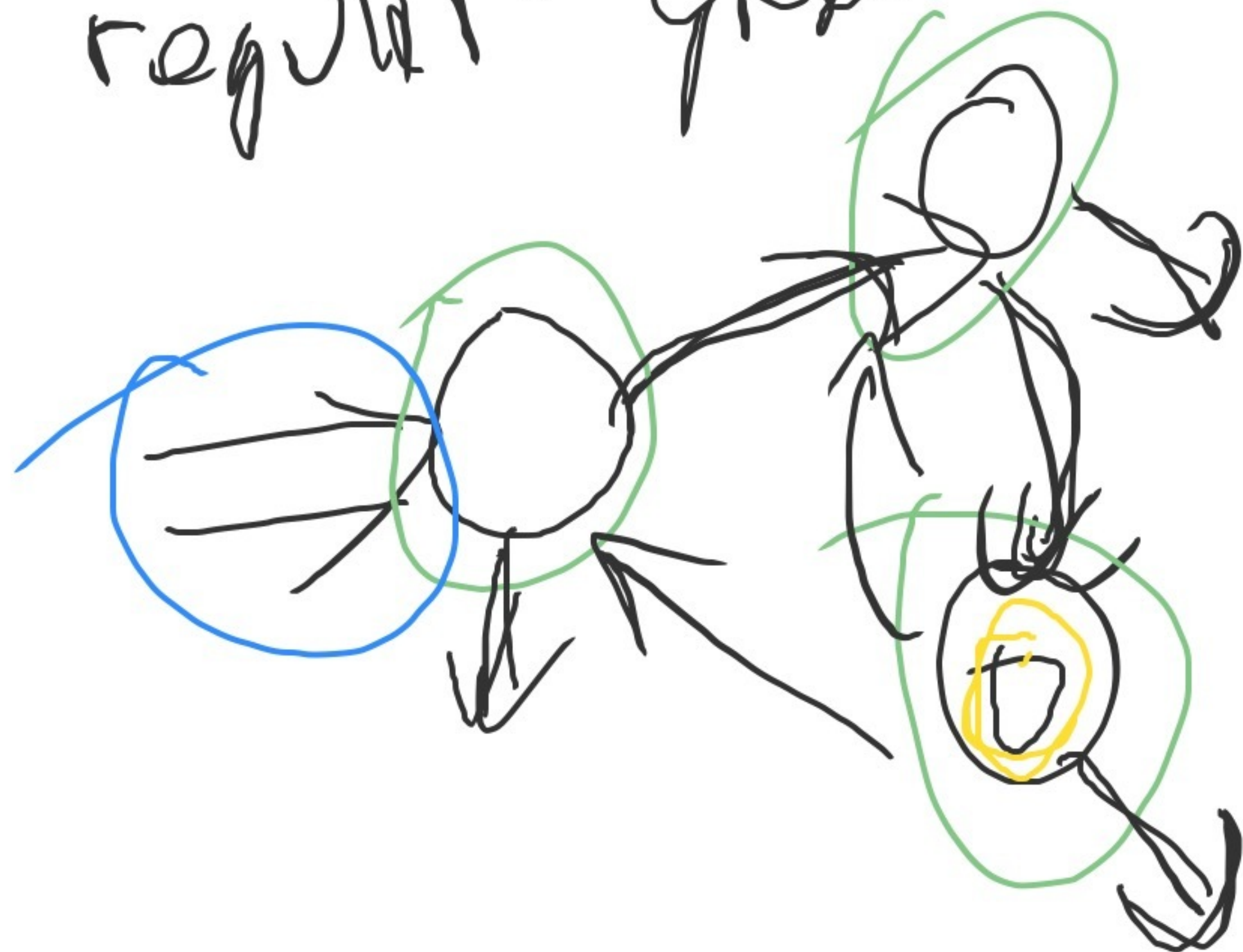


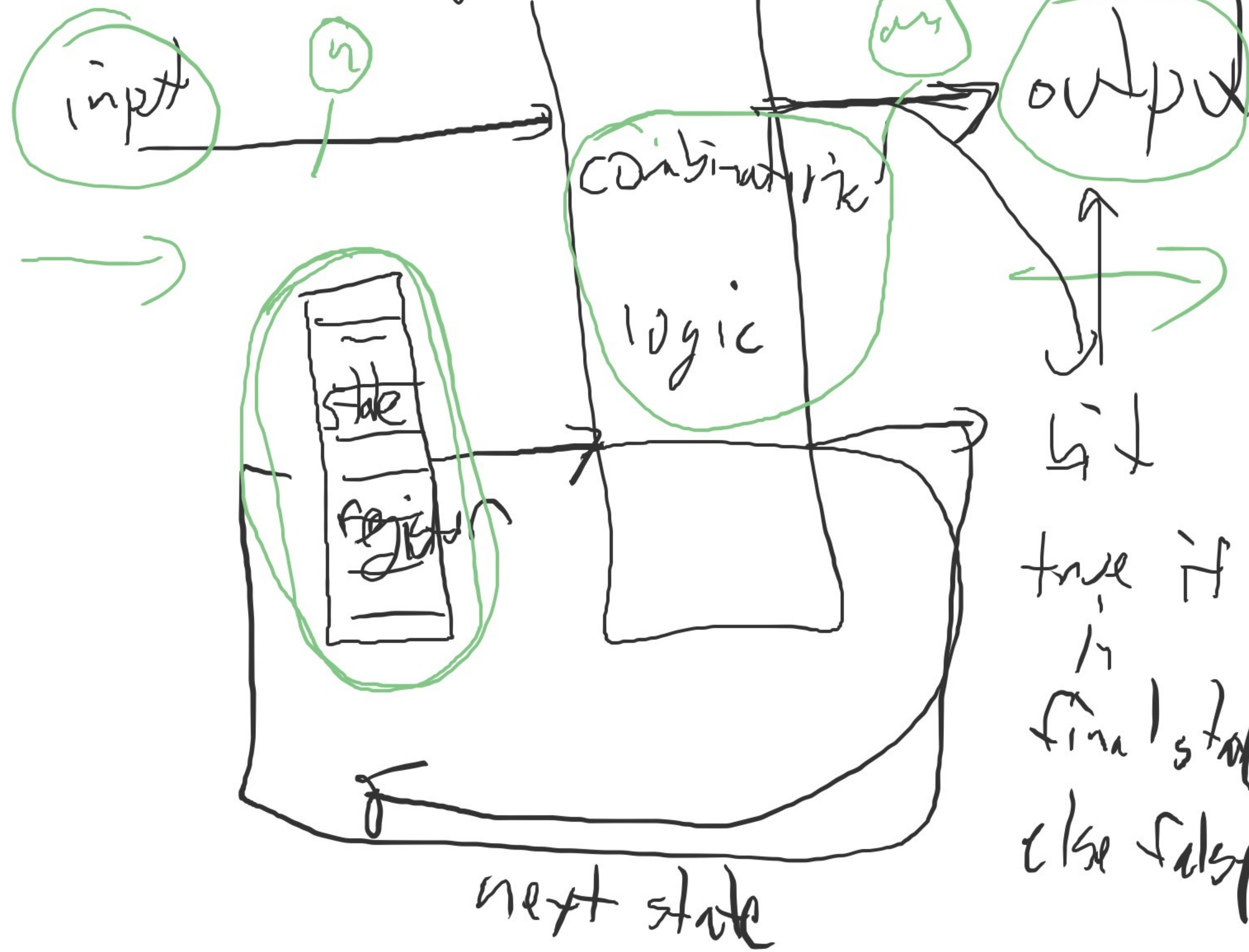
regular grammar



HMM

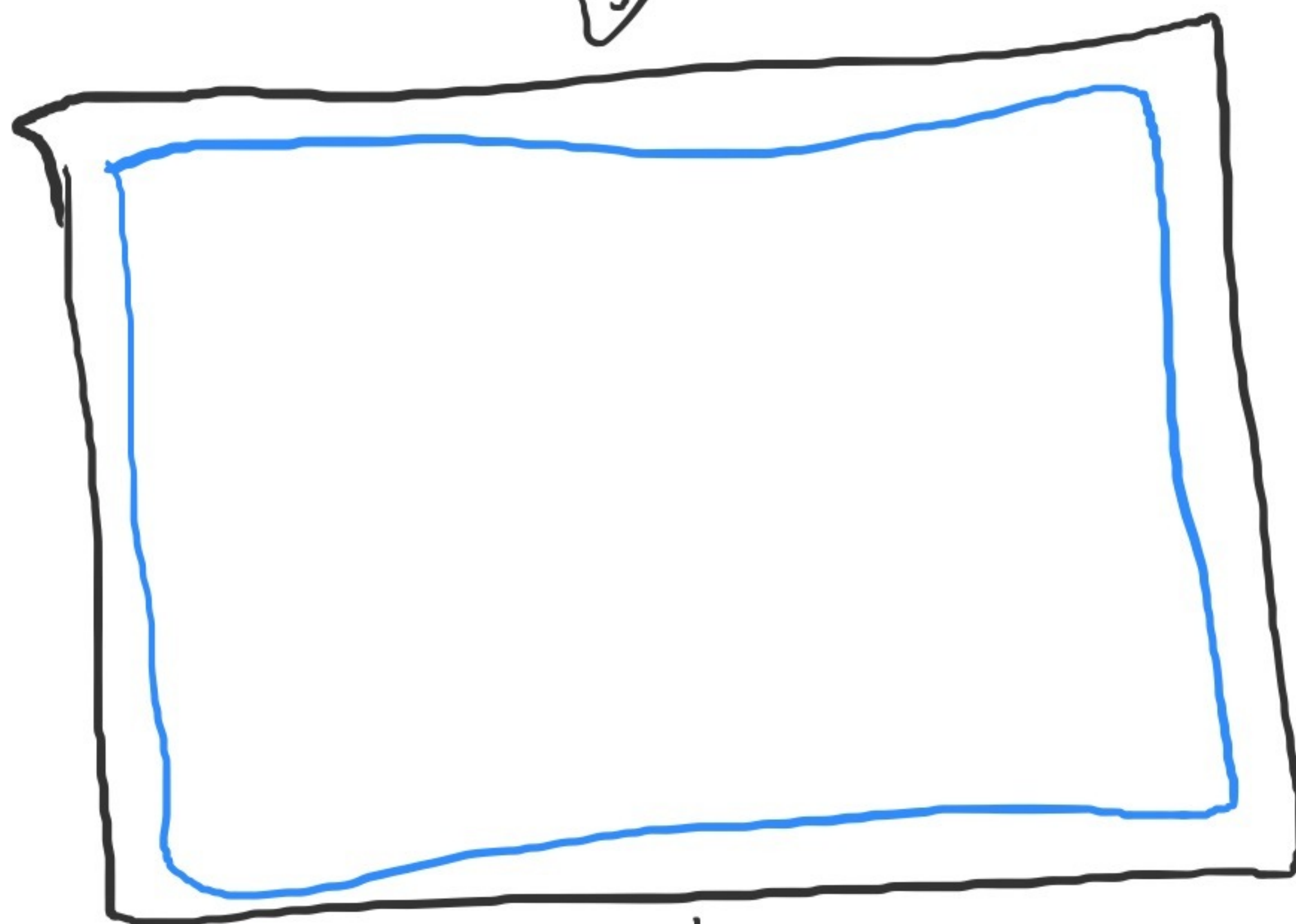
Finite State Machine Transducer

recognizer



recognizer
transducer

~~sequence~~

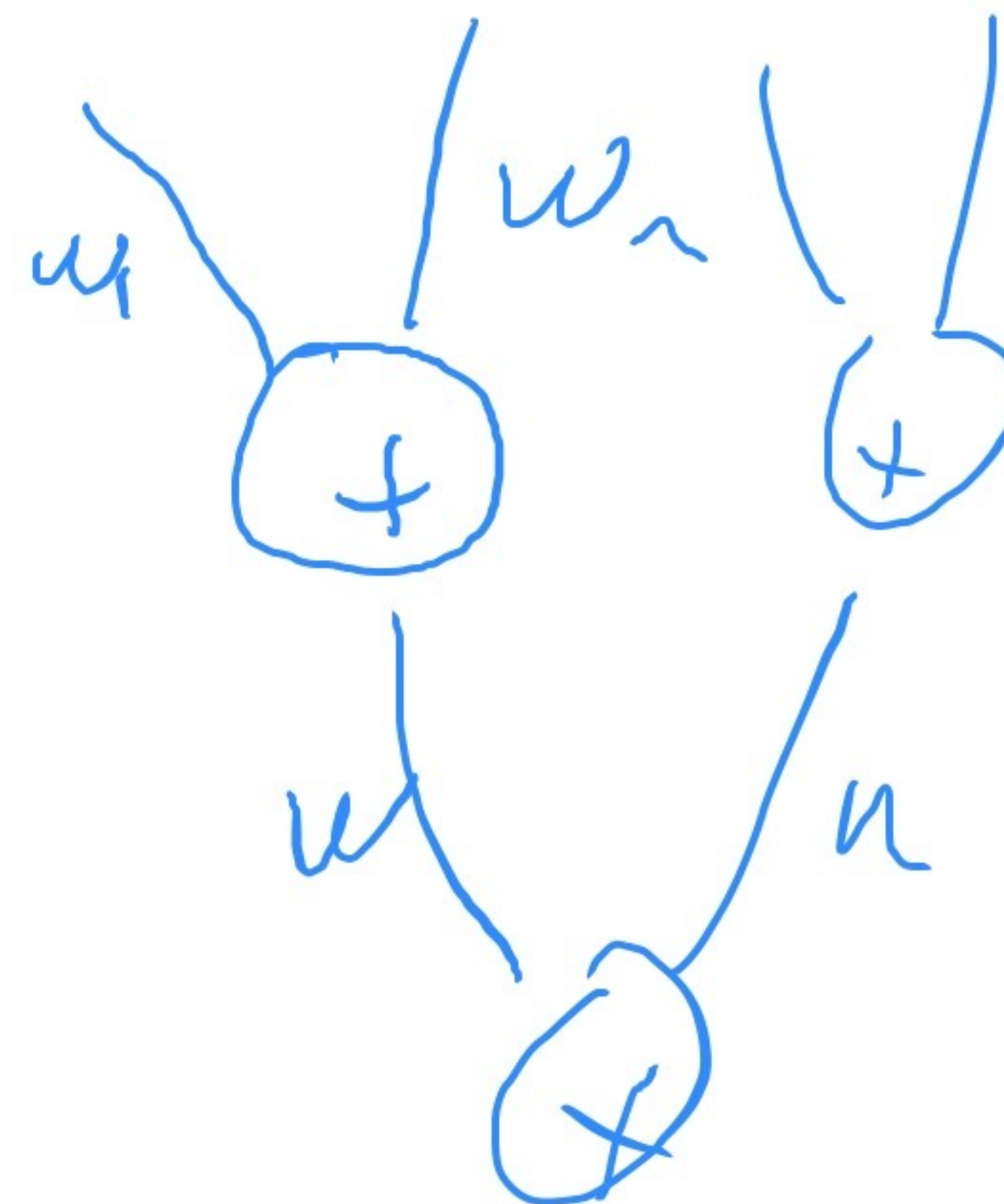


output ~~bit~~

true

false

sequence



wordy K

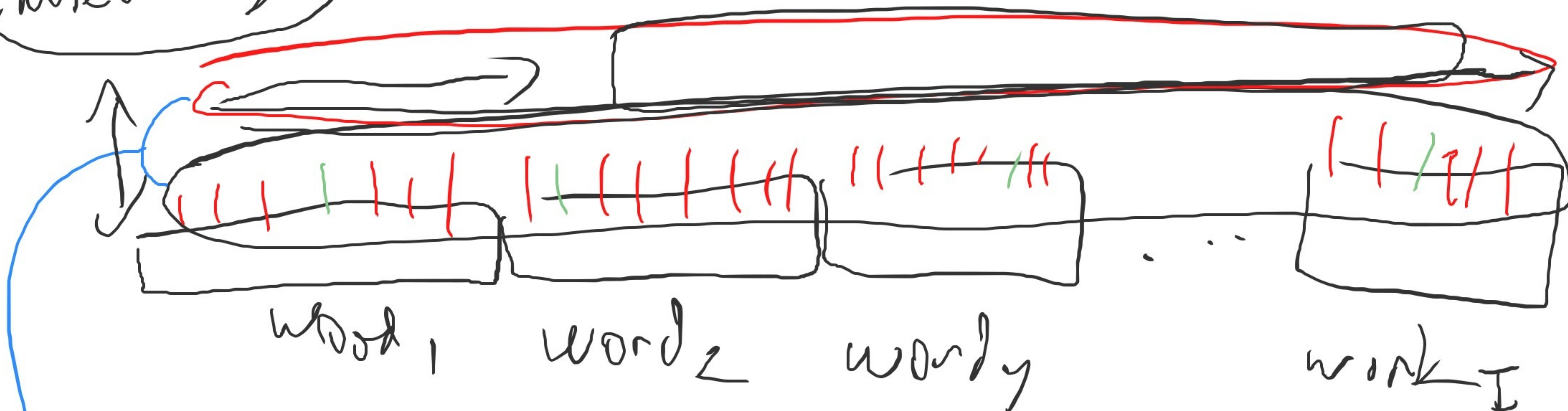
(inference)

