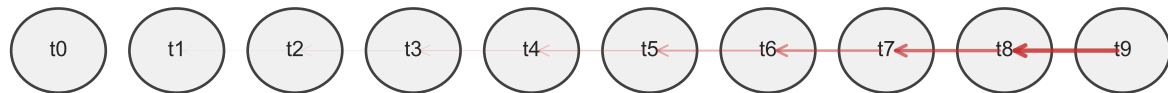


RNN: Vanishing Gradient

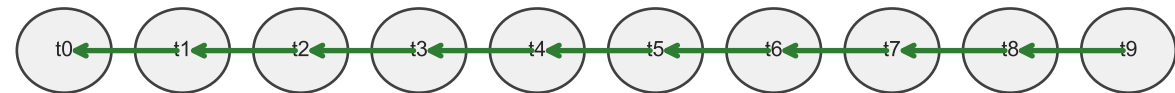


1.000 0.900 0.810 0.729 0.656 0.590 0.531 0.478 0.430 0.387

Gradient shrinks exponentially: $0.9^{10} \approx 0.35$

LSTM: Gradient Highway

Cell State Highway



1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00

Gradient preserved: $1.0^{10} = 1.0$