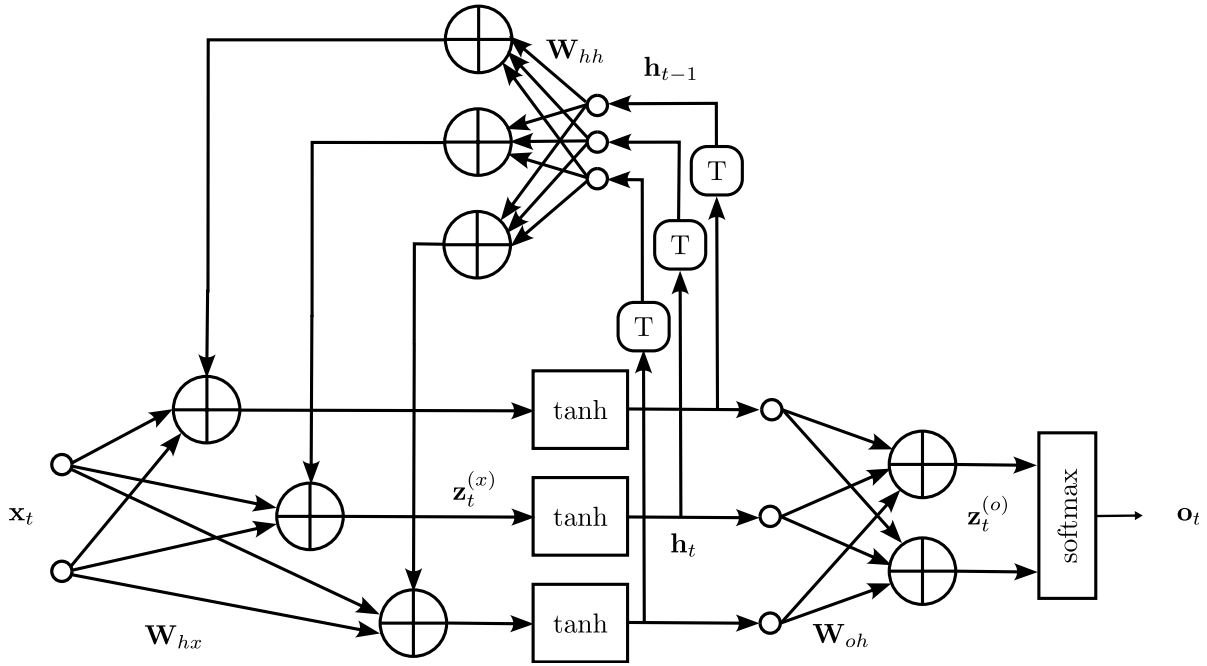
The Hopfield network



- Introduced in 1982 by John Hopfield.
- Structure previously introduced by W. A. Little in 1974.
- First to include the idea of recurrence.
- Used in optimization as the classic travelling salesman problem.
- Continuous state versions exist.

Structure of the Elman RNN

State equations



$$
\begin{aligned}
\mathbf{z}_t^{(x)} &= \mathbf{W}_{hx}^\top \mathbf{x}_t + \mathbf{W}_{hh}^\top \mathbf{h}_{t-1} + \mathbf{b}_h \\
\mathbf{z}_t^{(o)} &= \mathbf{W}_{oh}^\top \mathbf{h}_t + \mathbf{b}_o \\
\mathbf{h}_t &= \tanh(\mathbf{z}_t^{(x)}) \\
\mathbf{o}_t &= \mathbf{o}(\mathbf{z}_t^{(o)}) = \mathrm{softmax}(\mathbf{z}_t^{(o)})
\end{aligned}
\tag{1}
$$