(embeddings)

one hot
1 0 0 0 1 0 ... 0

↓
K

K



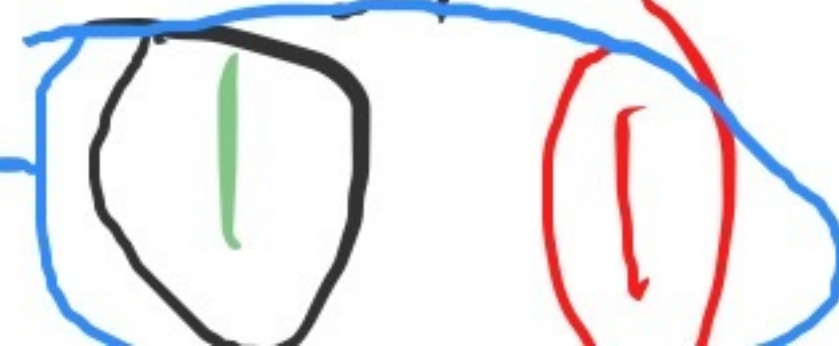
word₁

word₂

word_y

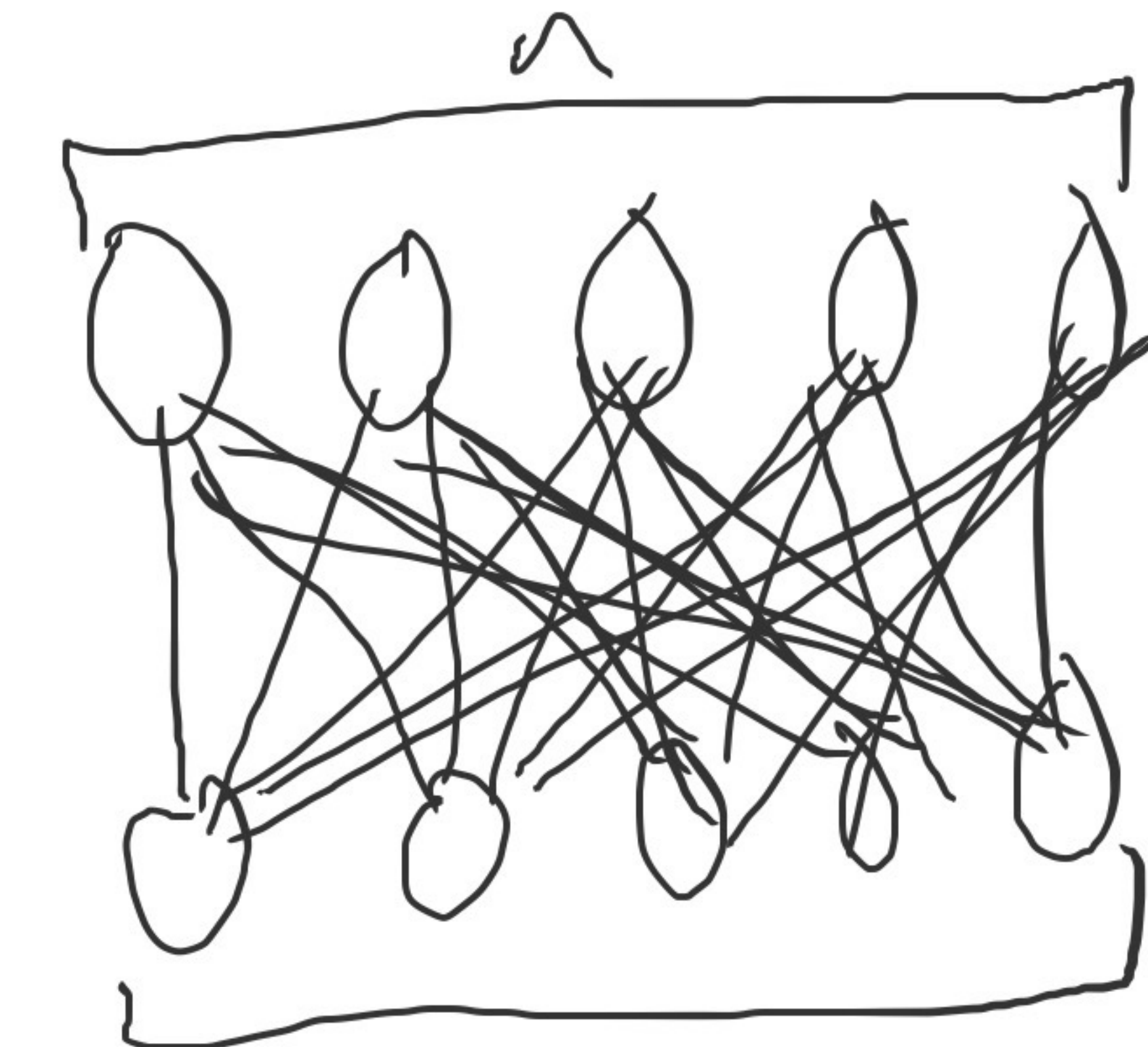
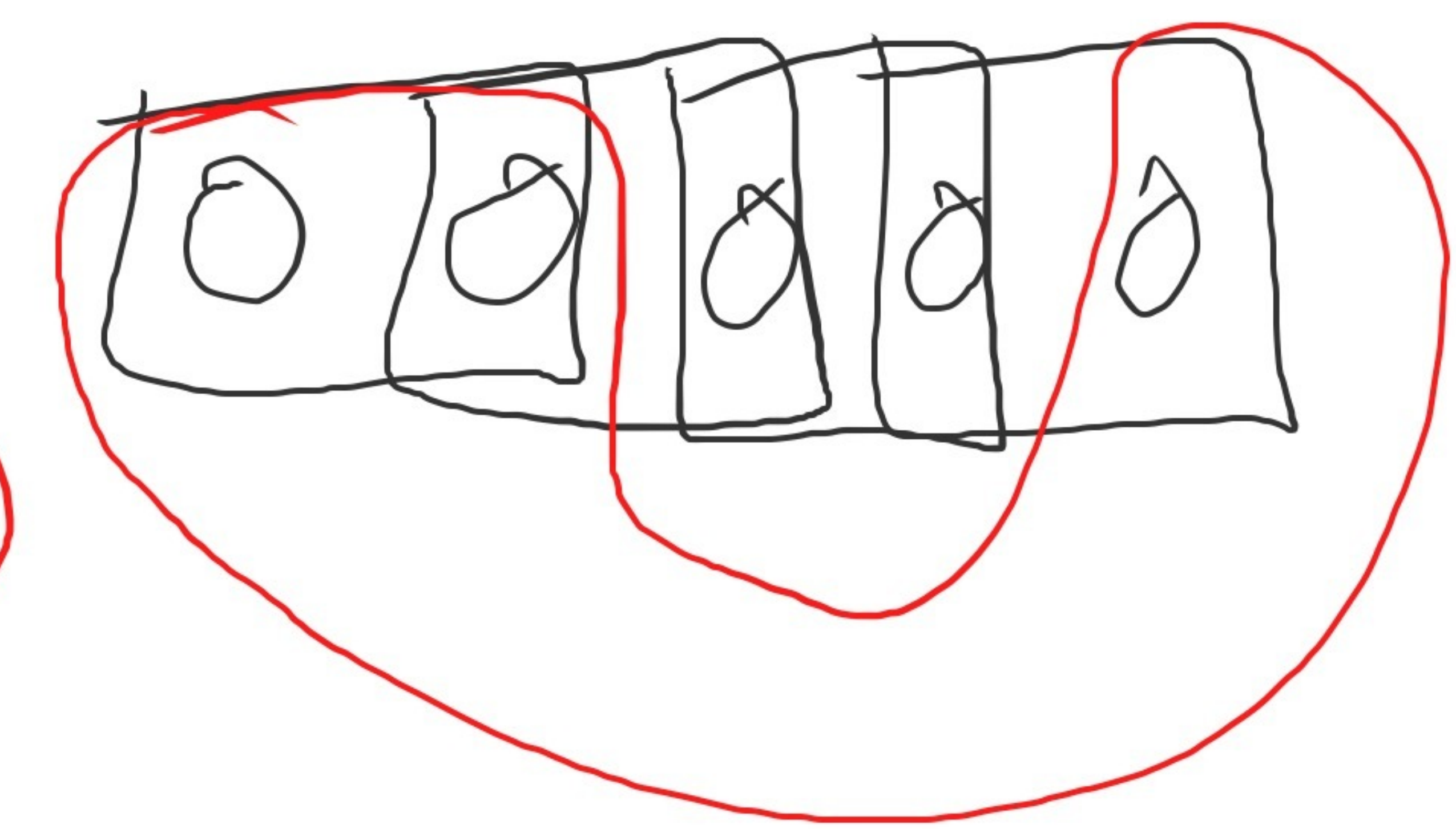
work₁

