S : state at time t

S' : state at time t+1

state $\in [2d, n]$

$D = [S^1 \dots, S^t, S^{t+1} \dots]$

Next steps
- ✓ modelling
- ○ tasks

DT : flatten(S) $[2dxn, 1]$ → MLP_θ → S'

NF : flatten(S) $[2dxn, 1]$ → MLP_θ → f → solve → $\begin{bmatrix} s \\ s' \end{bmatrix}$ → MLP_ϕ → S'

WC

1. architecture
2. hard to fit
3. learn signal is small.

NBF : s_1 $[2d, 1]$, s_2 $[2d, 1]$, s_3 $[2d, 1]$ → MLP θ → $\begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}$ $[nE, 1]$ → MLP_ϕ → f → solve → S'

NR : s_1, s_2, s_3 → MLP θ → $[e_1 \circ e_2]$, $[e_1 \circ e_3]$, $[e_2 \circ e_1]$, $[e_2 \circ e_3]$, $[e_3 \circ e_1]$, $[e_3 \circ e_2]$ $[2E, 1]$ → MLP ϕ_1 (g.f.) → f_{12}, f_{13}, f_{21}, f_{23}, f_{31}, f_{32} → $\oplus$ → $\begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}$ → solve → S'

* MLP_ϕ e.f. attention

* $s_4 \dots$