Structure of the Elman RNN

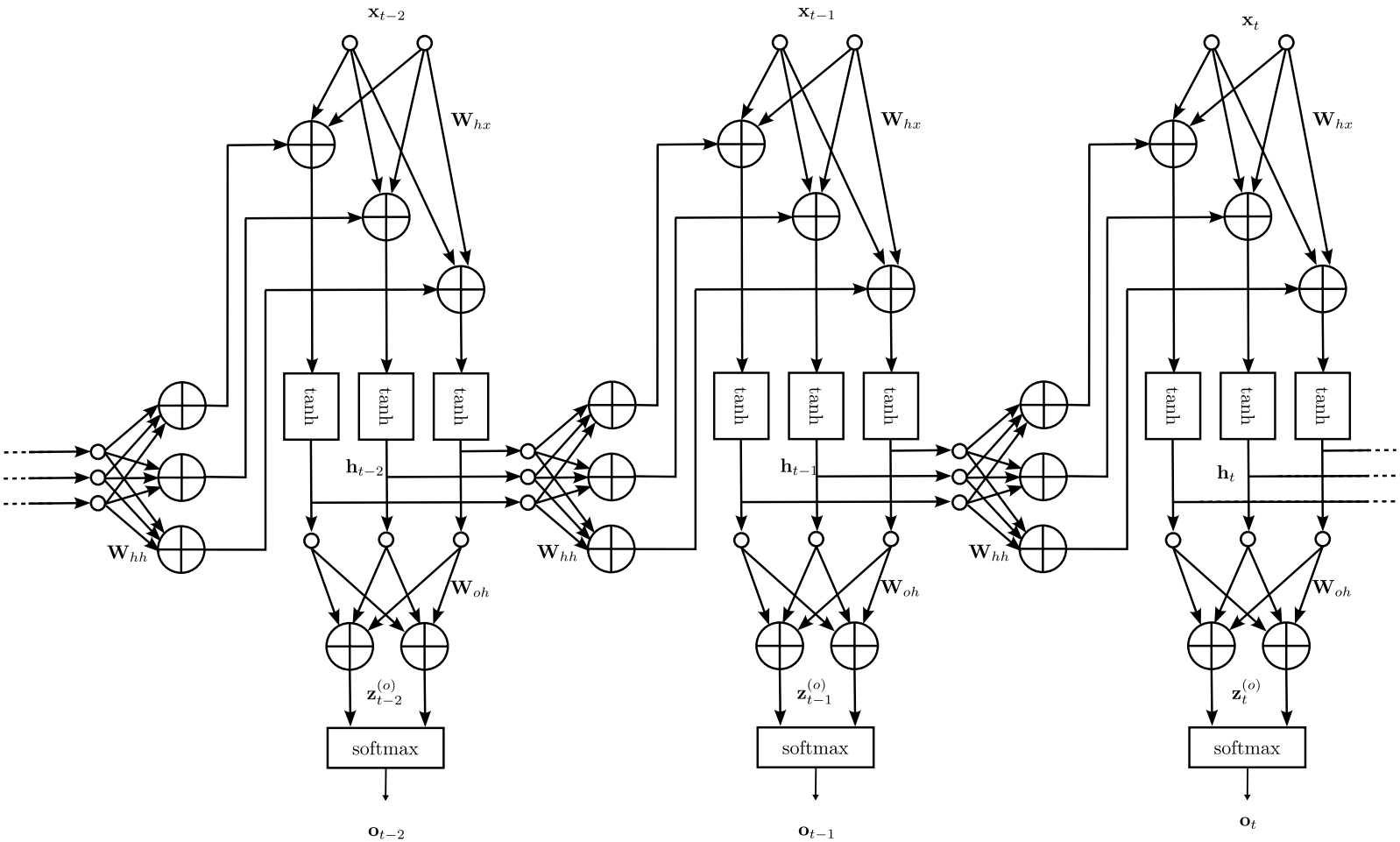
State equations

Introduced by J. L. Elman in 1990, which follows a similar structure by M. I. Jordan, 1986.

- The input matrix $\mathbf{W}_{HX}$ extracts features from the input sample.
- The output matrix $\mathbf{W}_{oh}$ transforms the hidden state into an output response, through a softmax activation.
- The hidden state matrix $\mathbf{W}_{hh}$ performs the feedback. It must store the dependencies between the past inputs and the present output.
- the hidden state itself is a summary of the past samples.
- The network is an infinite impulse response structure.

