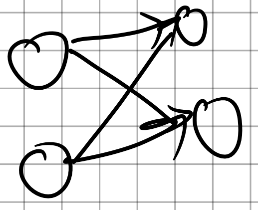

1. XOR:

$h_1[t] = \sigma(w_{11} x[t] + w_{12} s[t-1] + b_1)$

$h_2[t] = \sigma(w_{21} x[t] + w_{22} s[t-1] + b_2)$

Output

$y[t] = \sigma(v_1 h_1[t] + v_2 h_2[t] + b)$

$s[t] = y[t]$

$y[T] = s[T] = x[1] \oplus x[2] \oplus x[2] \oplus \cdots \oplus x[T]$

2.

$y[t] = XOR(x[t], s[t])$

$y[1] = x[1]$

$x[1] = XOR(x[1], s[1])$

$s[1] = 0$

$y[1] = XOR(x[1], 0) = x[1]$

3.

$y[t] = XOR(x[t], s[t])$

$y[1] = x[1]$

$s[t] = y[t-1]$ for $t > 1$

$y[t] = x[t] \oplus s[t] = x[t] \oplus y[t-1]$

$s[1] = 0$, $s[t] = y[t-1]$ for $t > 1$

$y[T] = x[1] \oplus x[2] \oplus \cdots \oplus x[T]$

4.

$s[t]$ - the state / hidden state (the RNN's memory)

$s[t]$ - Single binary component - Represents running XOR of all past inputs.

5.

$y[t]$ is the running XOR of all inputs up to time t.