↓ bit



Sparsity
tying

prior
learning bias



n^2 weights



$\bar{N} \rightarrow AN$

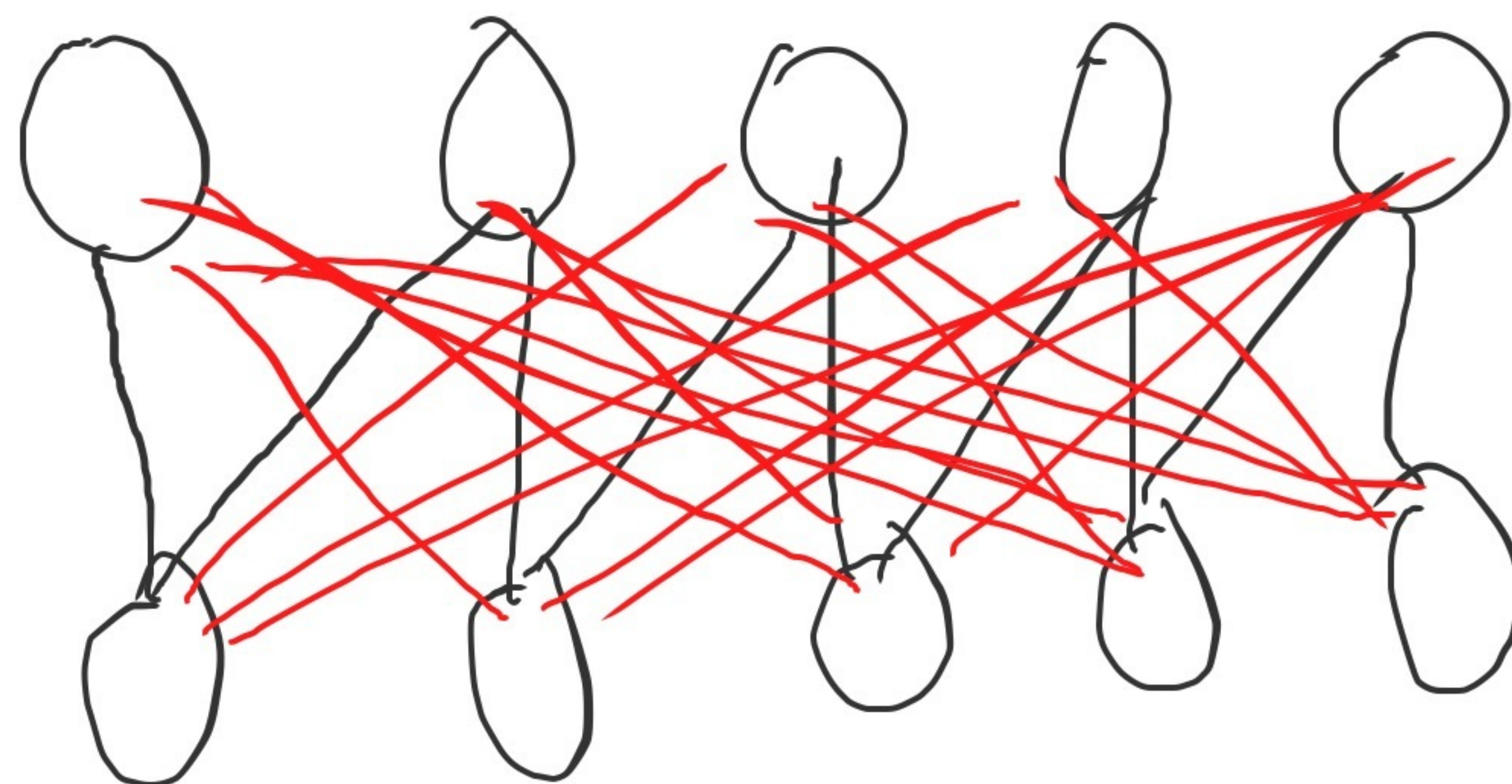
$\bar{N} \rightarrow NP$

~~what ball~~ did you buy

ball in the court

I bought a red ball

$NP \rightarrow DN$
locality



sparsity

locality

— zero

tying

$W=0$

prior

learning bias

the ball

$NP \rightarrow D N$

the ball is red

John caught

the ball

John threw

the ball

into the basket

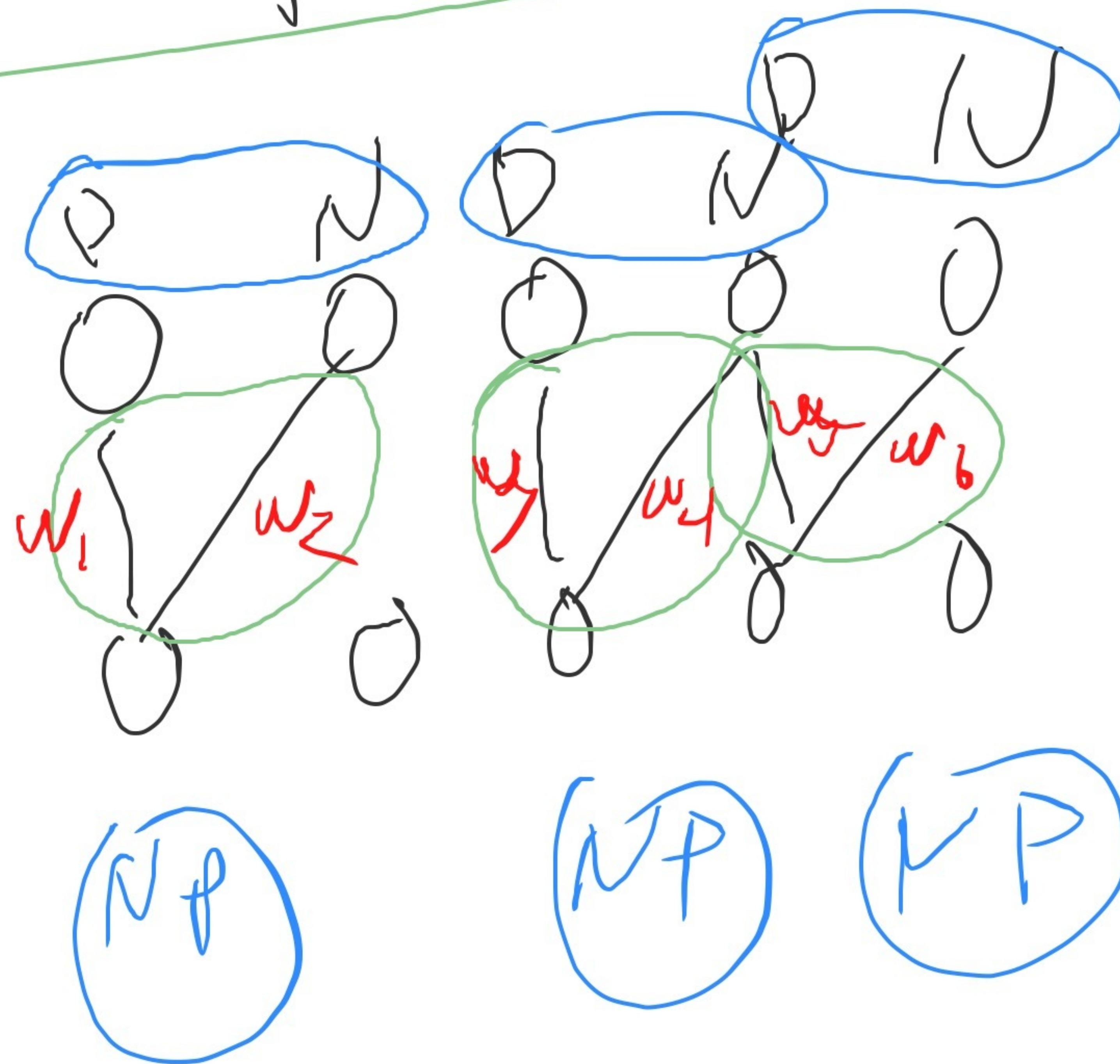
$S \rightarrow NP VP$

$VP \rightarrow V NP PP$

$VP \rightarrow V XP$

positive independence criterion

locality

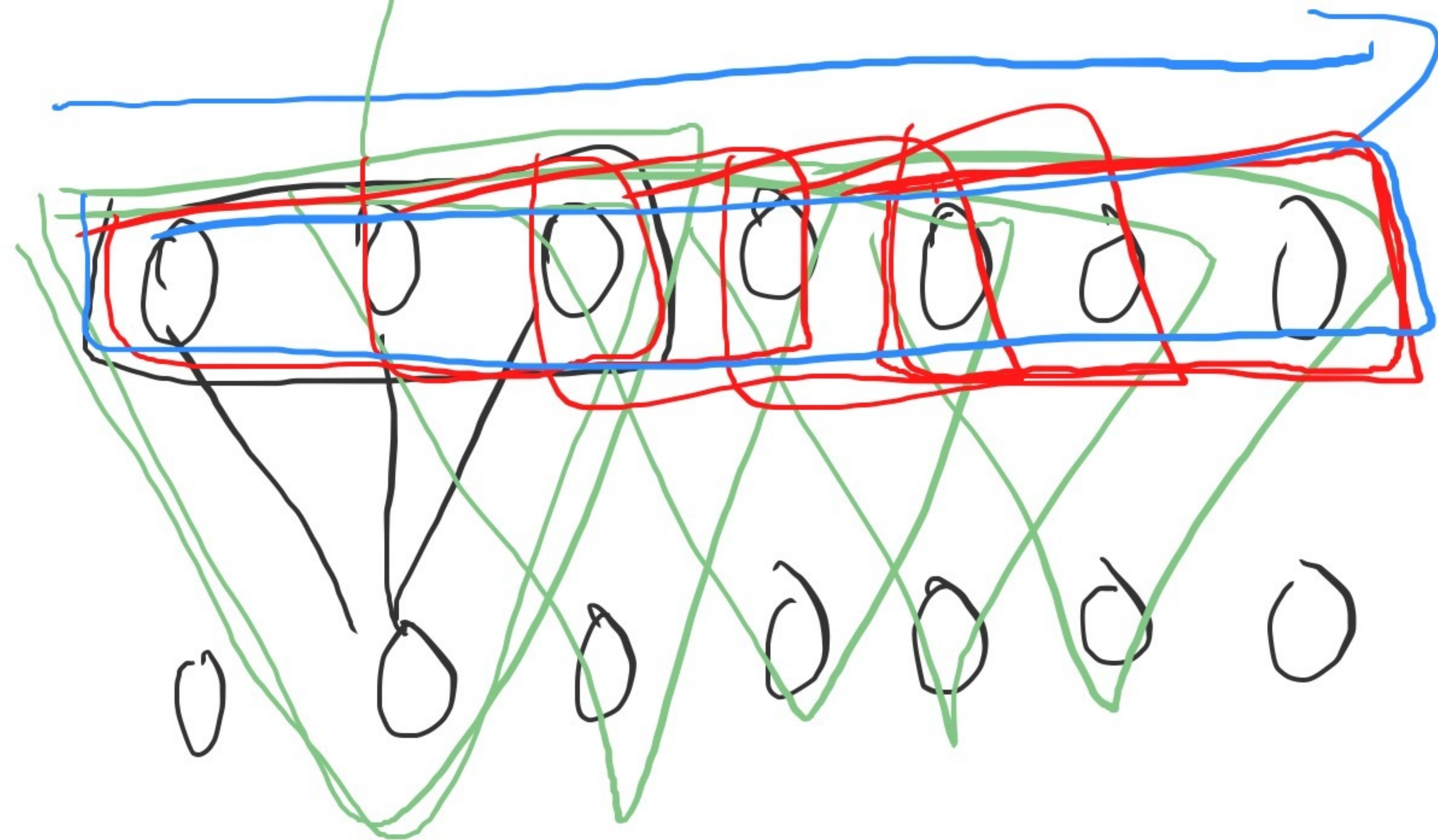


$$w_1 = w_3 = w_5$$
$$w_2 = w_4 = w_6$$

tying
sparsity

dot product

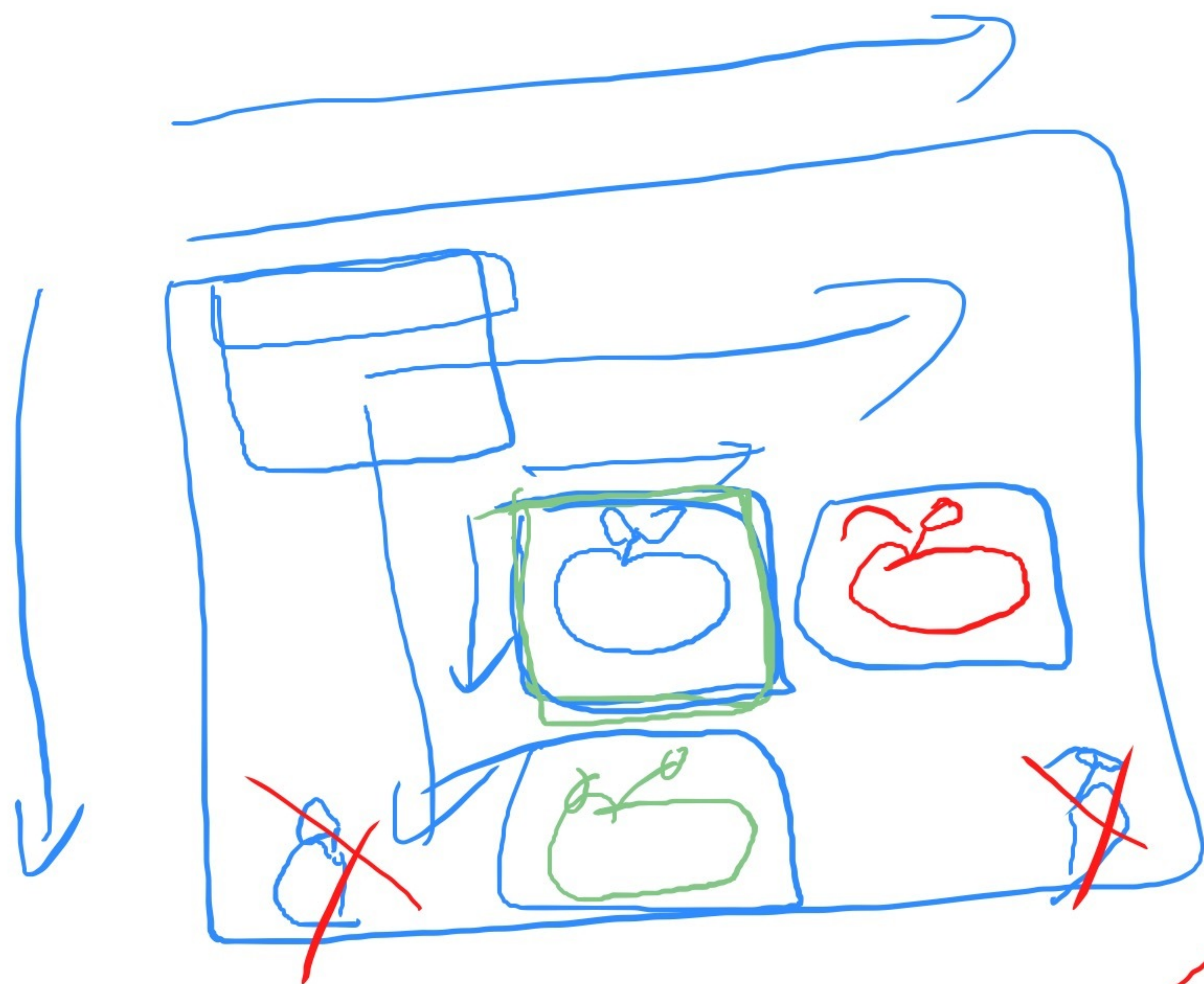
$$\vec{w} \cdot \vec{x}$$



↓ dimension

convolution

1D convolutional neural network



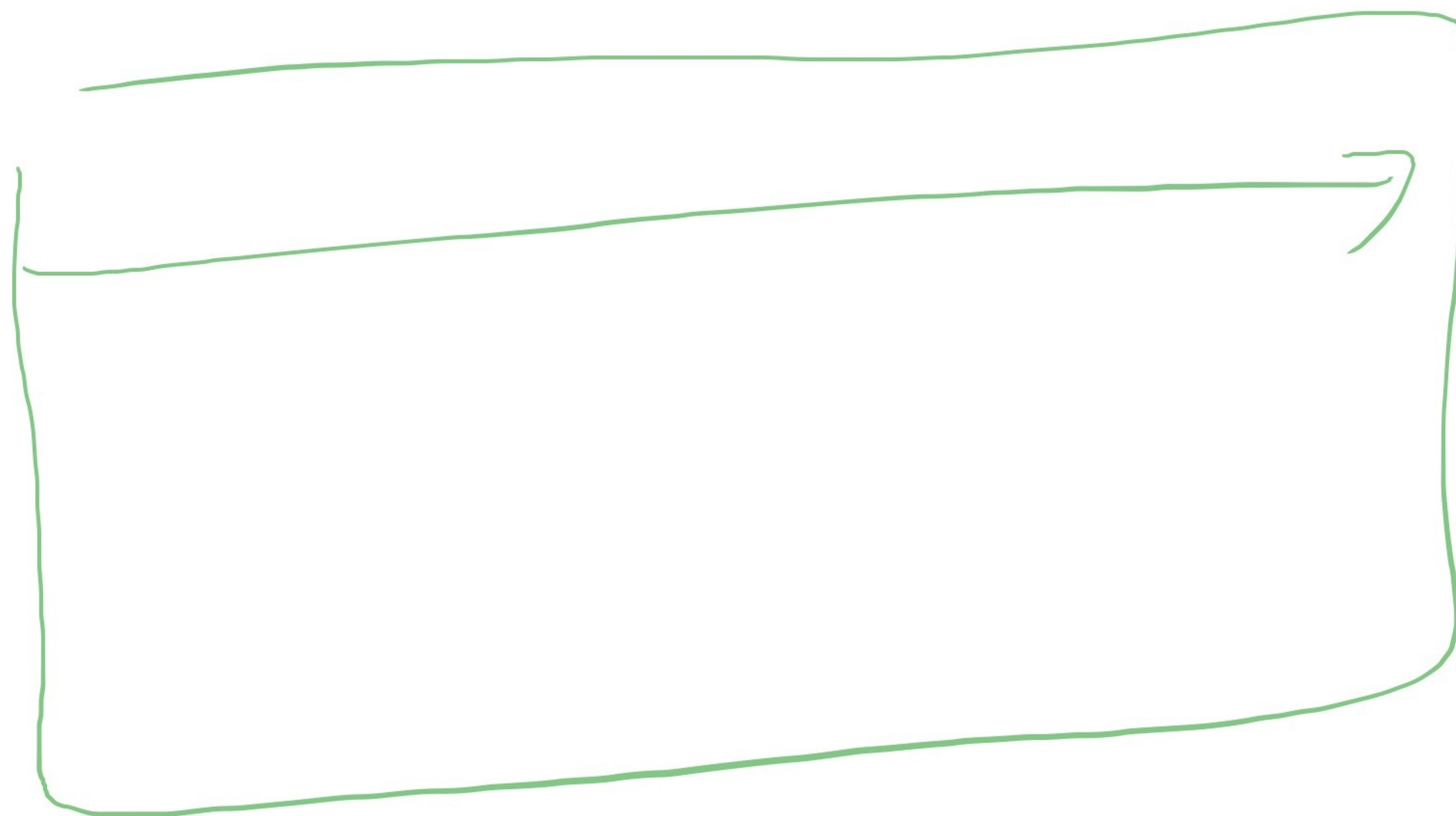
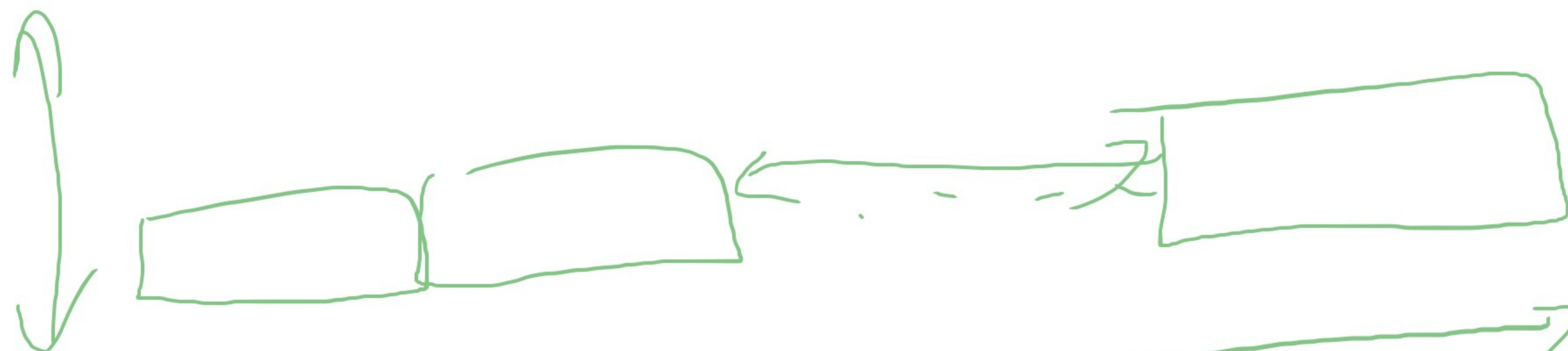
locality

