

can be mitigated by introducing skip connections in the hidden-to-hidden path, as illustrated in figure 10.13c.

10.6 Recursive Neural Networks

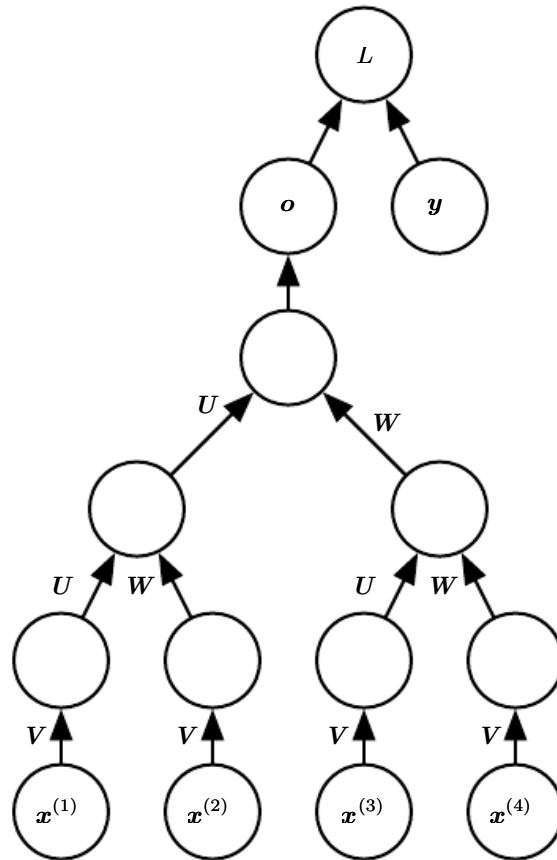


Figure 10.14: A recursive network has a computational graph that generalizes that of the recurrent network from a chain to a tree. A variable-size sequence $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(t)}$ can be mapped to a fixed-size representation (the output \mathbf{o}), with a fixed set of parameters (the weight matrices \mathbf{U} , \mathbf{V} , \mathbf{W}). The figure illustrates a supervised learning case in which some target \mathbf{y} is provided which is associated with the whole sequence.

Recursive neural networks² represent yet another generalization of recurrent networks, with a different kind of computational graph, which is structured as a deep tree, rather than the chain-like structure of RNNs. The typical computational graph for a recursive network is illustrated in figure 10.14. Recursive neural

²We suggest to not abbreviate “recursive neural network” as “RNN” to avoid confusion with “recurrent neural network.”