

↓
John ate the apple

one-hot encoding

abc j
| 0 ... 0 1 0 ... 1 |

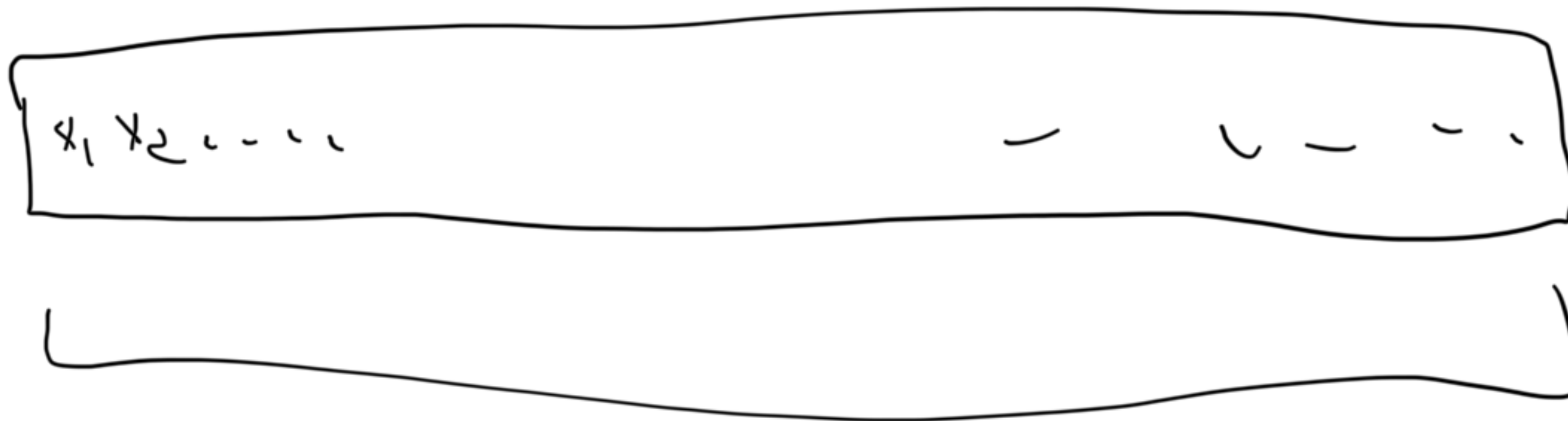
26 John
| 0 ... 0 1 0 ... 0 |
90k

embedding

w_1, w_2
similar

$\mathbb{R}^n$

etc



$\| \text{eat, drink} \| <$