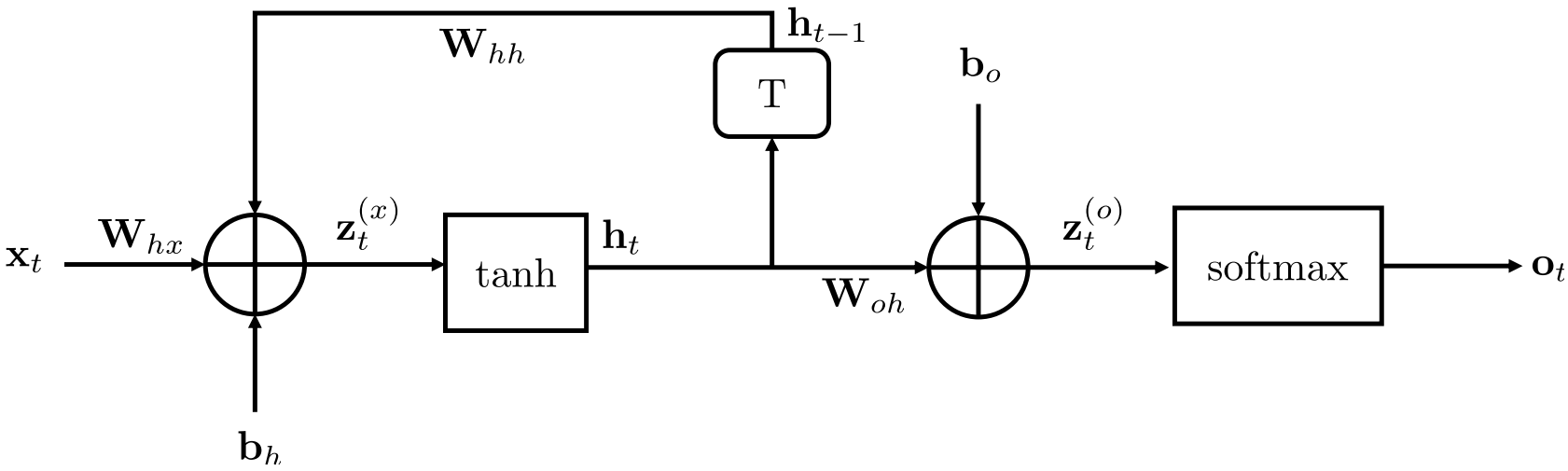
Structure of the Elman RNN

The Elman RNN unfolded



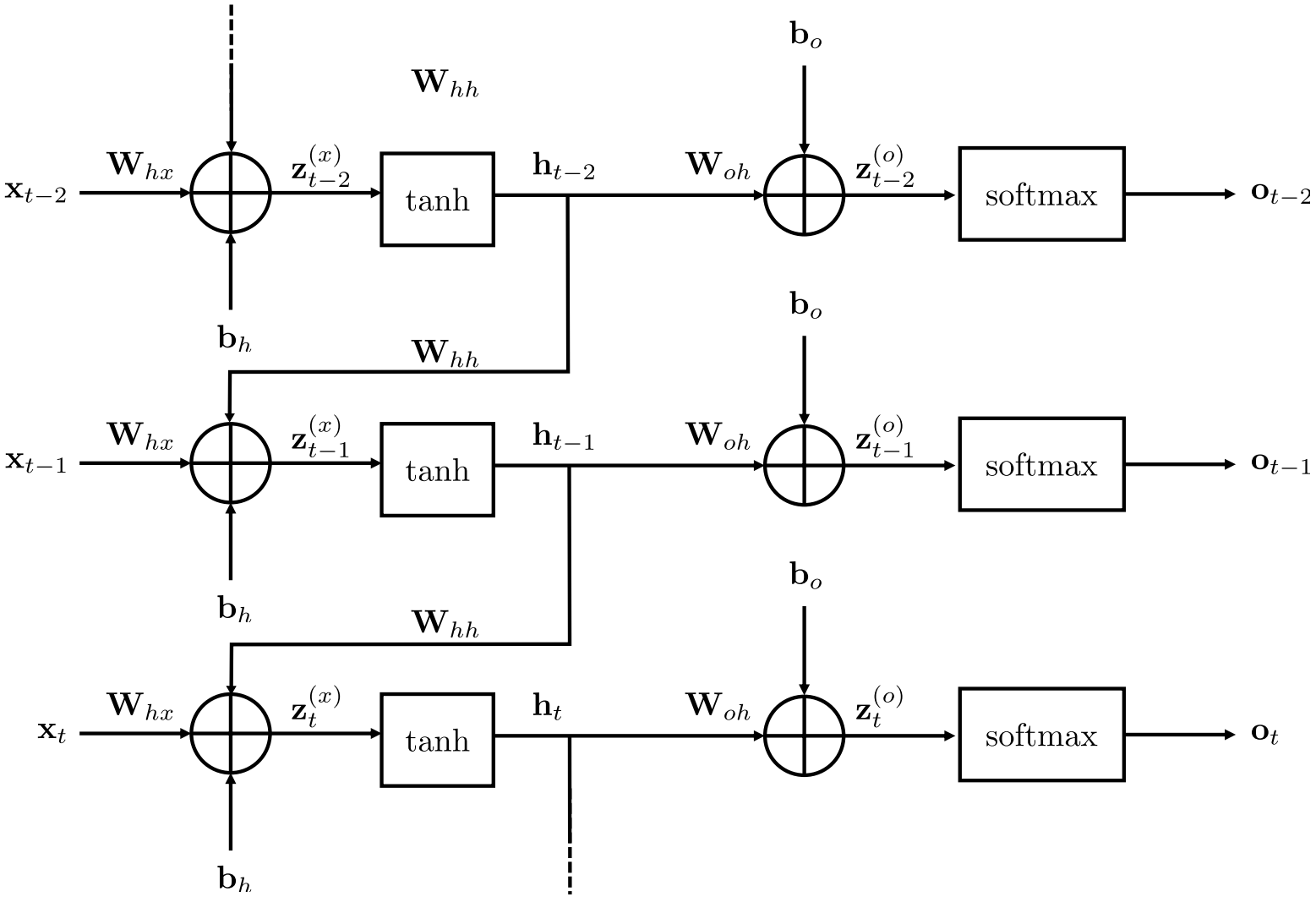
Structure of the Elman RNN

A more compact view of the RNN



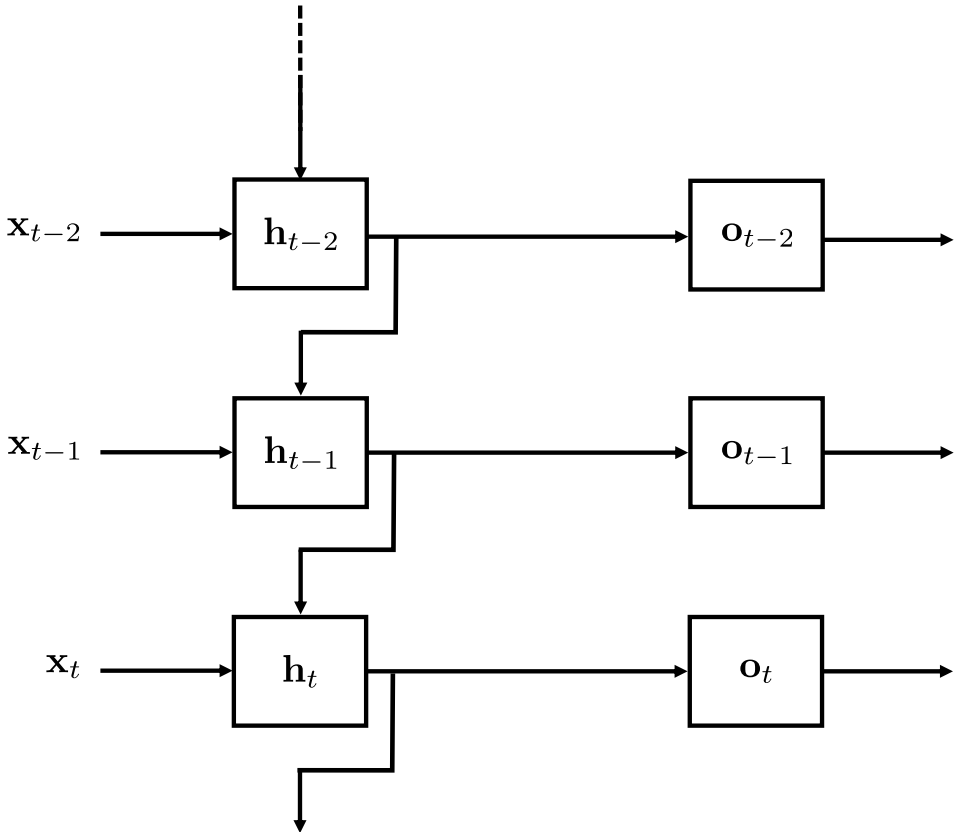
Structure of the Elman RNN

Unfolded compact representation of the RNN



Structure of the Elman RNN

Unfolded even more compact representation



- Input $\mathbf{x}_t$ is transformed with $\mathbf{W}_{hx}$,
- Hidden state $\mathbf{h}_{t-1}$ is transformed with $\mathbf{W}_{hh}$ and bias $\mathbf{b}_h$.
- The output of this box, transformed with a tanh, is $\mathbf{h}_t$.