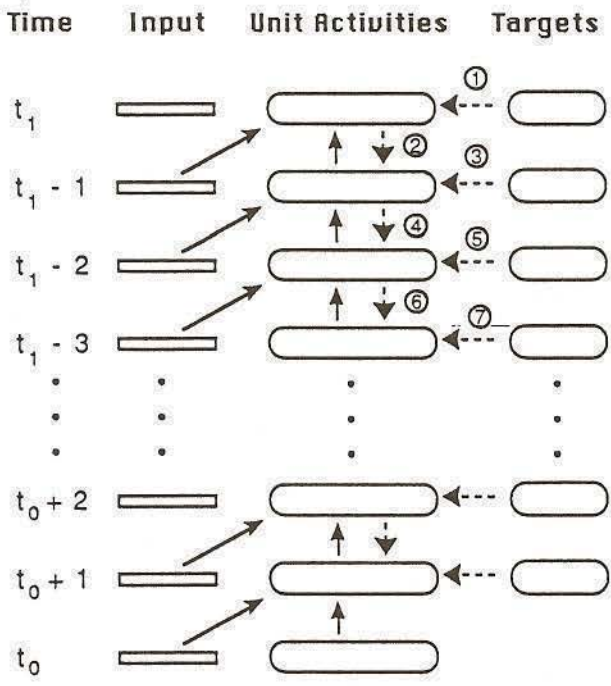

Figure 4. A schematic representation of the storage and processing required for epochwise BPTT. All input, unit output, and target values for every time step from t_0 and t_1 are stored in the history buffer. The solid arrows indicate how each set of unit output values is determined from the input and unit outputs on the previous time step. After the entire epoch is complete, the backward pass is performed as indicated by the dashed arrows. Each even-numbered step determines the virtual error from later time steps, while each odd-numbered step corresponds to the injection of external error. Once the backward pass has been performed to determine separate δ values for each unit and for each time step back to $t_0 + 1$, the partial derivative of the negative error with respect to each weight can then be computed.