 $\| \text{eat, sleep} \|$

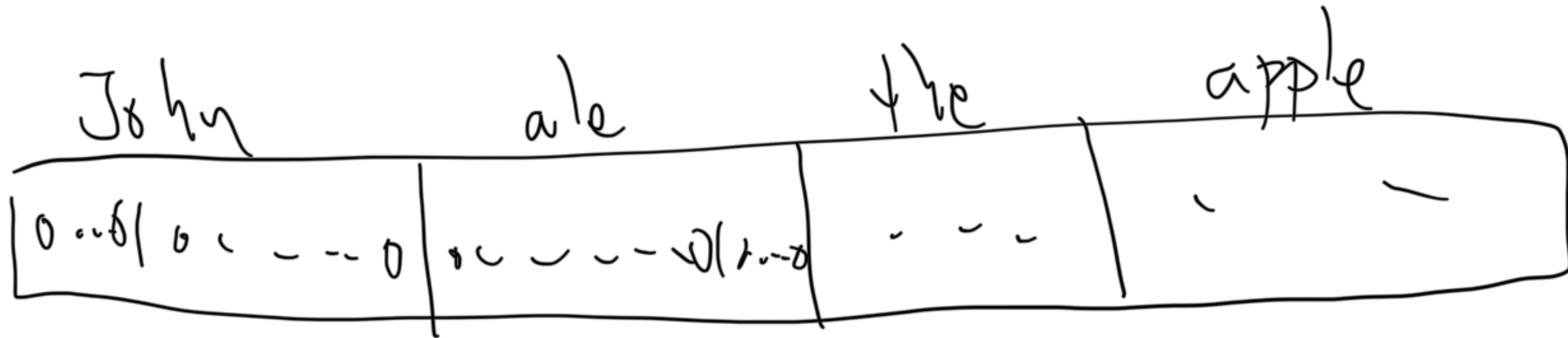
king
queen

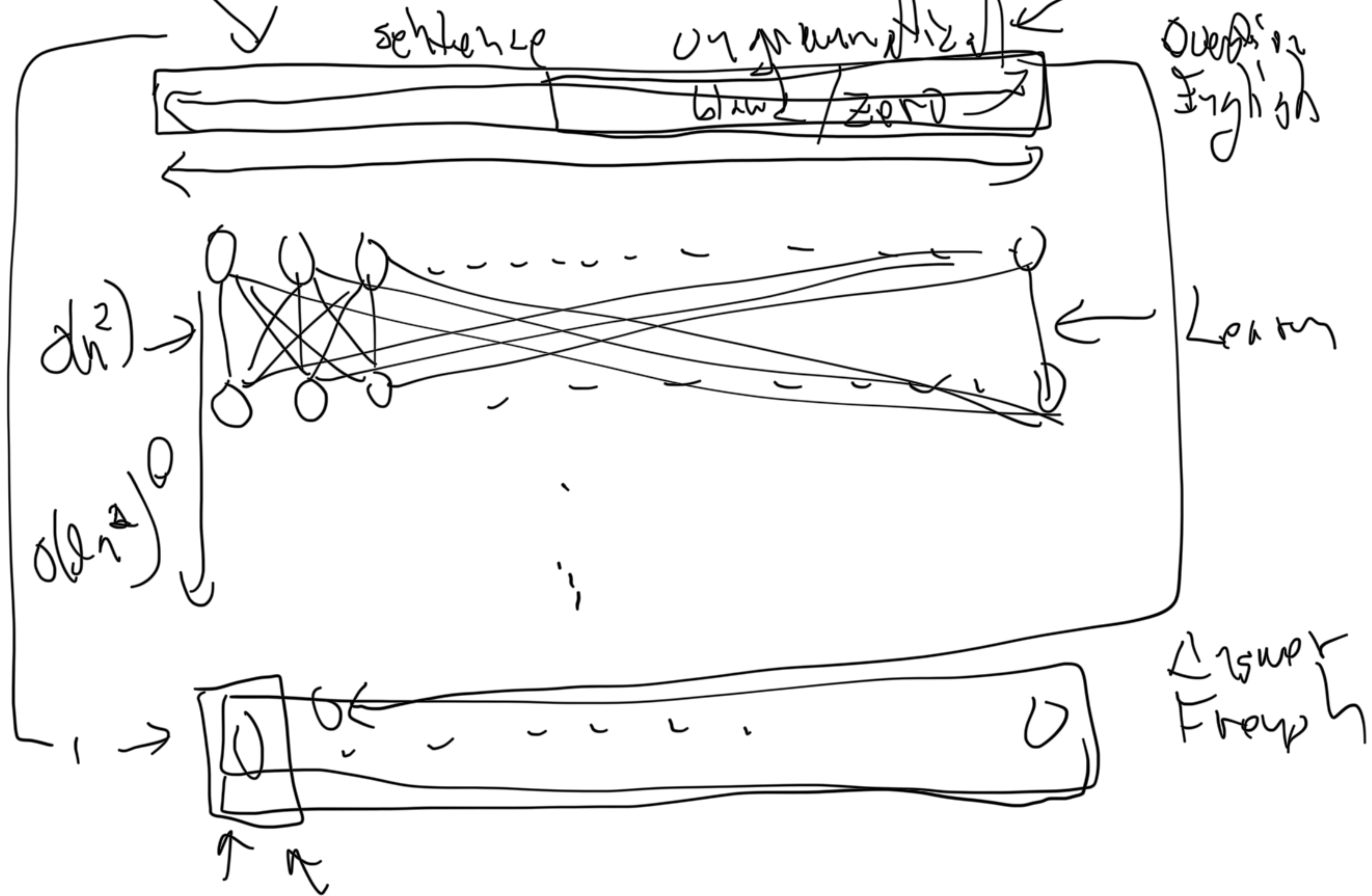
queen - king - many & woman

man

woman

John ate the apple





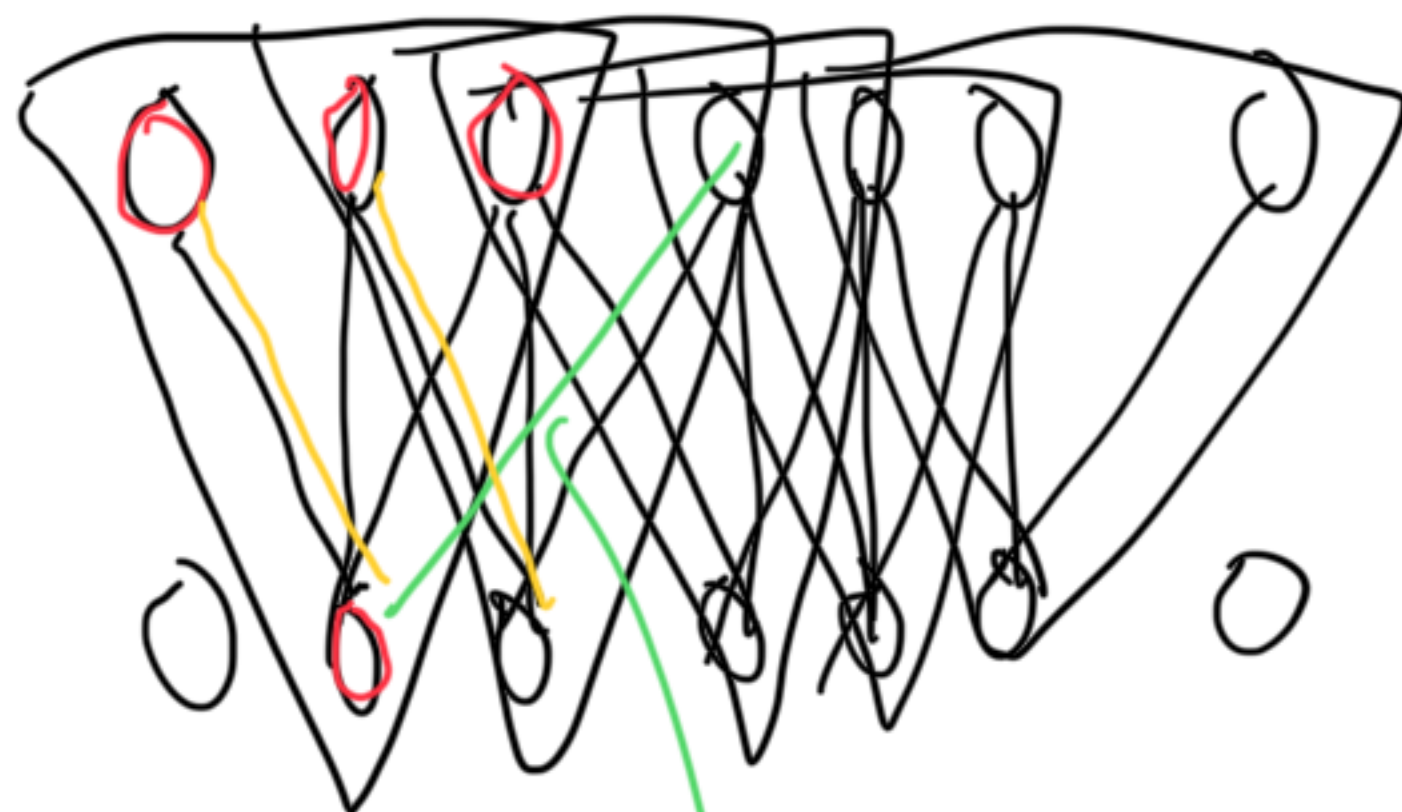
