



## Legend

- unweighted connection
- weighted connection
- - - connection with time-lag
- branching point
- $\odot$  multiplication
- $+$  sum over all inputs
- $\sigma$  gate activation function (always sigmoid)
- $g$  input activation function (usually tanh)
- $h$  output activation function (usually tanh)