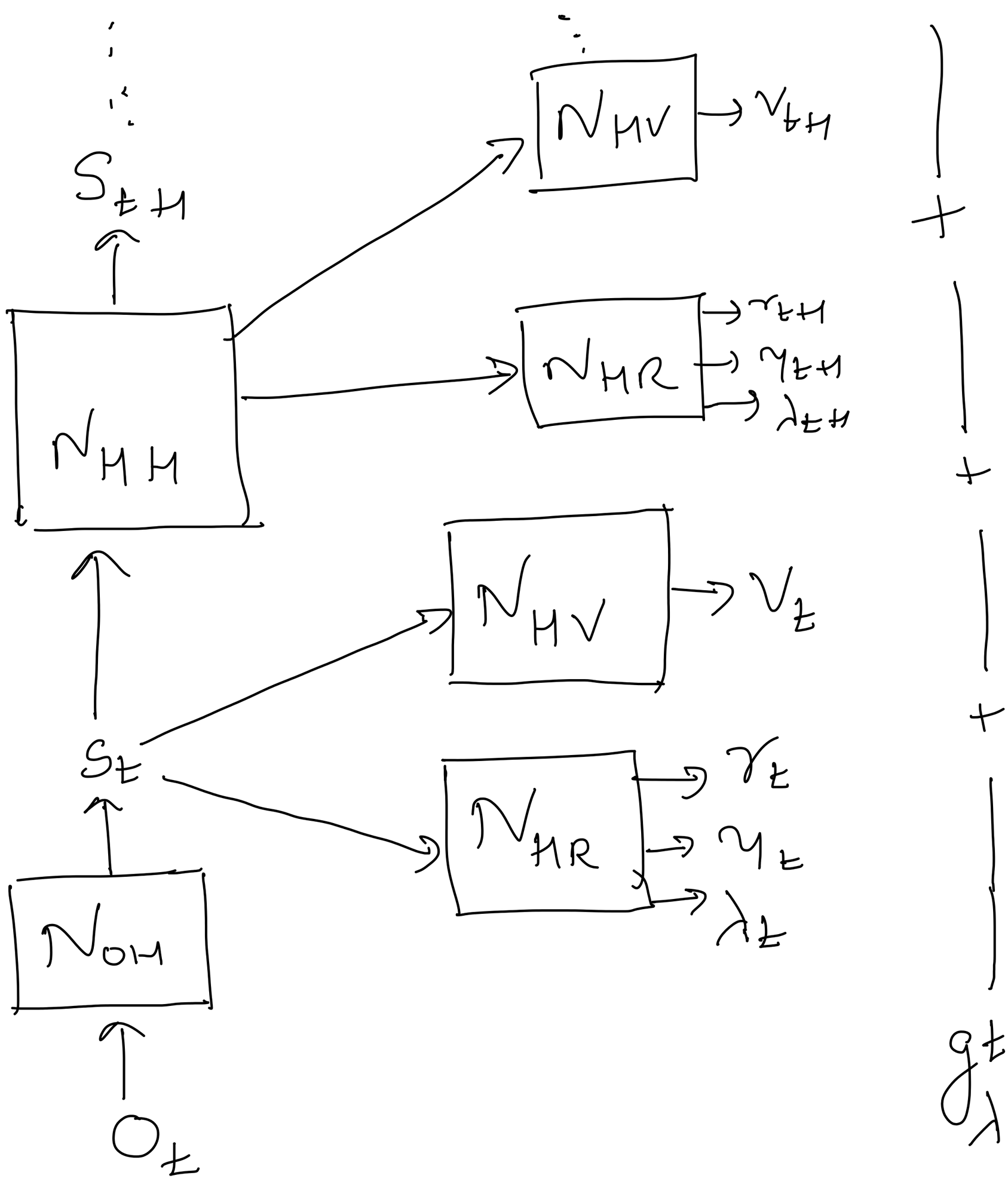
Architecture

Thursday, September 24, 2020 10:56 PM



$O_t \Rightarrow$ Observations from CMAPPs data (24)

$S_t \Rightarrow$ Hidden state for observation O_t

$S_{t+1} \Rightarrow$ Next Hidden state

$N_{OH} \Rightarrow$ observation to hidden state $\Rightarrow (21, 32, 16, 4)$

$N_{HR} \Rightarrow$ hidden state to reward $\Rightarrow (4, 32, 16, 3)$

$N_{HV} \Rightarrow$ hidden state to value $\Rightarrow (4, 32, 16, 1)$

$N_{HH} \Rightarrow$ hidden state to next hidden state $\Rightarrow (4, 32, 16, 4)$

→ We can start with simple k-step return(g_k) and then we can integrate TD(λ) return (g_λ).

→ As discussed, we will be using Monte-Carlo return (g) in our loss function $E[(g_\lambda - g)^2]$

→ Later, we would also like to try out back up ($R + \gamma V(s')$) as the target value.