Language in local
Grammaticality

$NP \rightarrow \overline{DN}$

$VP \rightarrow \overline{VNP}$



Sparsity (local)



$w = 0$
constant

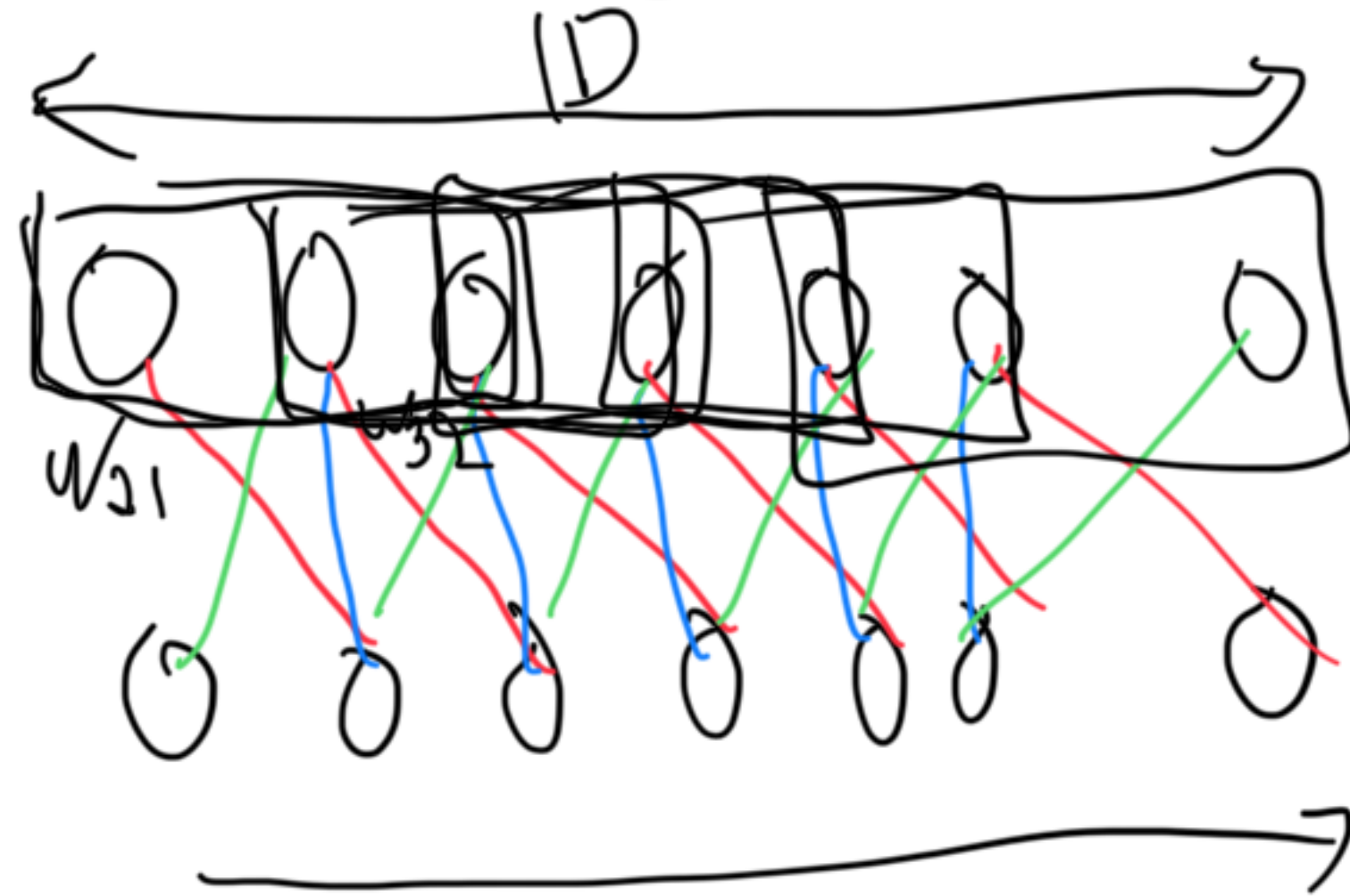
learning bias
prior
incorporating
expert knowledge