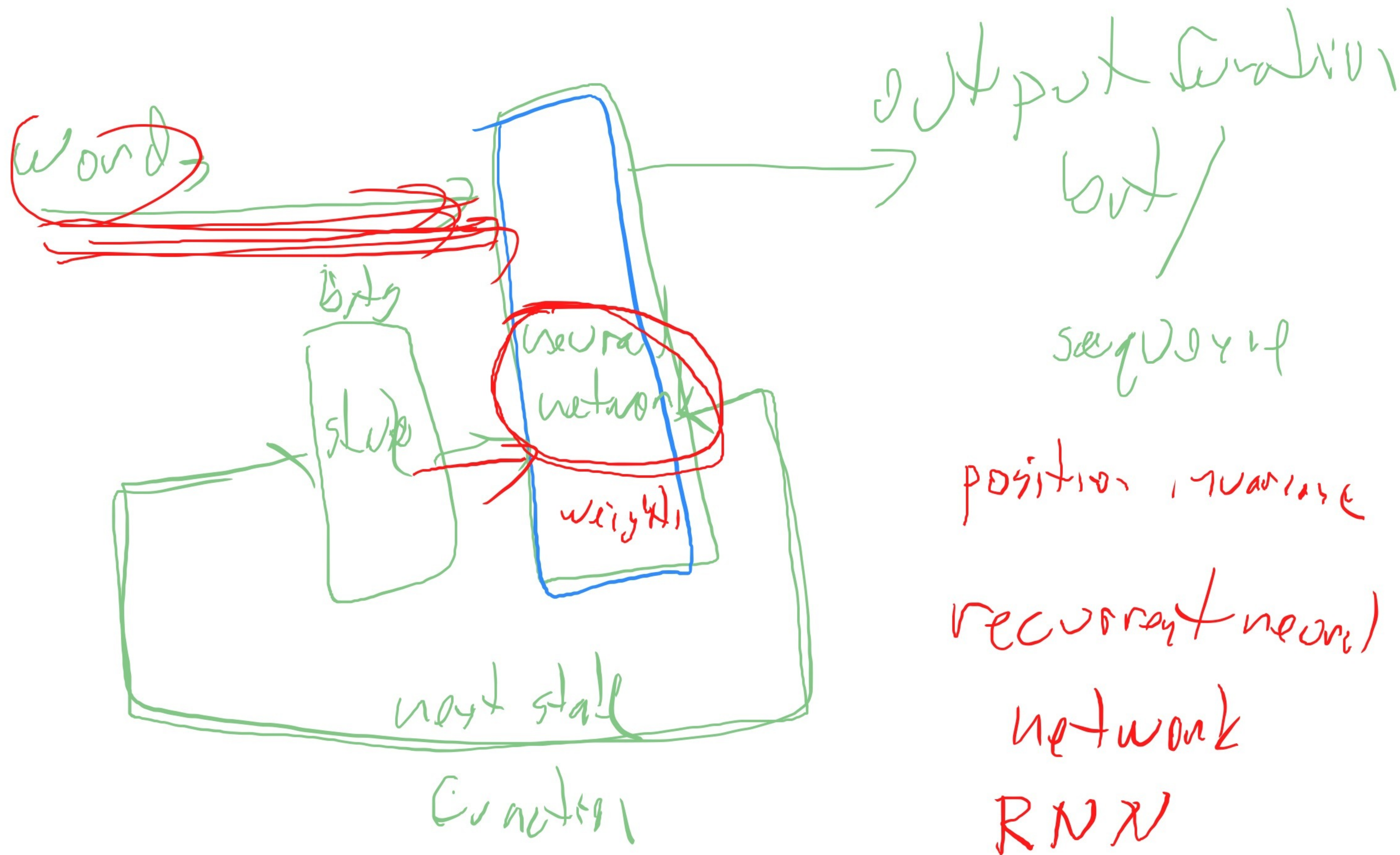
translation invariance

2D CNN

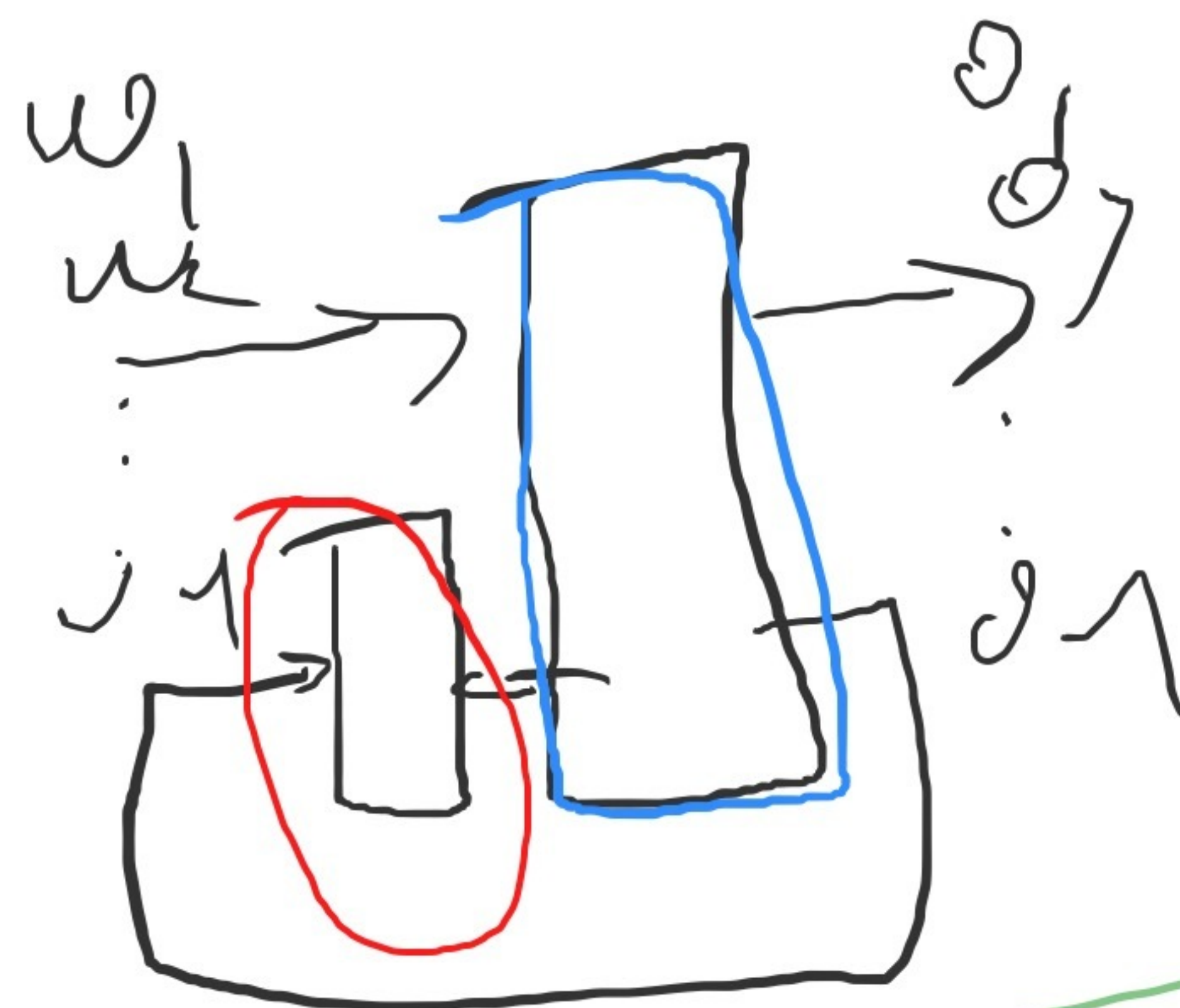
3D CNN



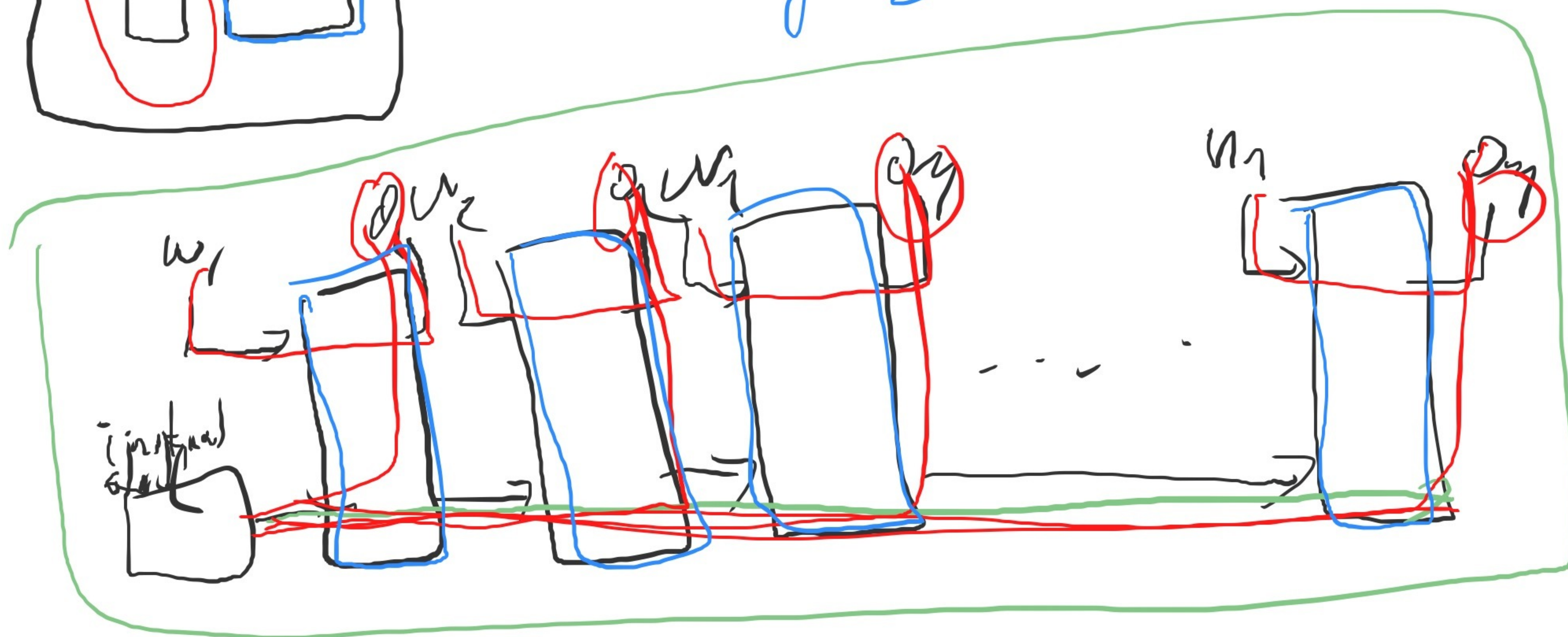




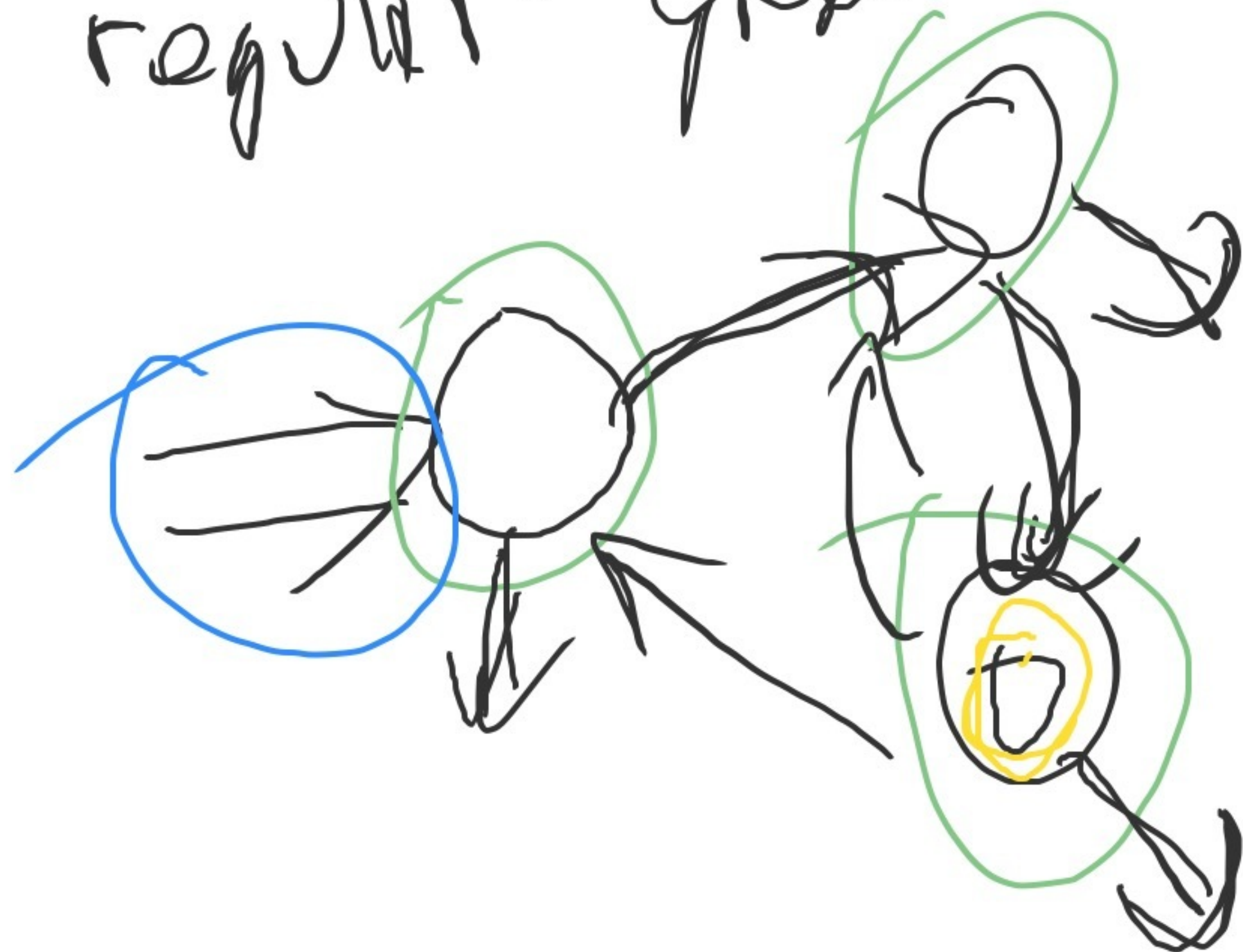
unrolling



tying

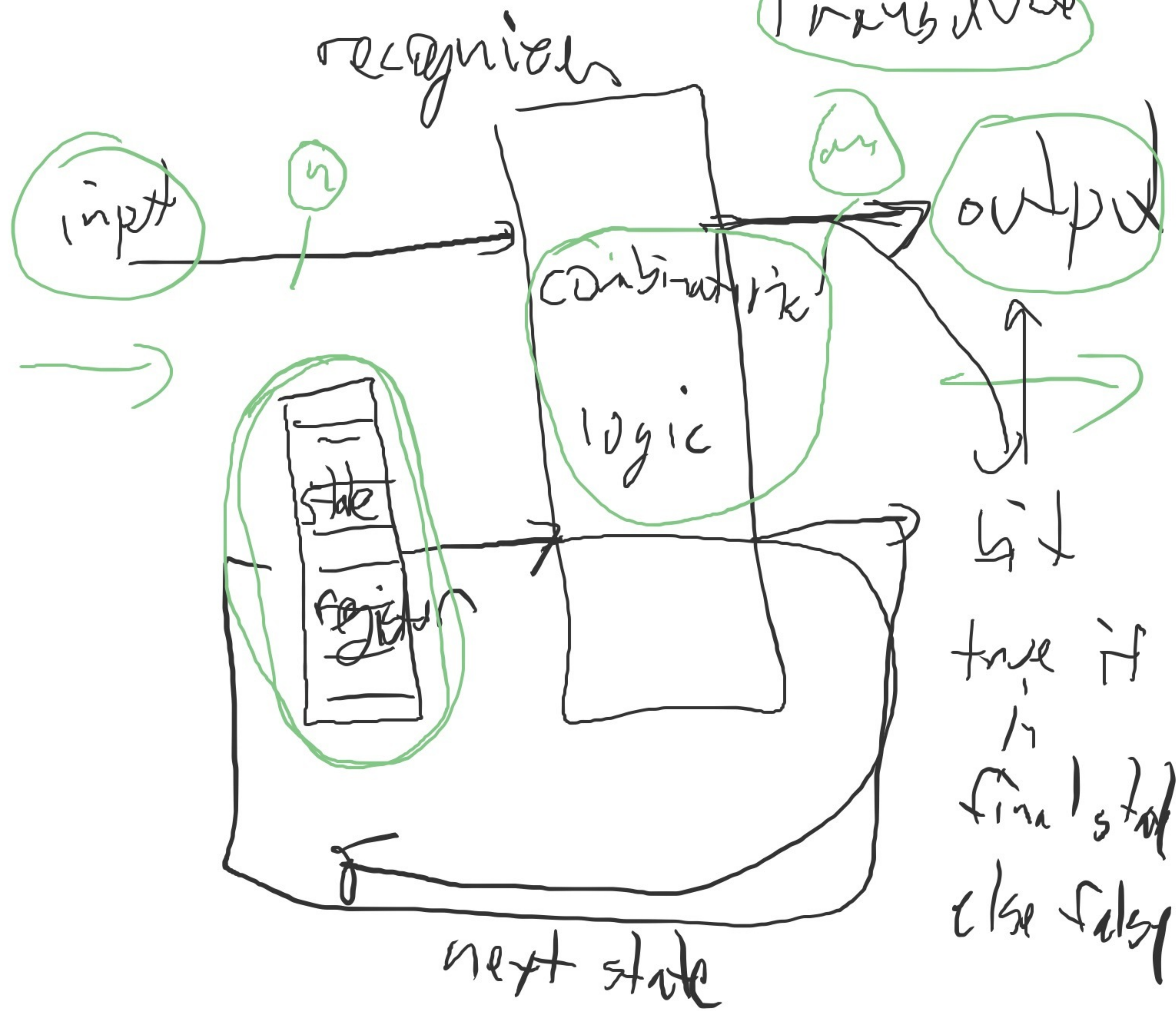


regular grammar



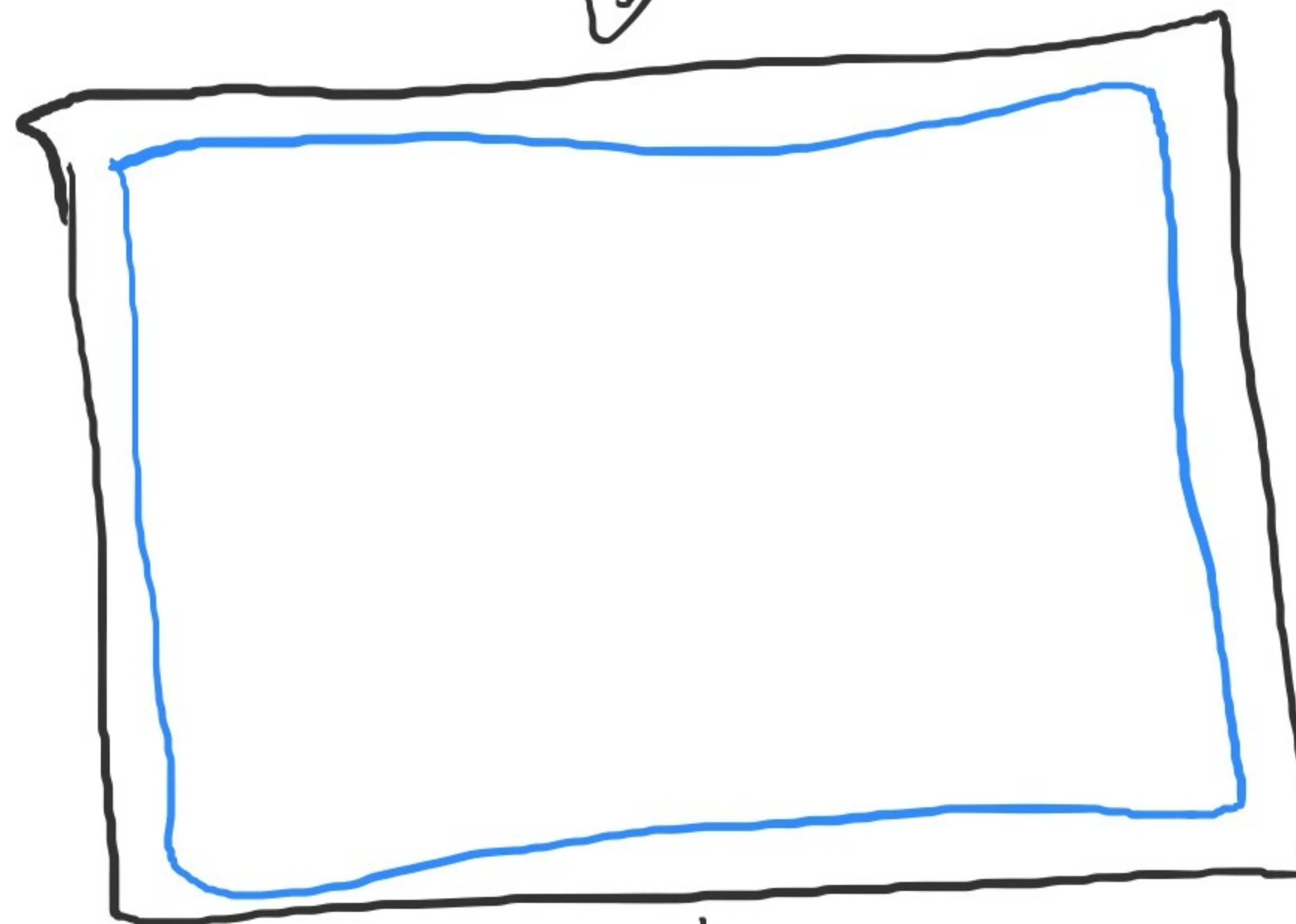
HMM

Finite State Machine Transducer



recognizer
transducer

~~sequence~~

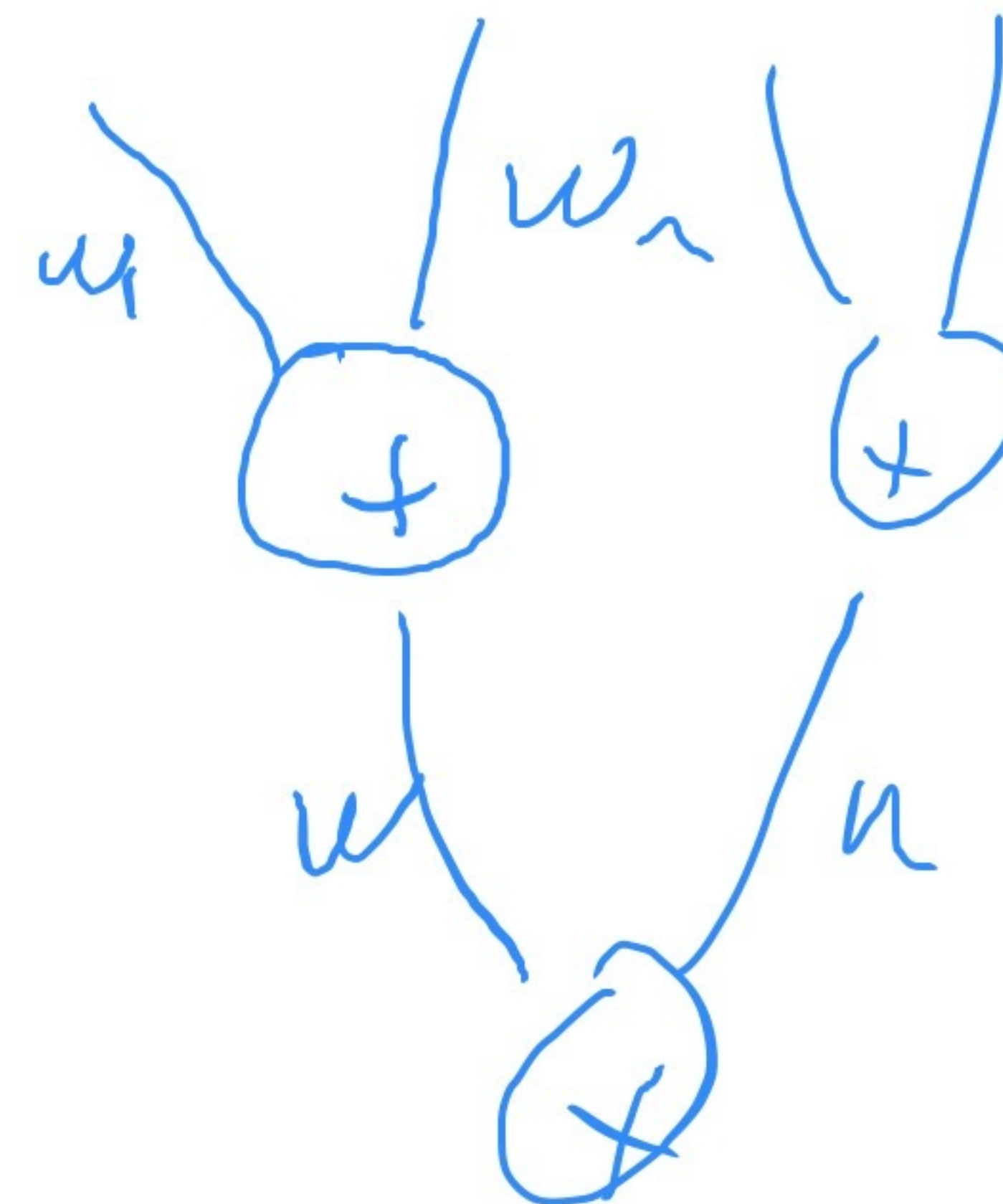


output ~~bit~~

true

false

sequence



wordy K

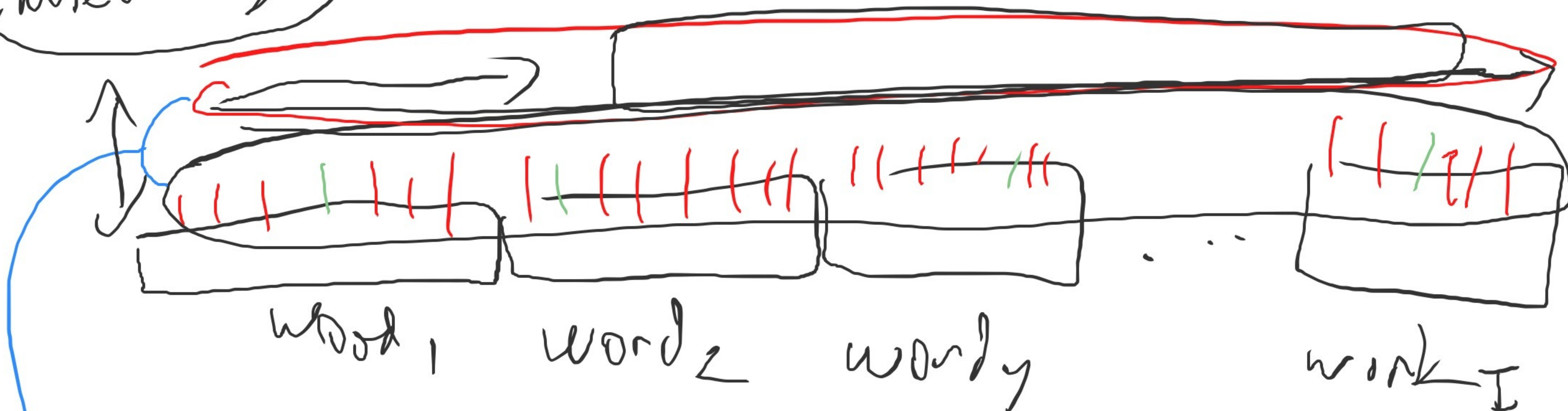
(inference)

