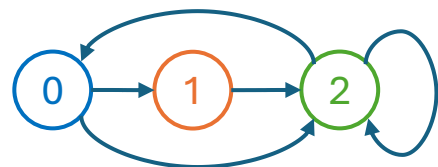


Common-used



or

$$x_{k+1} = f(x_k)$$

Type 1: Unaligned Sequence

1,0,2,2,1,2,...

$$\begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix}, \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, \begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix}, \dots$$

Transformer

 $x_1, x_2, x_3, \dots, x_{k-1}$ x_k

Type 2: Aligned Sequence

$$y_k = f(x_k)$$

$$\begin{bmatrix} 0.5 \\ 0.5 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1.0 \end{bmatrix}, \begin{bmatrix} 0.3 \\ 0.4 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0.7 \end{bmatrix}, \dots$$

Transformer

 $x_1, y_1, x_2, y_2, \dots, x_k$ x_k

This work

$$\mathcal{H} = \{h_1, h_2, h_3\}$$

	a_1	a_2	a_3
h_1	0	1	1
h_2	1	0	1
h_3	1	0	0



Transformer

hypothesis prefix \oplus $x_1, y_1, x_2, y_2, \dots, x_k, y_k, >$ $h_?$

$$\mathcal{D} = \{(x_k, y_k)\}_{k=1}^K$$



Given Hypothesis space \mathcal{H}
and dataset \mathcal{D} ,
can Transformer identify
the correct hypothesis?