Language is position invariant
Grammaticality

beginning
the apple fell to the ground

NP end
I ate the apple that fell to the ground

weight tying & sparsity =

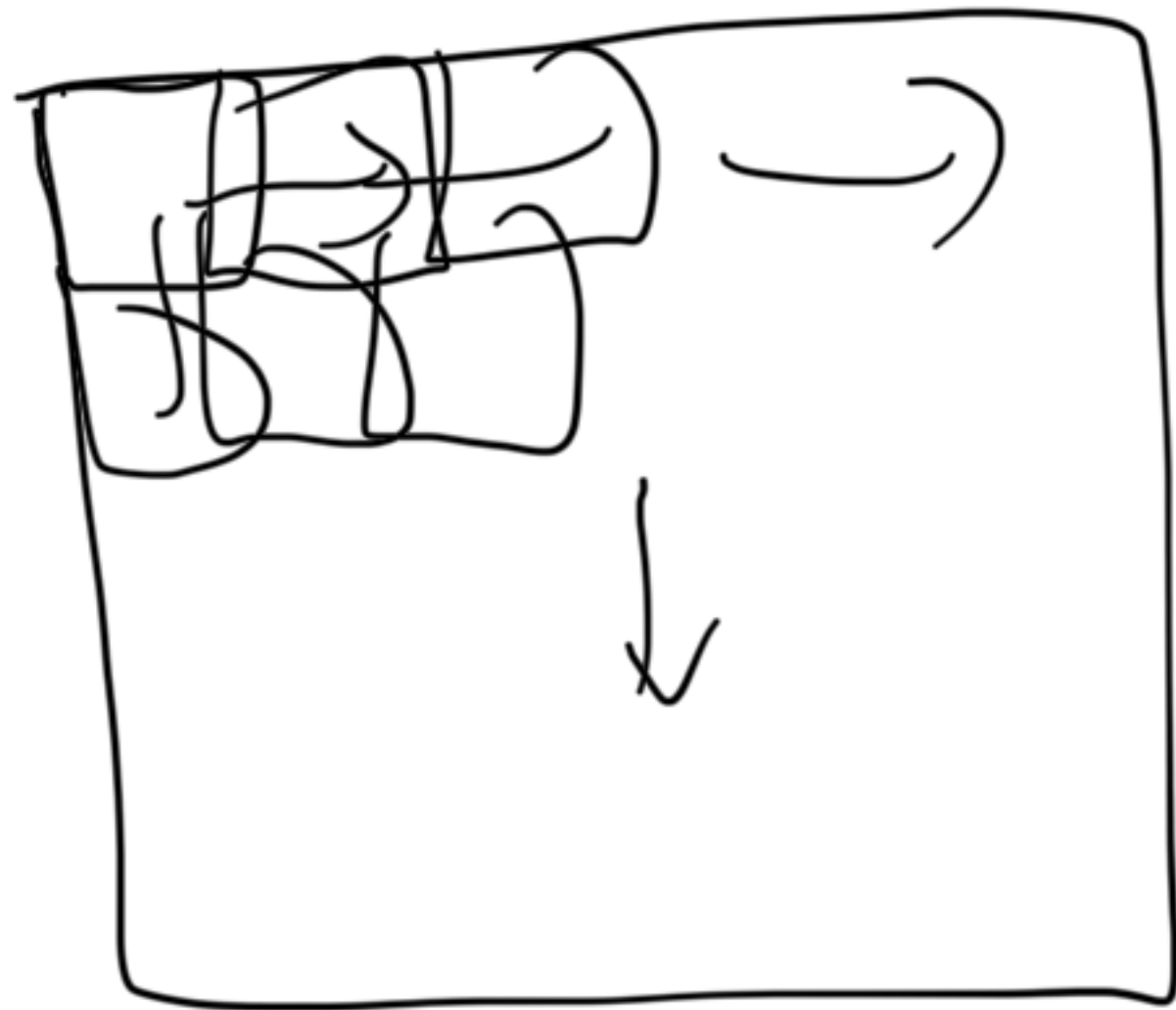


convolutional
neural
network

1D CNN

$$w_{21} = w_{32}$$

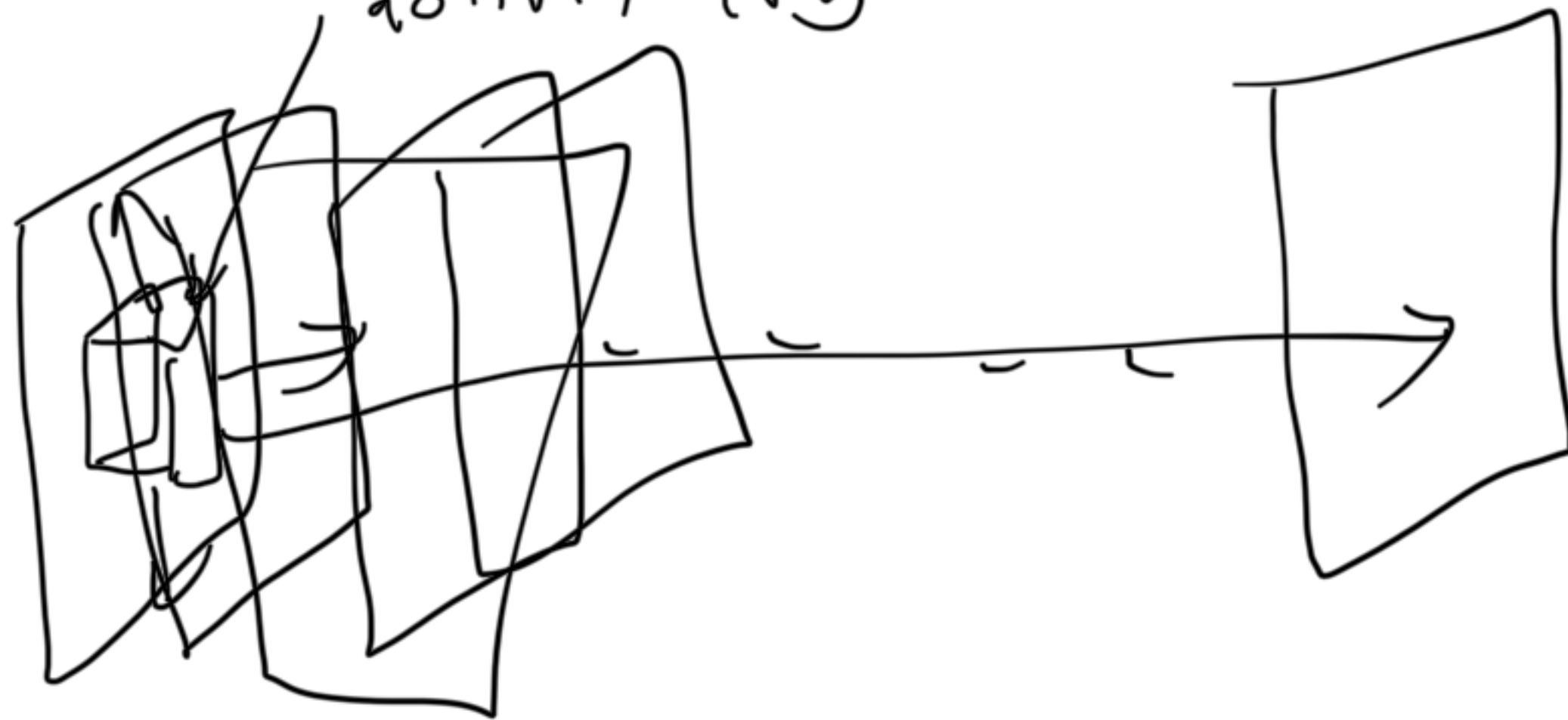
convolution

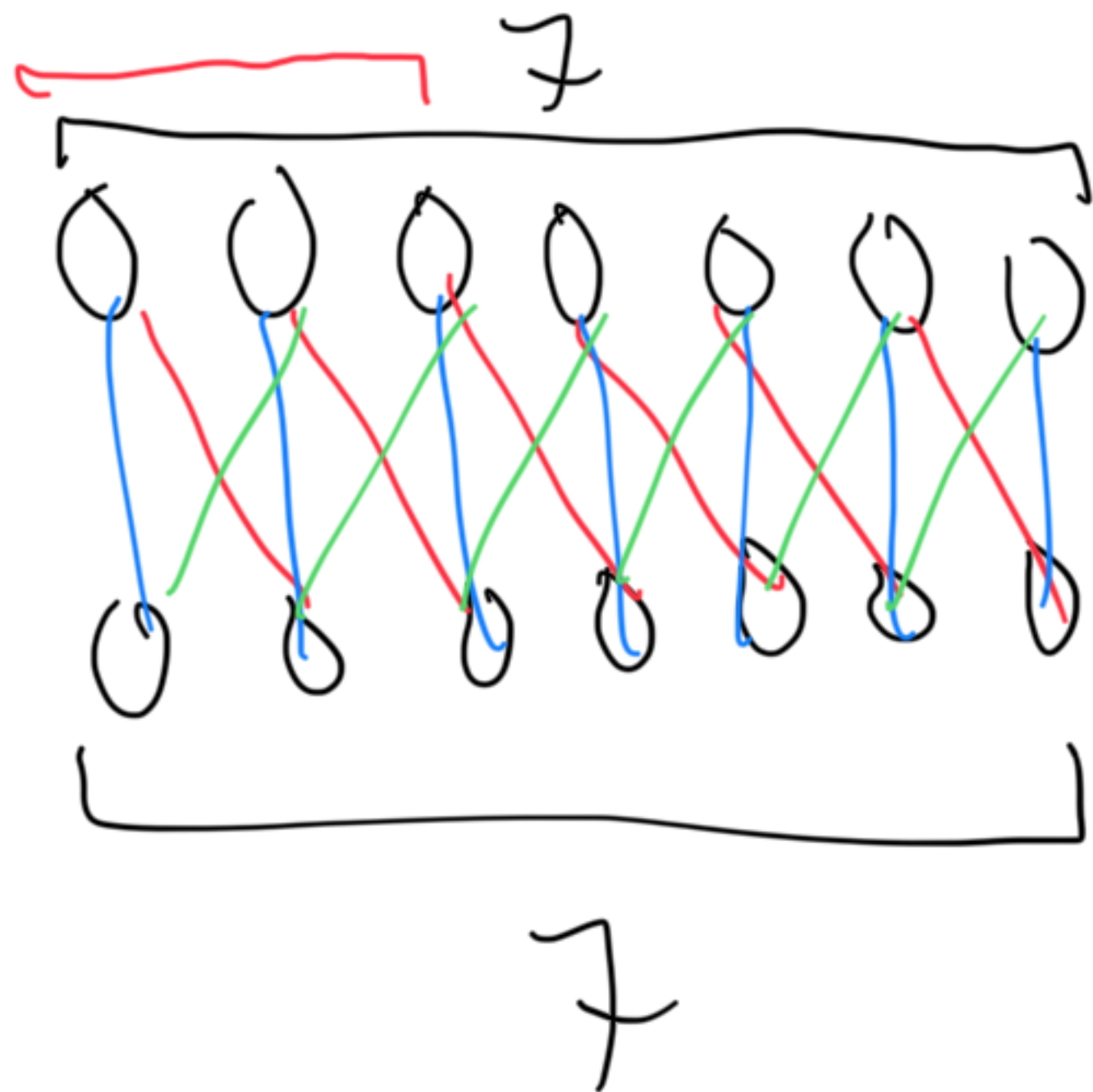