(embeddings)

one hot
1 0 0 0 1 0 ... 0

↓
K

K



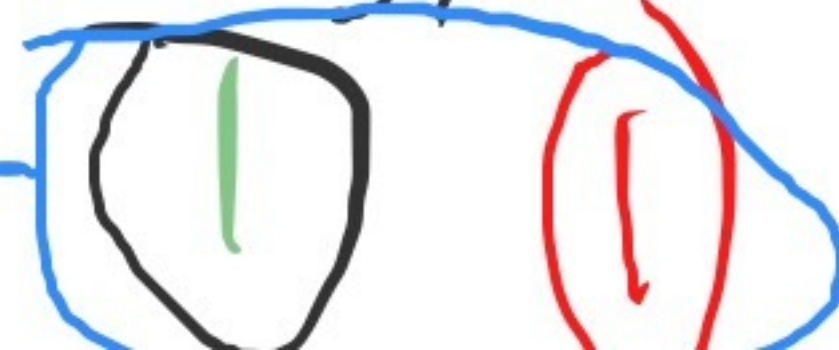
word₁

word₂

word_y

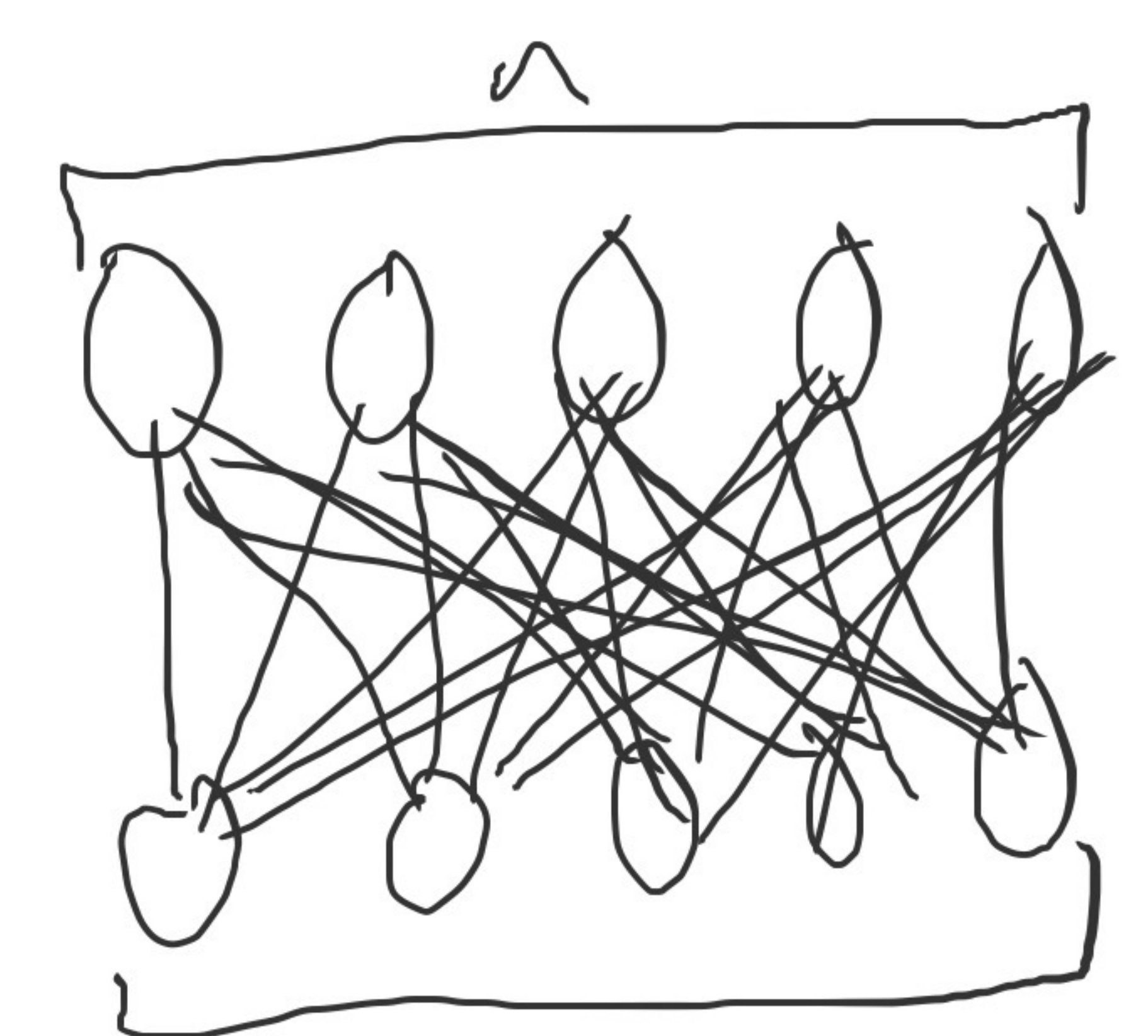
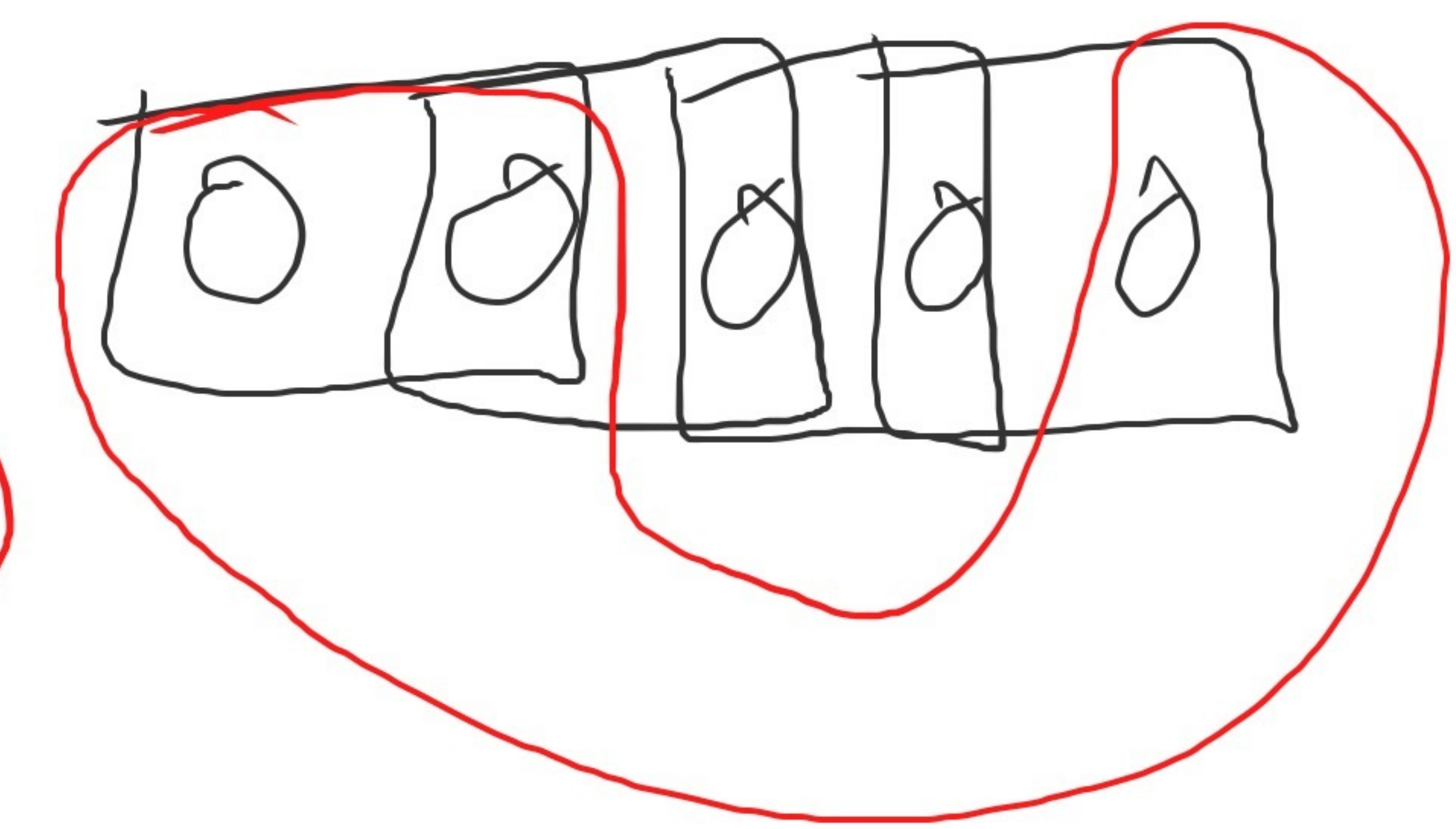
work₁

