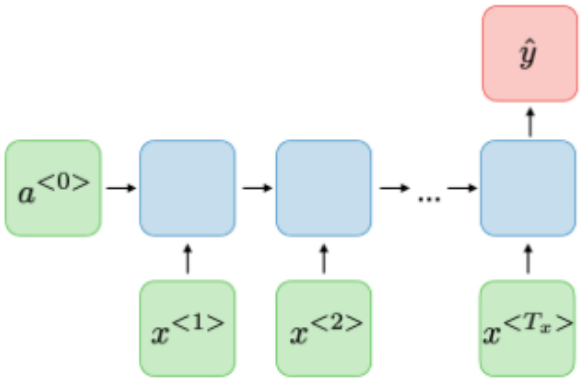
Xianchi zhang / echo chen



For each timestep *t*, the activation $a^{<t>}$ and the output $y^{<t>}$ are expressed as follows:

$a^{<t>} = g_1(W_{aa}\, a^{<t-1>}\, W_{ax}\, x^{<t>} + b_a)$

$y^{<t>} = g_2(W_{ya} a^{<t>} + b_y)$

Input： ["three", "one", "four", "one", "five", "two", "five", "three", "five"]

One-hot encoding: “one”=[1,0,0,0,0]. “two”=[0,1,0,0,0] “three”=[0,0,1,0,0] “four”=[0,0,0,1,0] “five”=[0,0,0,0,1]

$g_1 = (w_1 a^{<t-1>} + w_2 x^{<t>})$ $w_2 = [\log5, \log4, \log3, \log2, \log1]^T$ $w_{1.} = 1$

$g_2 = e^{W_{ya}{}^{<t>}}$ $w_y = 1$

$a_0 = 0$

$a_1 = 0 + \log3$

$a_2 = \log3 + \log1$

.

.

.

$a_9 = \log(3)+\log(1)+\log(4)+\log(1)+\log(5)+\log(2)+\log(5)+\log(3)+\log(5)$

$y_9 = e^{a9} = 9000$