2D CNN
computer vision
local
spatially
invariant

3D CNN

action / verb

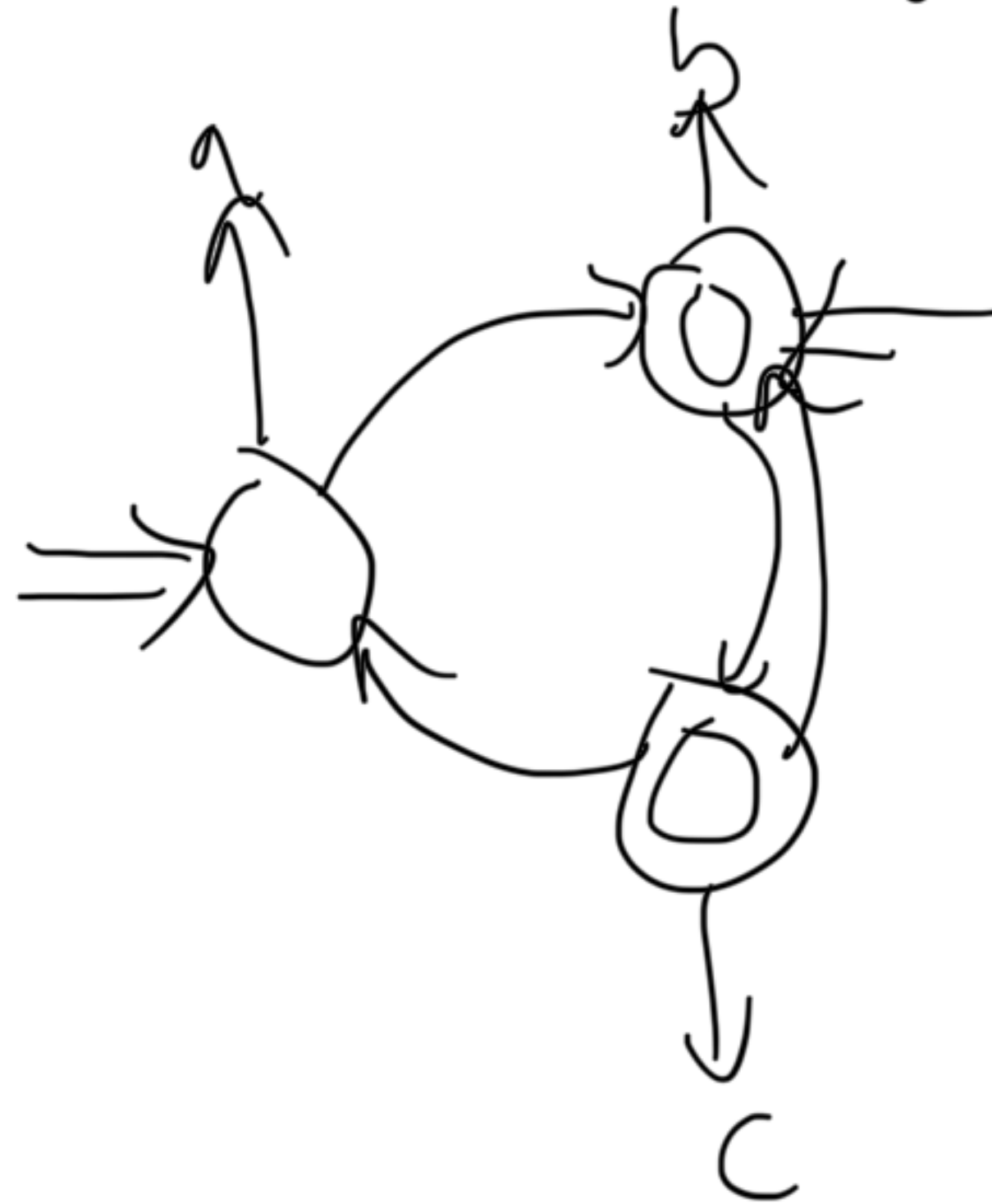




$$7^2 = 49$$

$$3 < 49$$

temporal sequence



FSM