↓ bit



Sparsity
tying

prior
learning bias



n^2 weights



$\bar{N} \rightarrow AN$

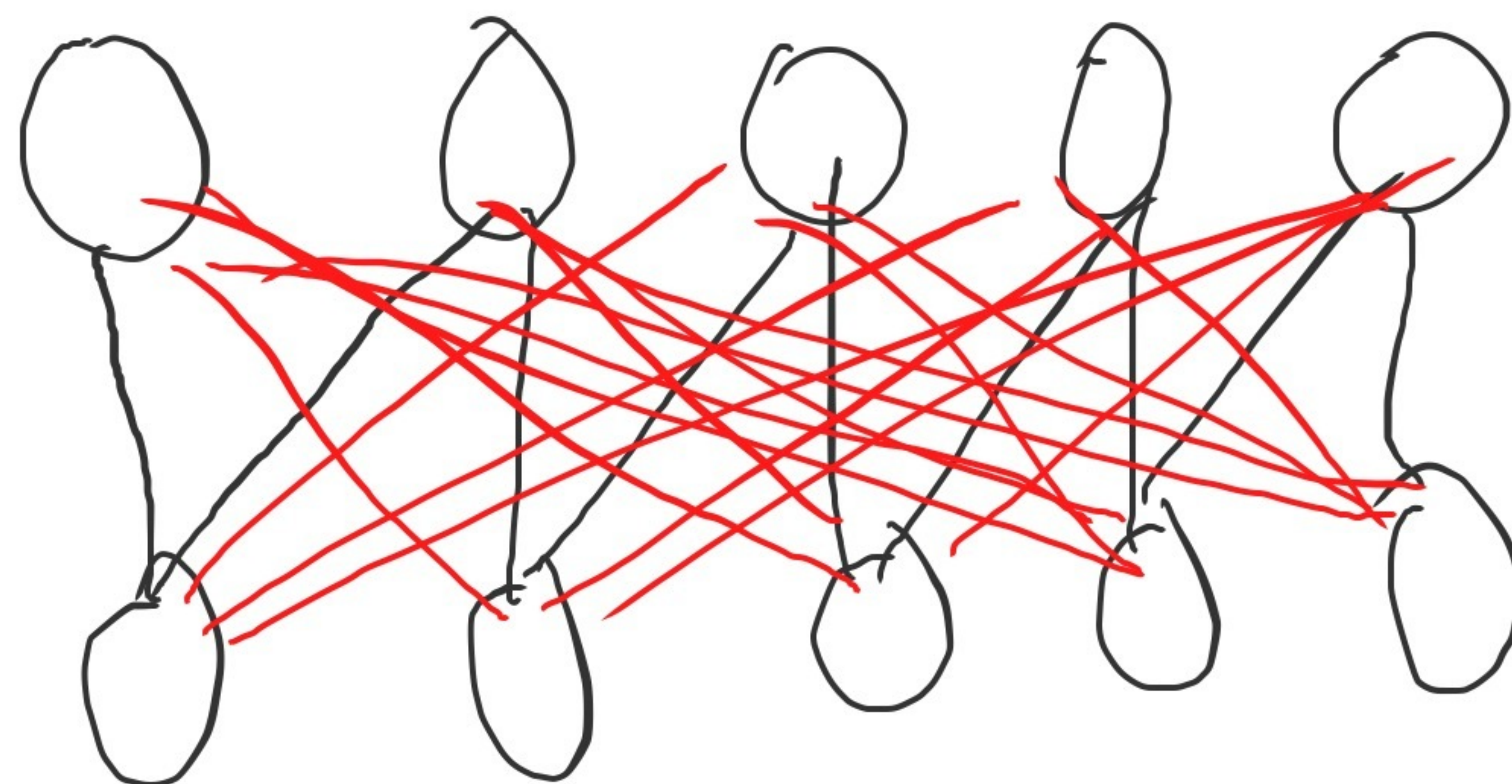
$\bar{N} \rightarrow NP$

~~what ball~~ did you buy

ball in the court

I bought a red ball

$NP \rightarrow DN$
locality



sparsity

locality

— zero

tying

$W=0$

prior

learning bias

the ball

$NP \rightarrow D N$

the ball is red

John caught the ball

John threw the ball into the basket

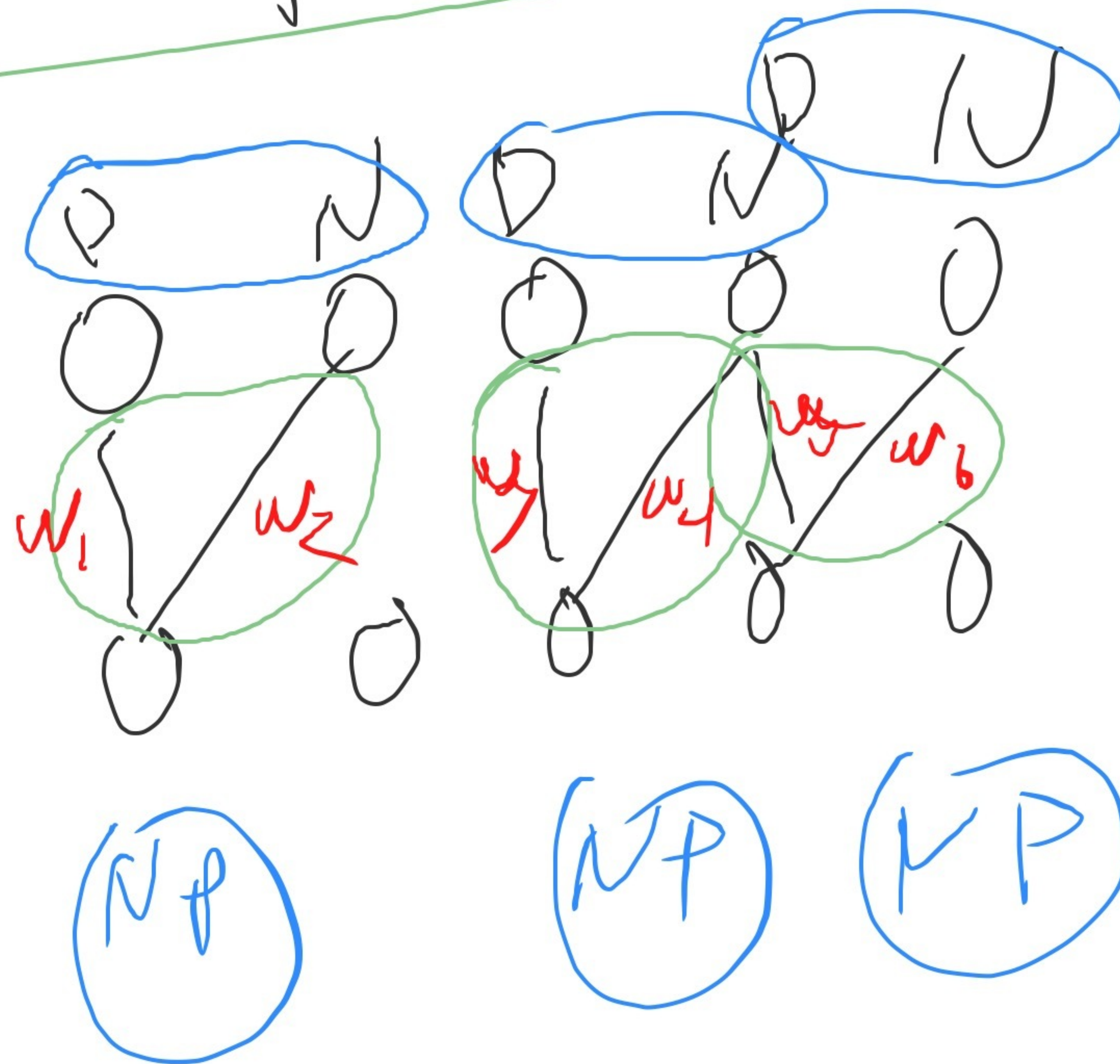
$S \rightarrow NP VP$

$VP \rightarrow V NP PP$

$VP \rightarrow V XP$

positive independence criterion

locality

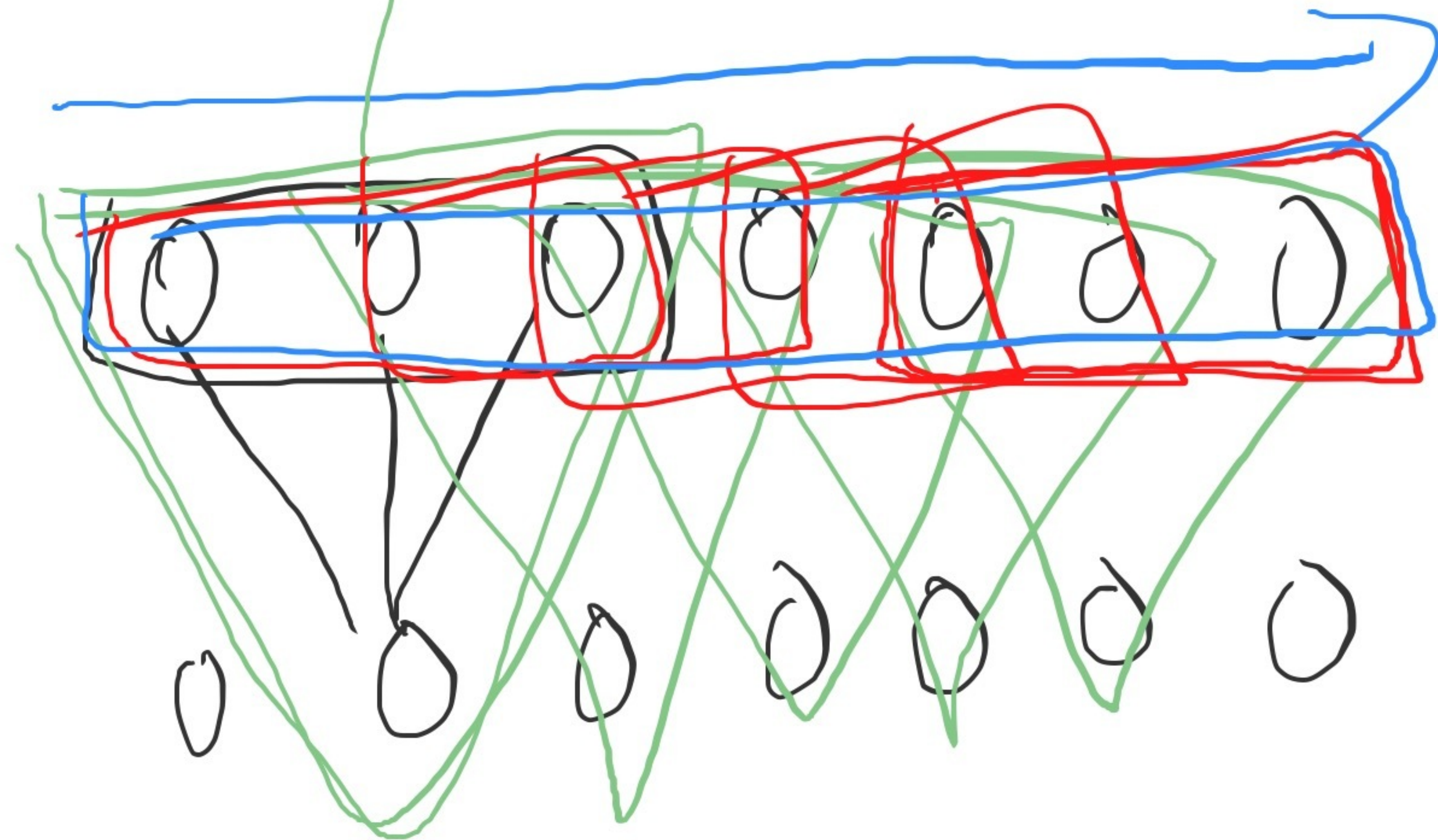


$$w_1 = w_3 = w_5$$
$$w_2 = w_4 = w_6$$

tying
sparsity

dot product

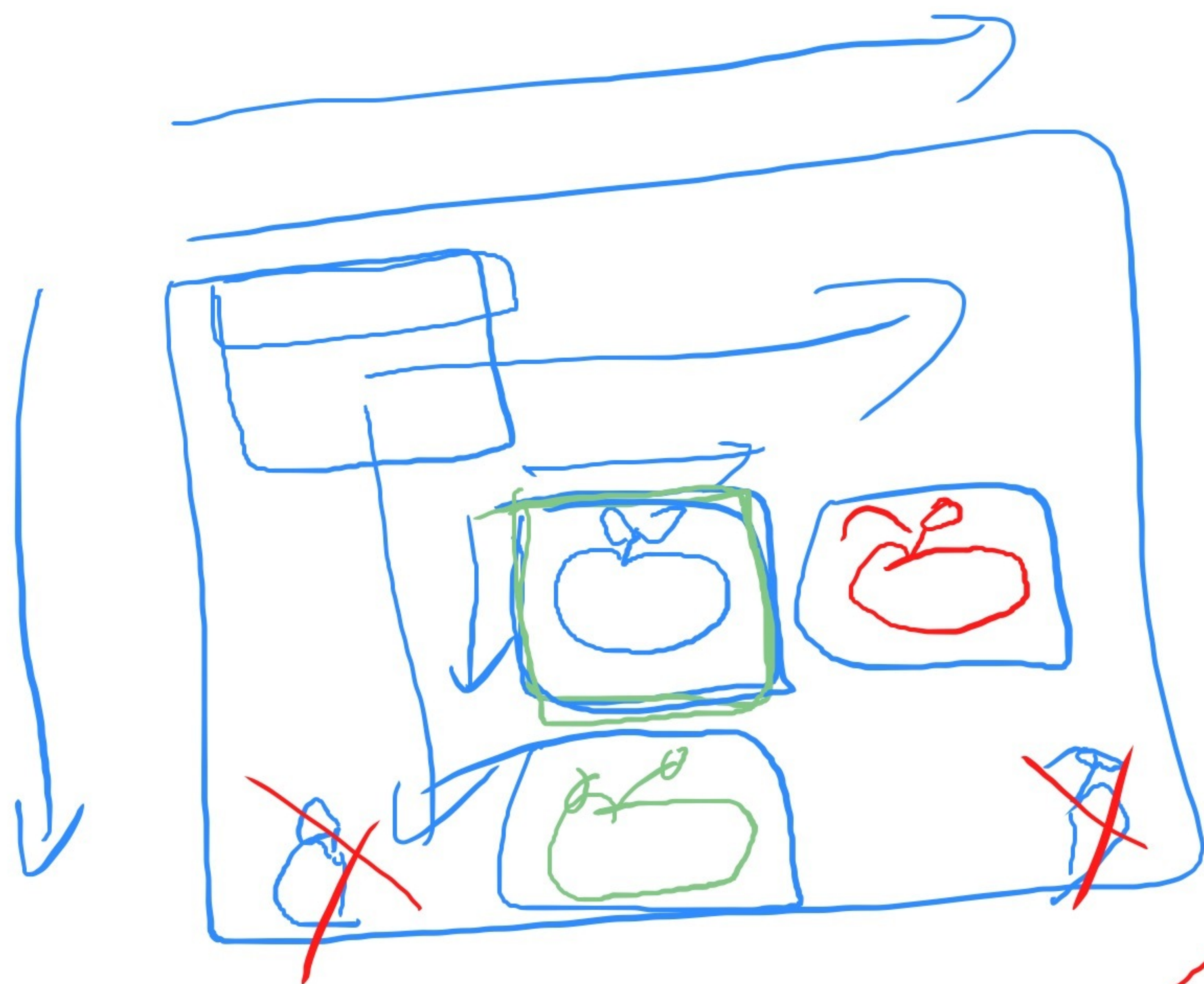