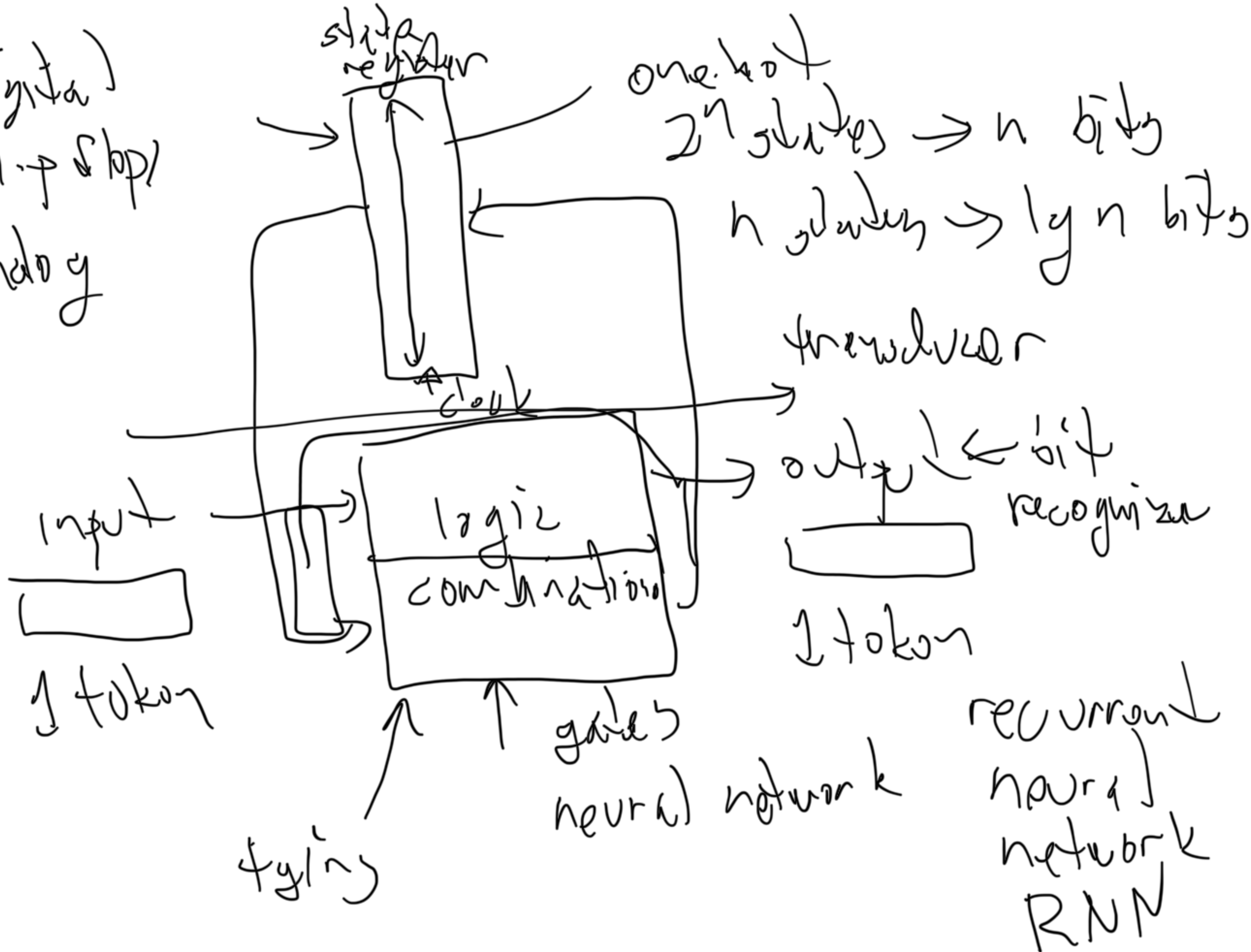
recognizer FA

transducer

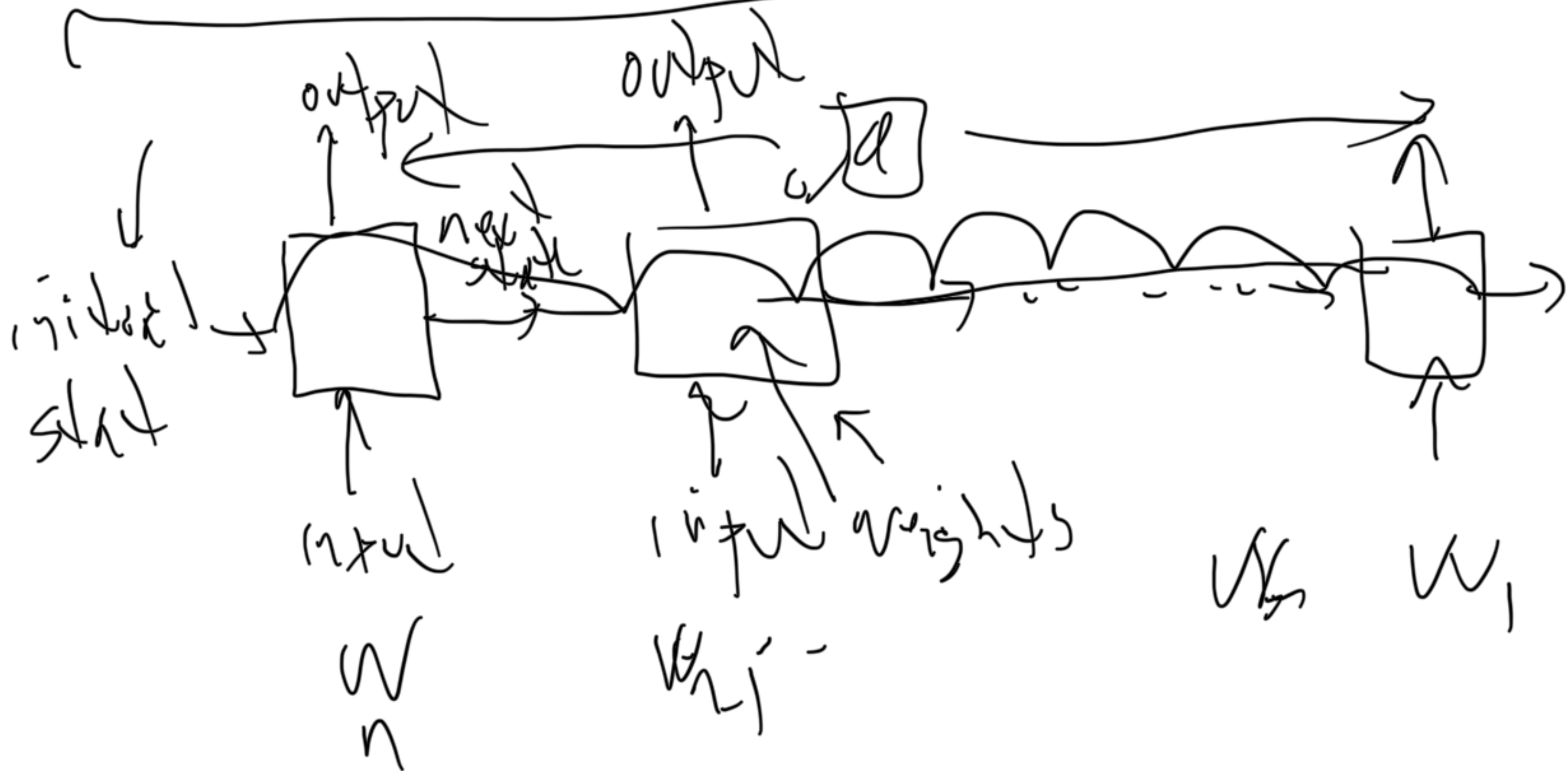
probabilistic

HMM

digital
slip & bp/
analog



loop unrolling



$$\begin{aligned}
 & a^n \\
 & a < 0 \\
 & a > 0
 \end{aligned}
 \quad
 \begin{aligned}
 & n \rightarrow \infty \\
 & a^n \rightarrow 0 \\
 & a^n \rightarrow \infty
 \end{aligned}$$