$$\vec{w} \cdot \vec{x}$$



↓ dimension

convolution

1D convolutional neural network



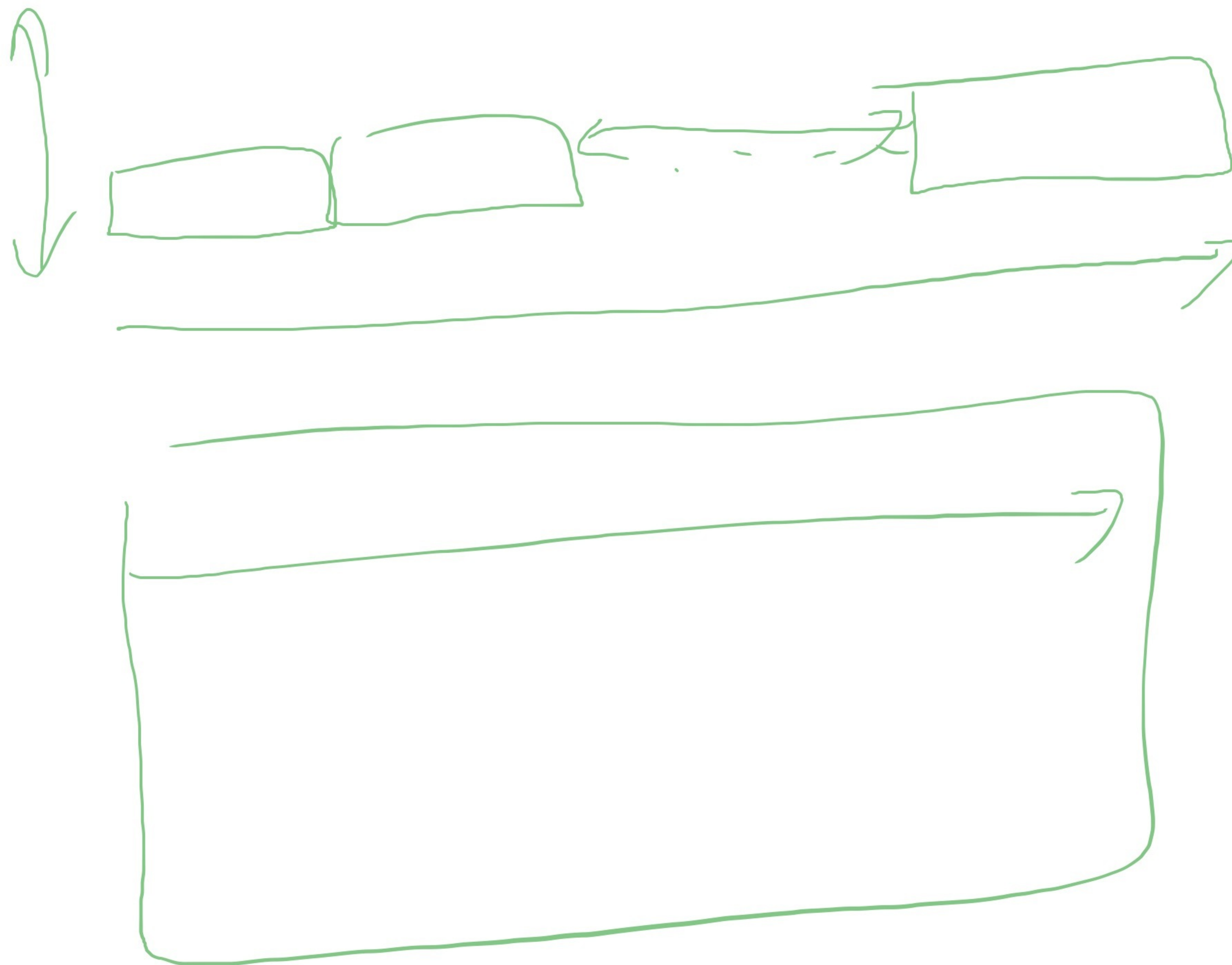
locality

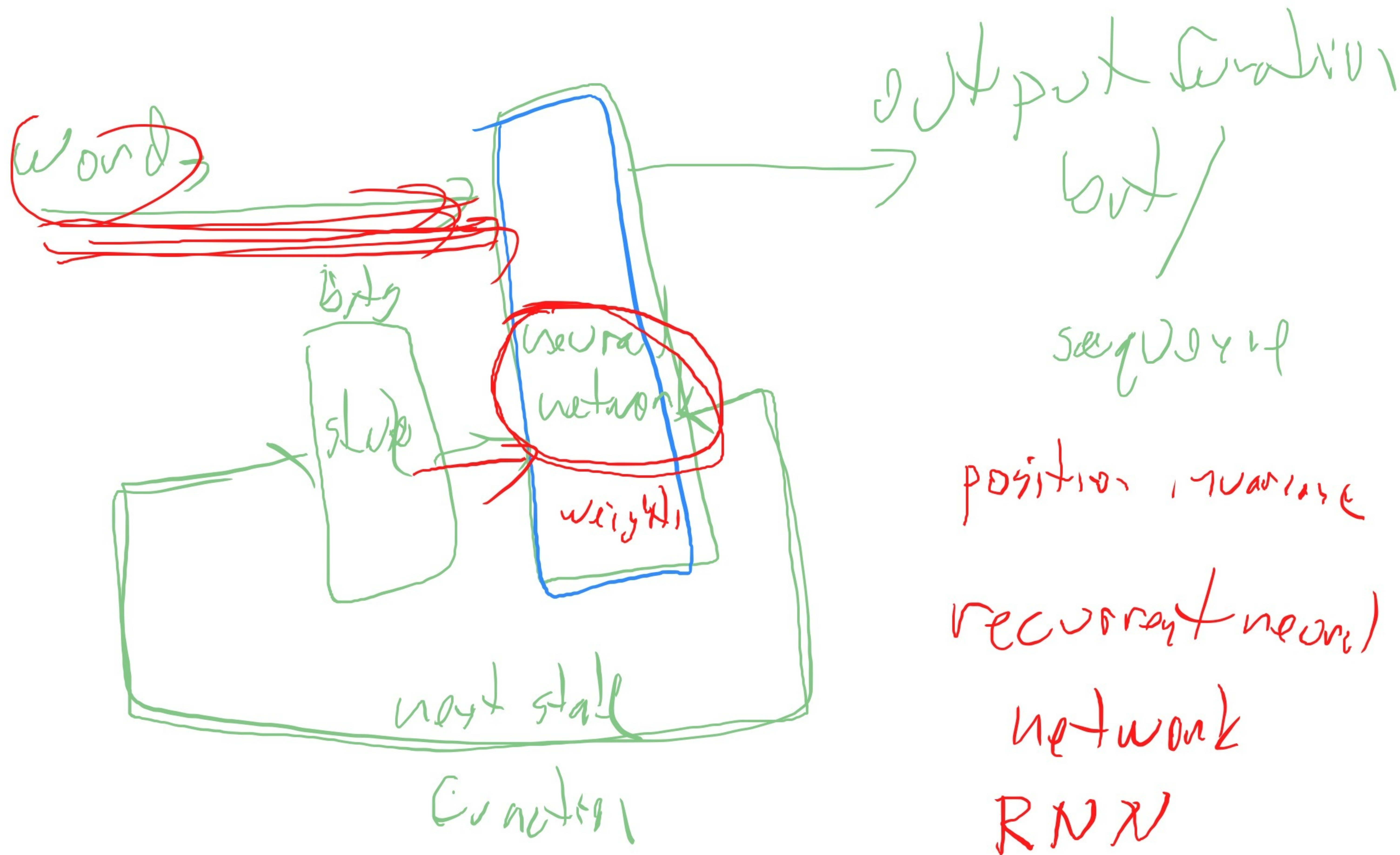
position
translation
invariance

2D CNN

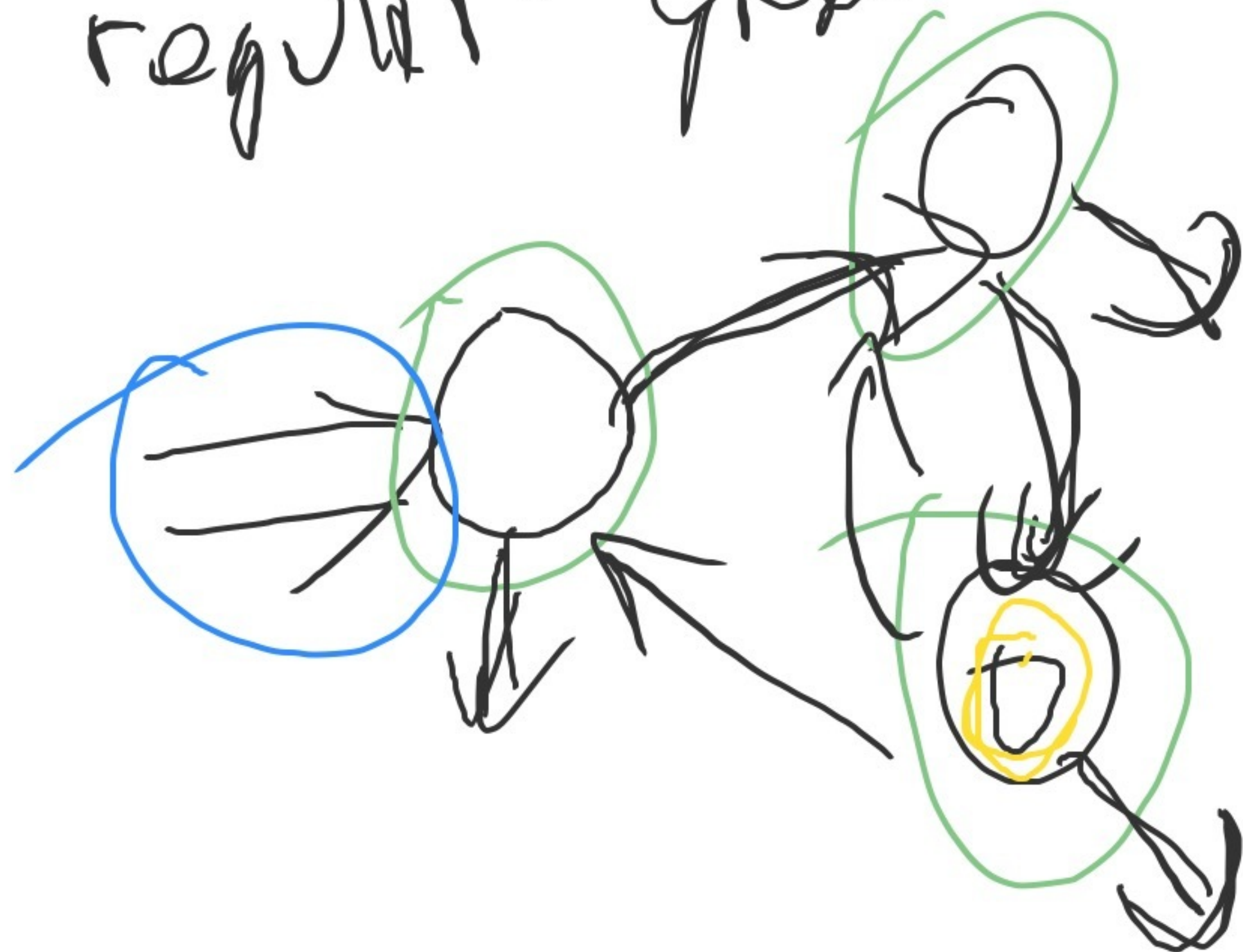
3D CNN







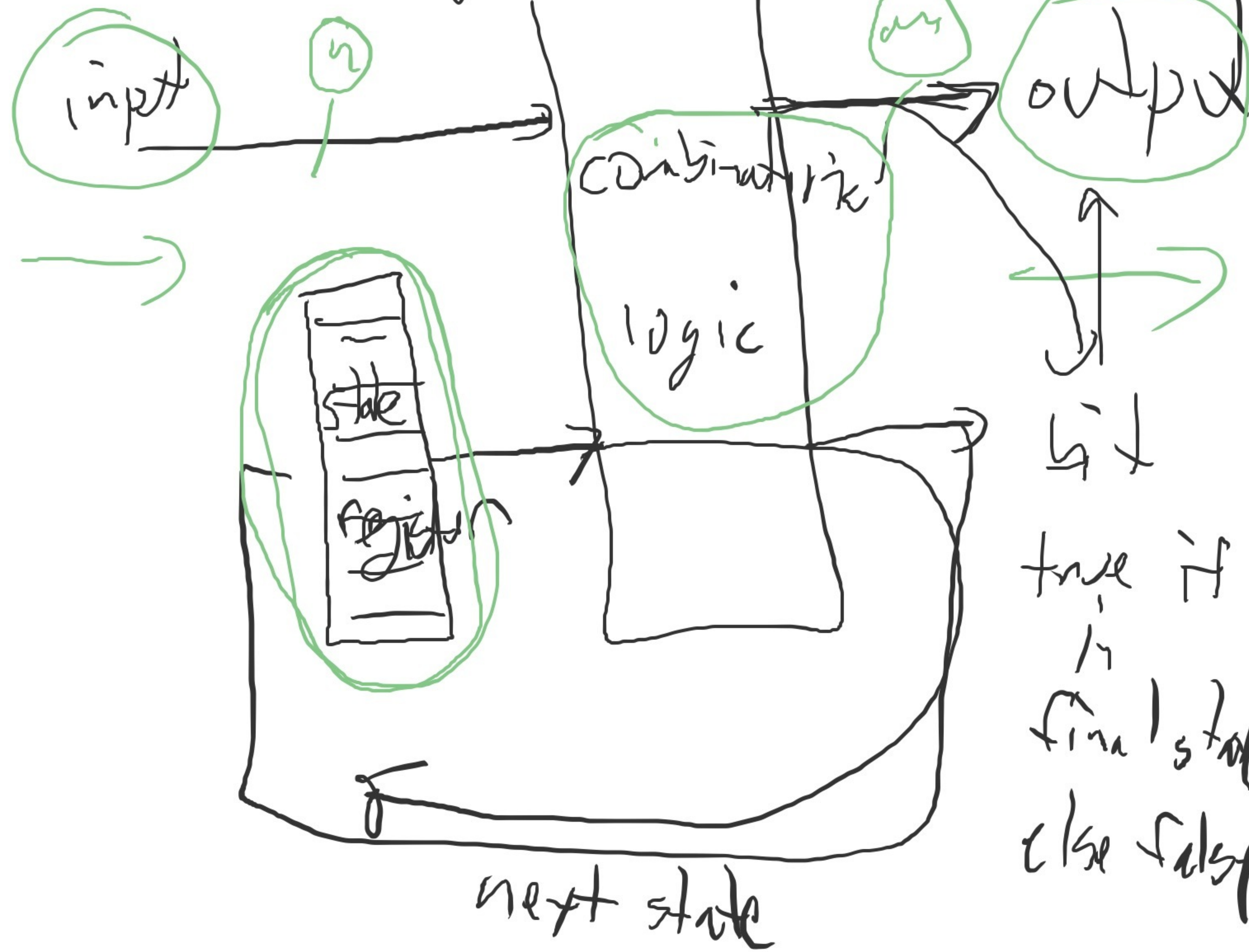
regular grammar



HMM

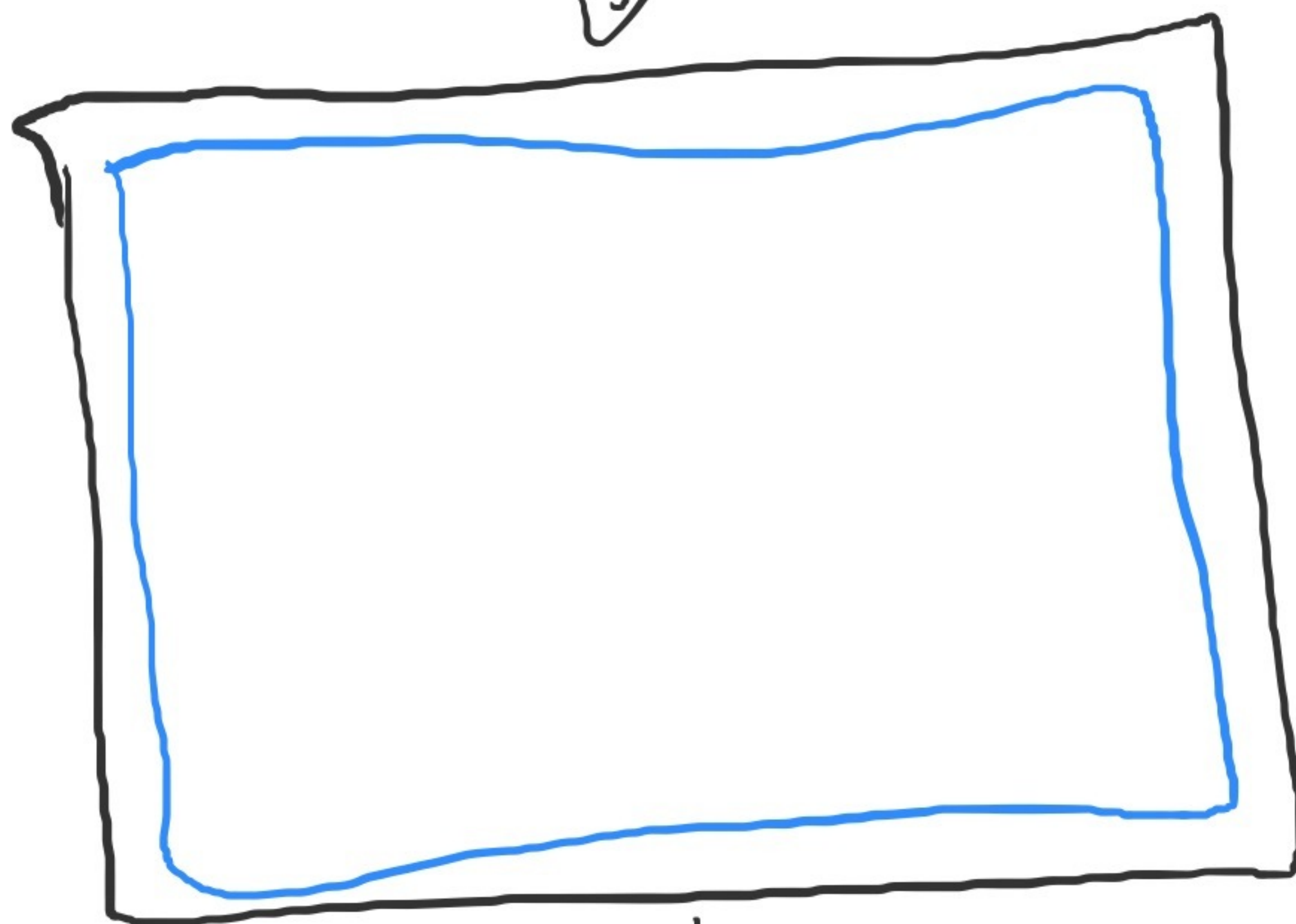
Finite State Machine Transducer

recognizer



recognizer
transducer

~~sequence~~

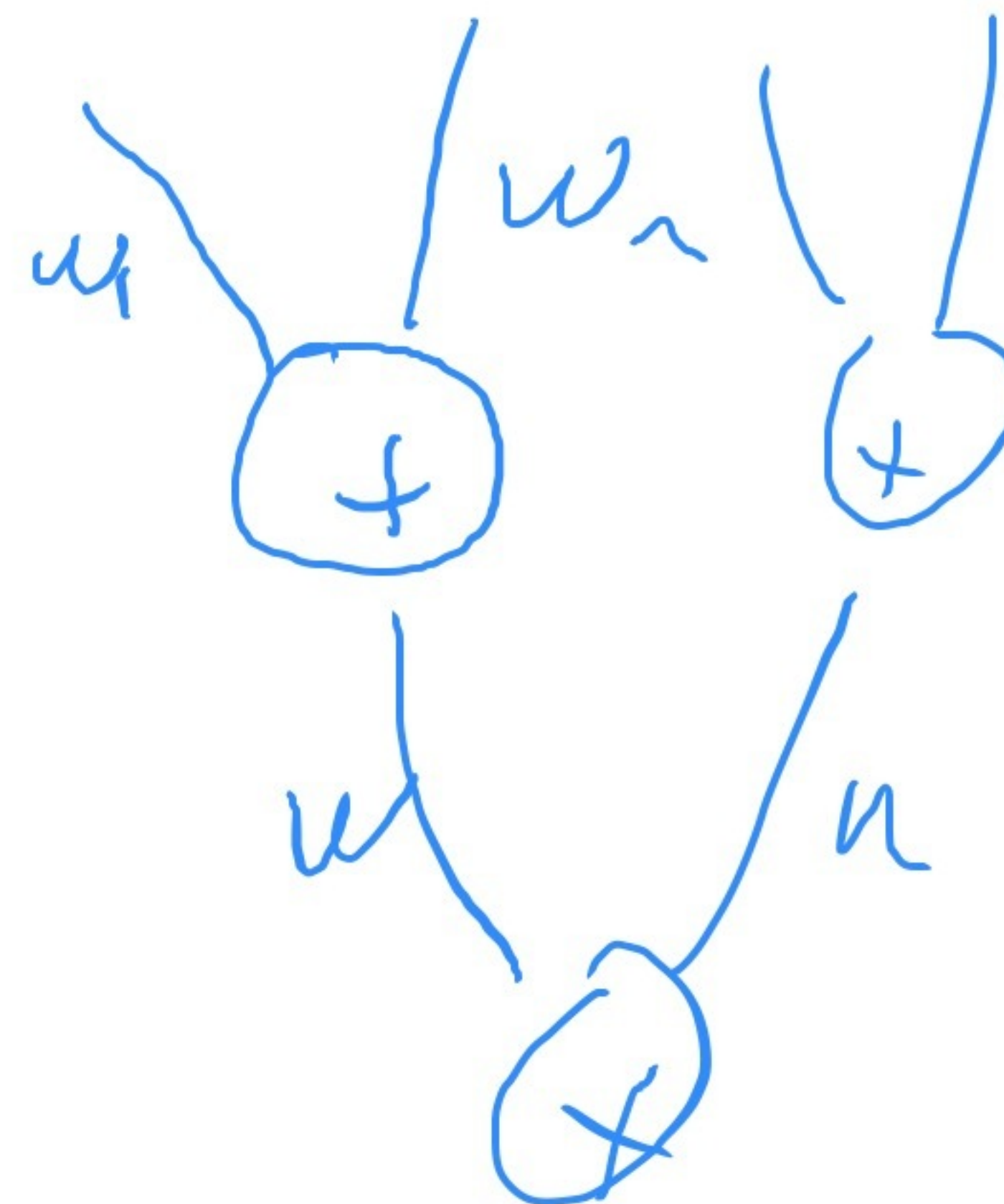


output ~~bit~~

true

false

sequence



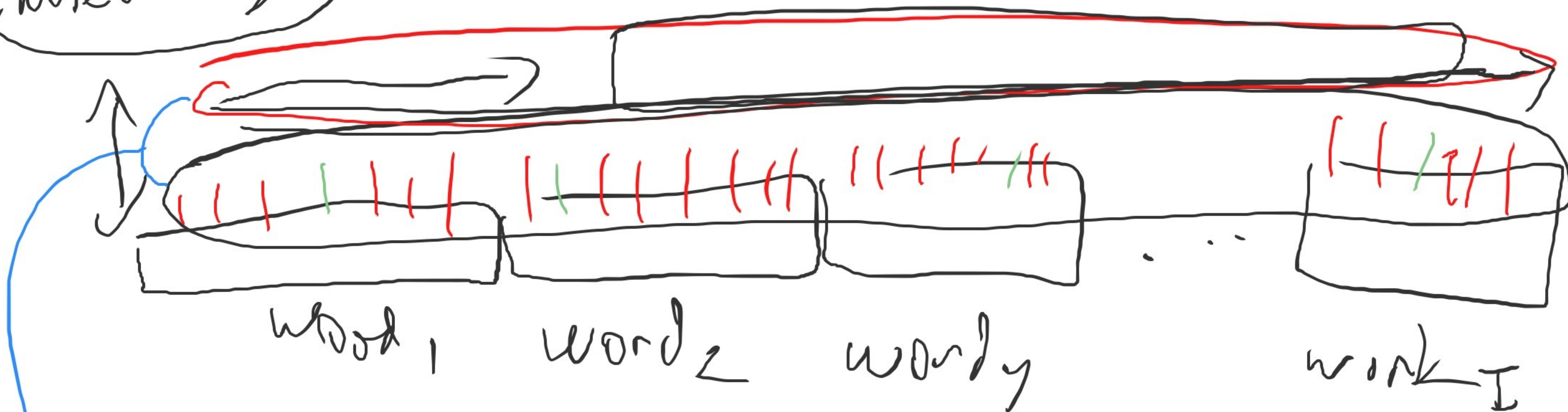
wordy K

(inference)

(embeddings)

one hot
1 0 0 0 1 0 ... 0

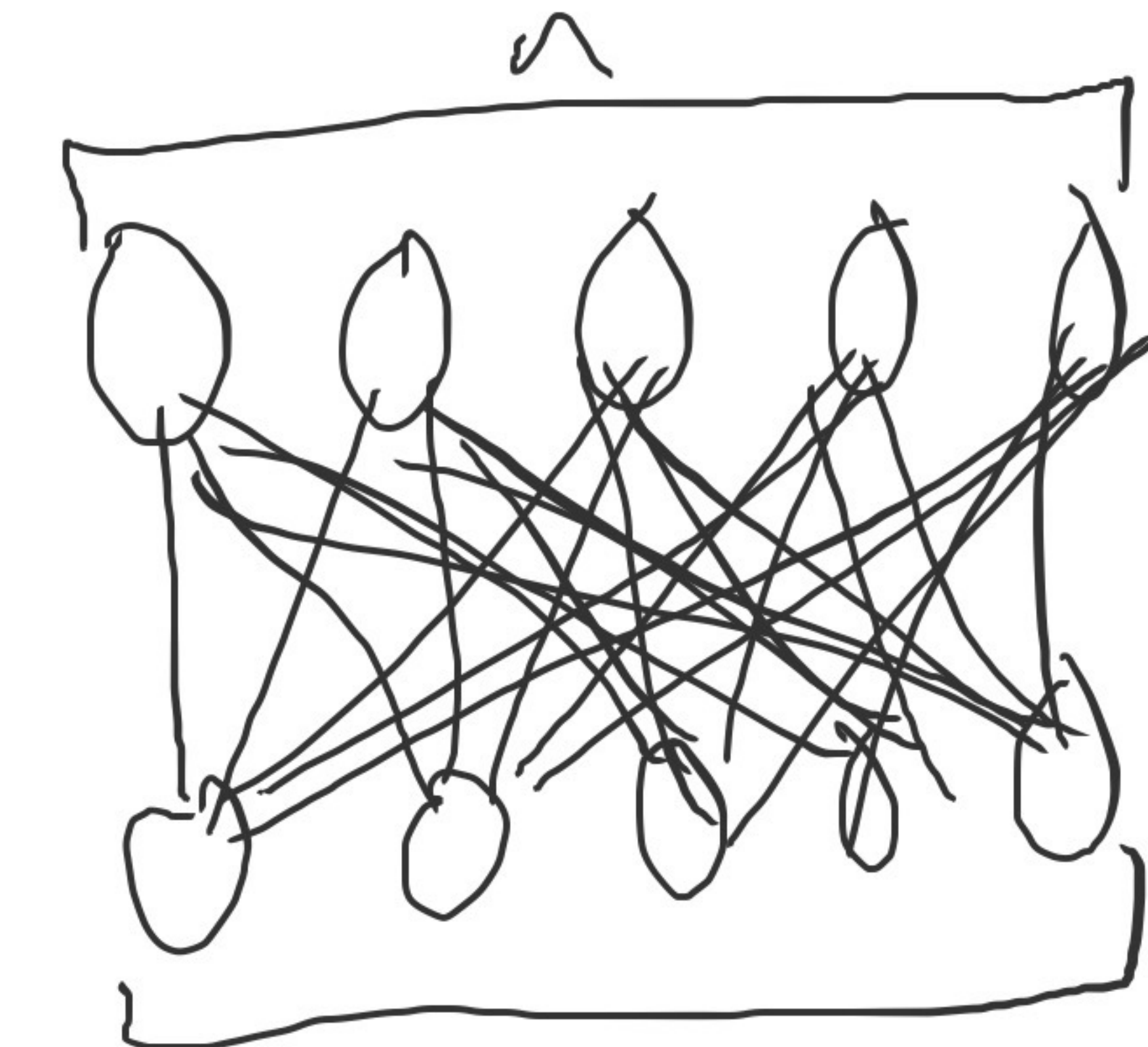
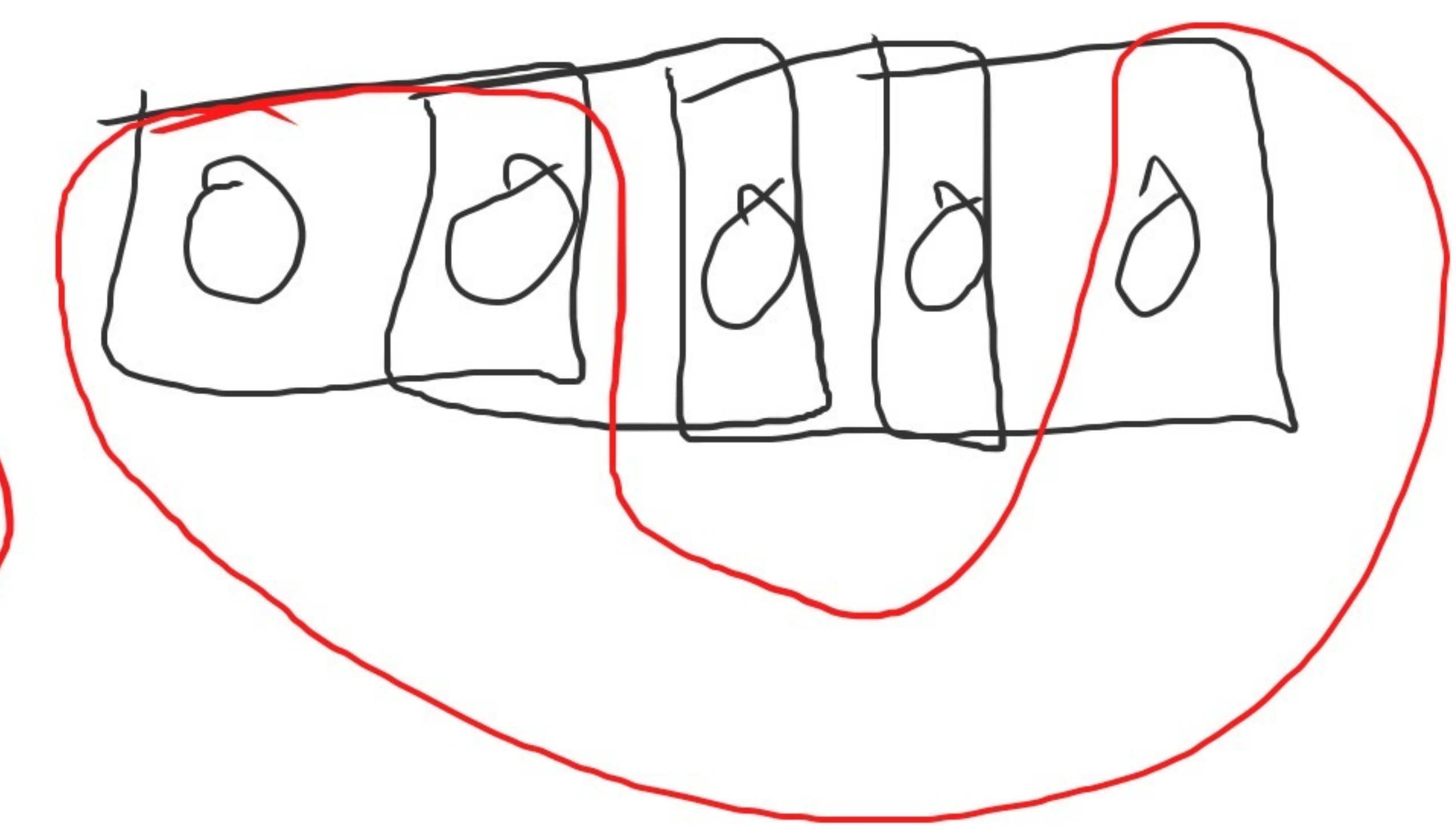
↓
K



↓ bit
1 1

Sparsity
tying

prior
learning bias



n^2 weights



$\bar{N} \rightarrow AN$

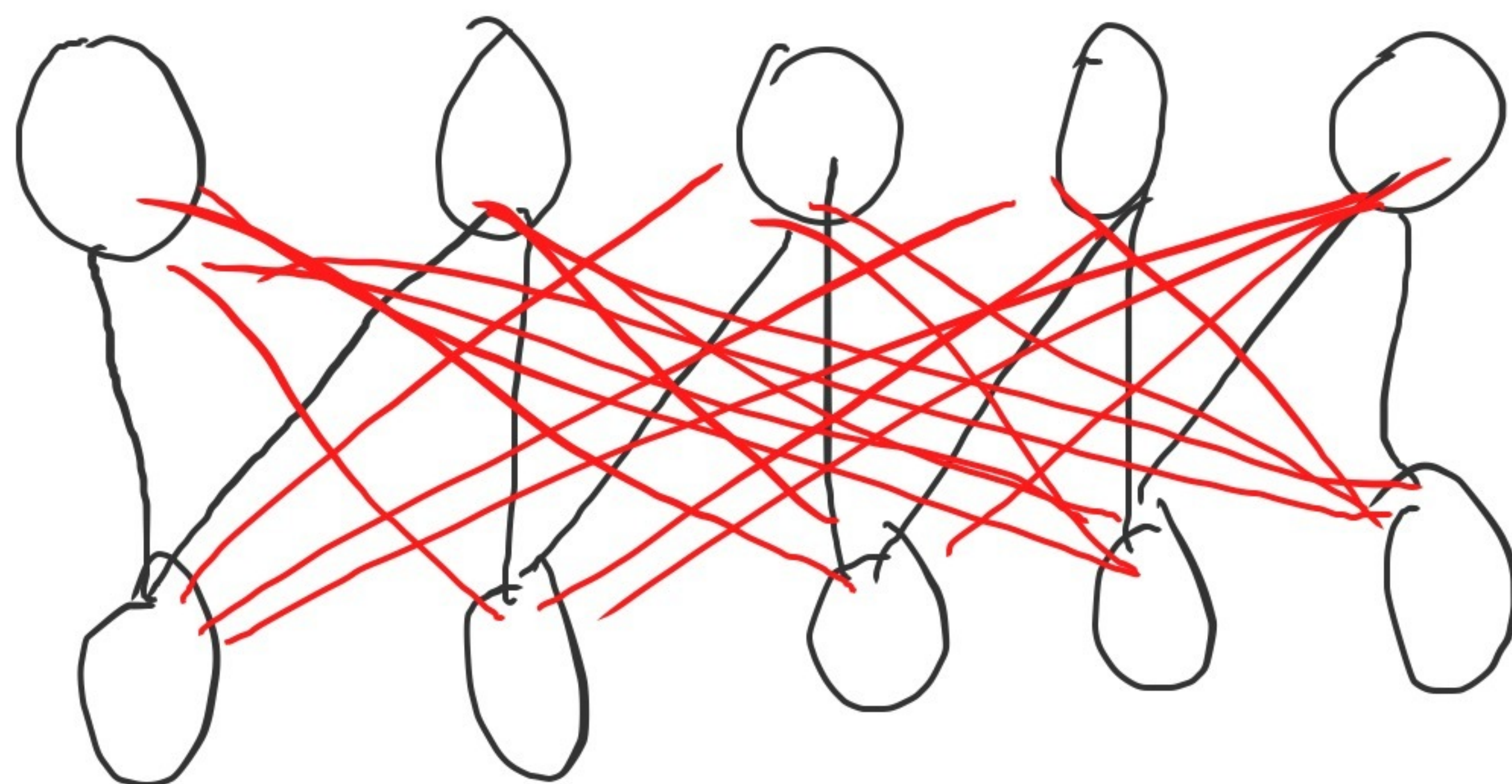
$\bar{N} \rightarrow NP$

~~what ball~~ did you buy

ball in the court

I bought a red ball

$NP \rightarrow DN$
locality



sparsity

locality

— zero

tying

$W=0$

prior

learning bias

the ball

$NP \rightarrow D N$

the ball is red

John caught

the ball

John threw

the ball

into the
basket

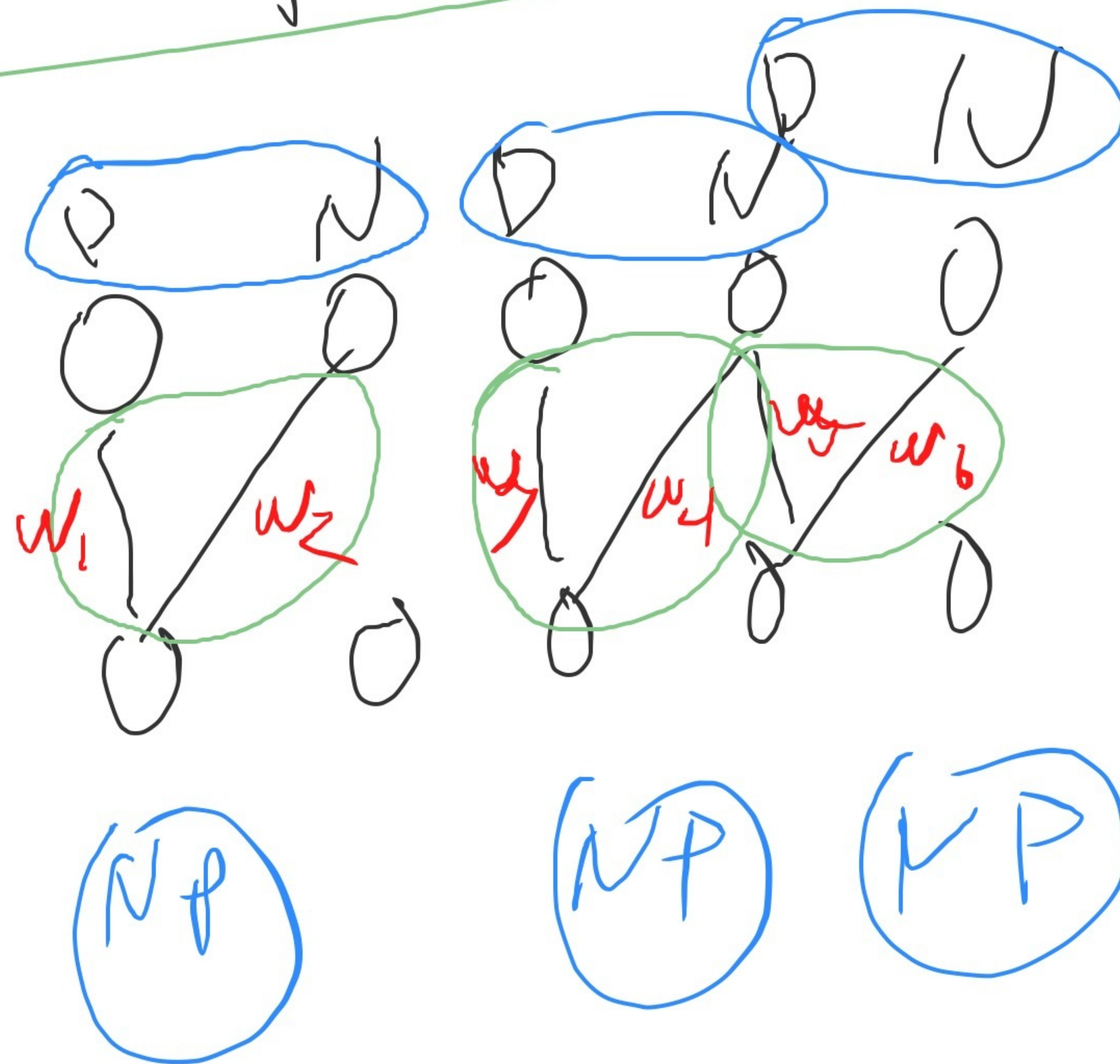
$S \rightarrow NP VP$

$VP \rightarrow V NP PP$

$VP \rightarrow V NP$

positive independence criterion

locality

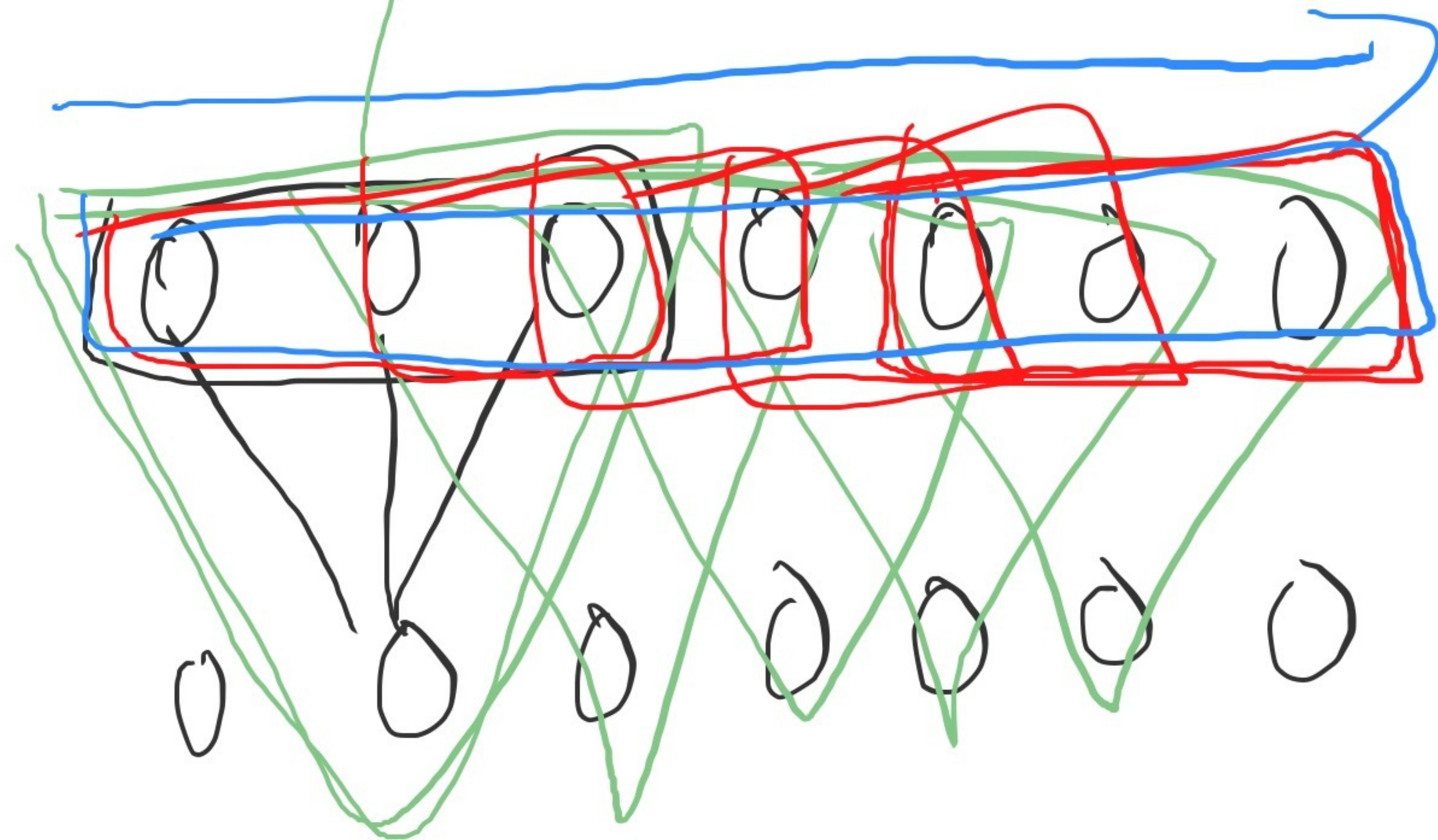


$$w_1 = w_3 = w_5$$
$$w_2 = w_4 = w_6$$

tying
sparsity

dot product

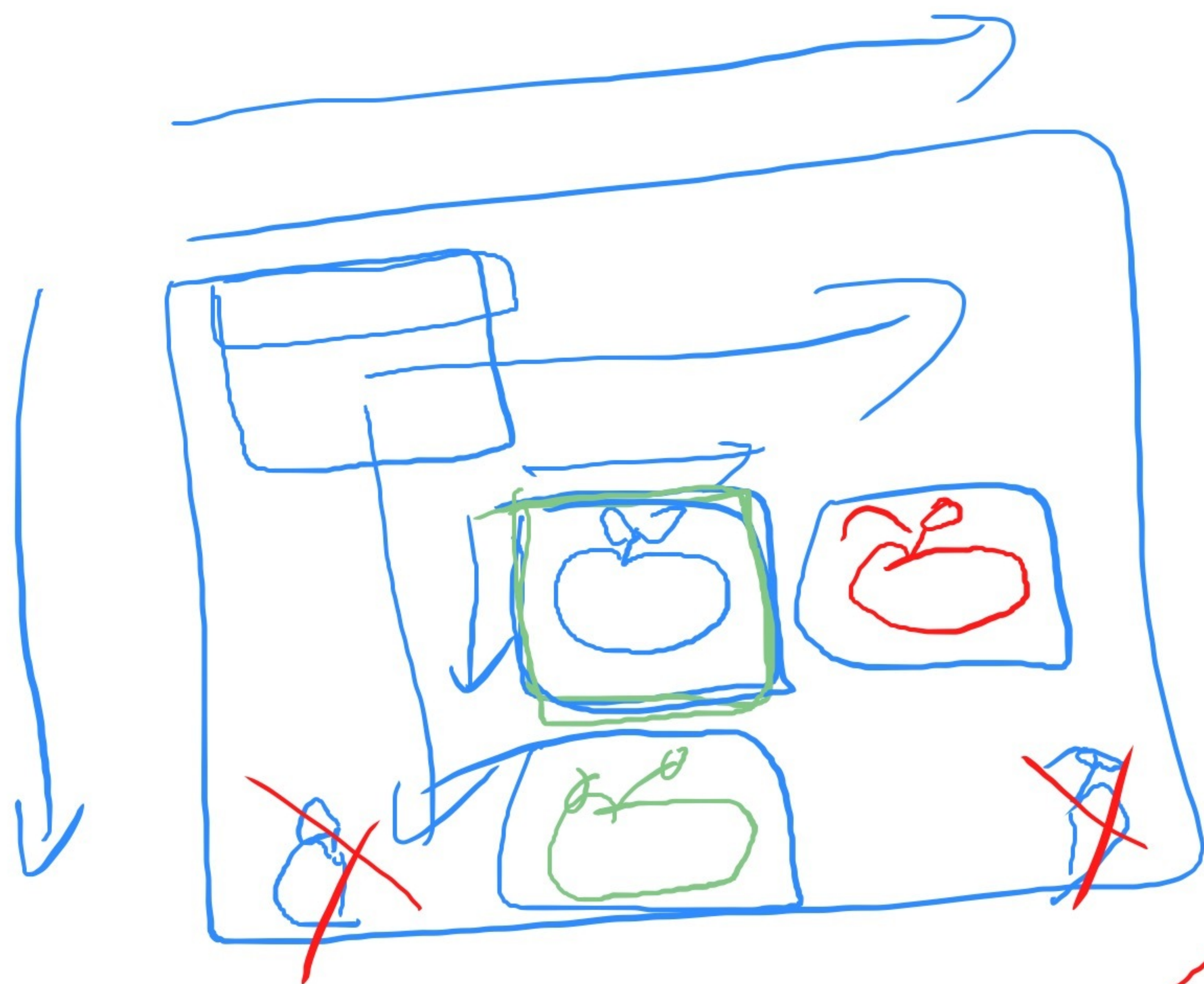
$$\vec{w} \cdot \vec{x}$$



↓ dimension

convolution

1D convolutional neural network



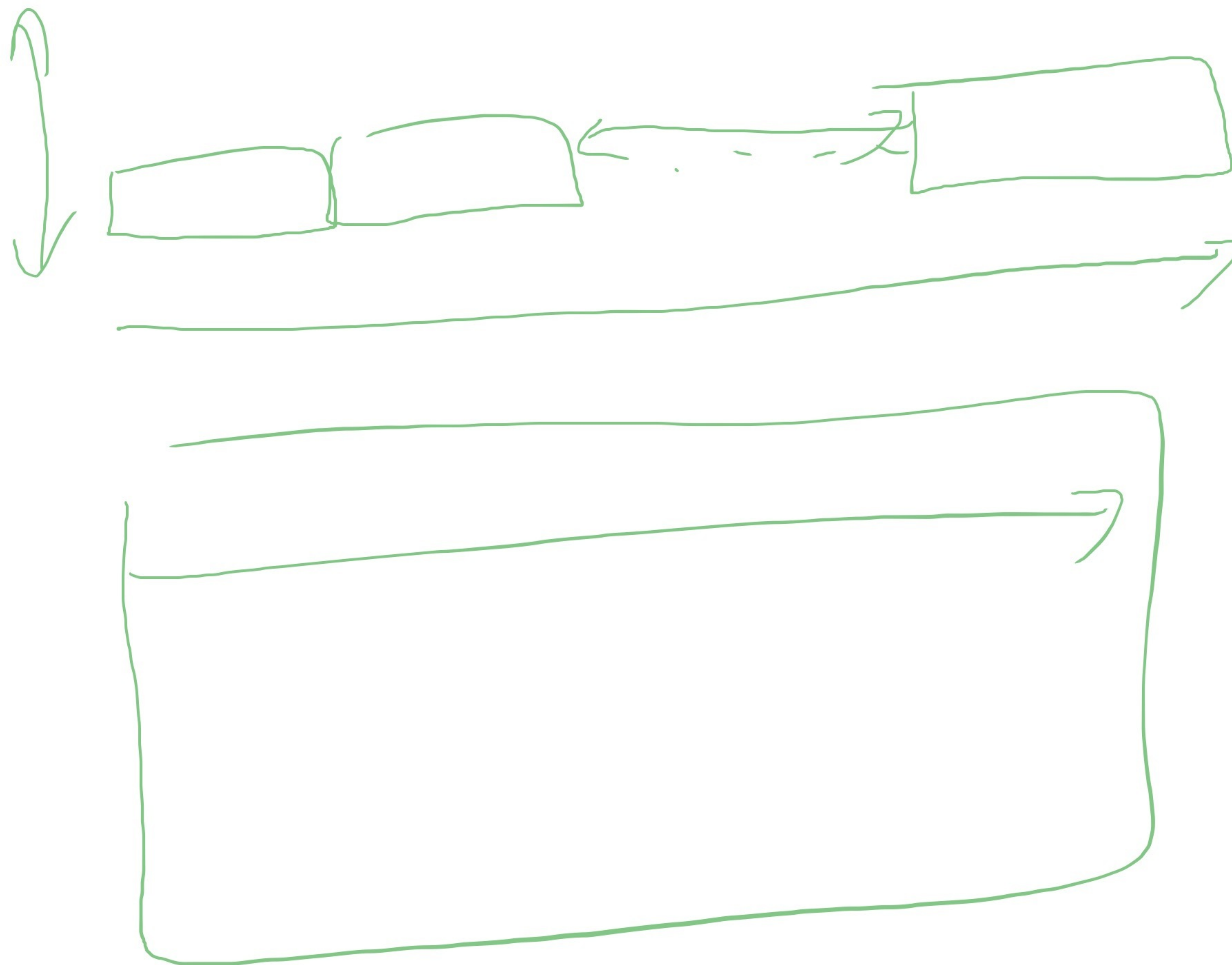
locality

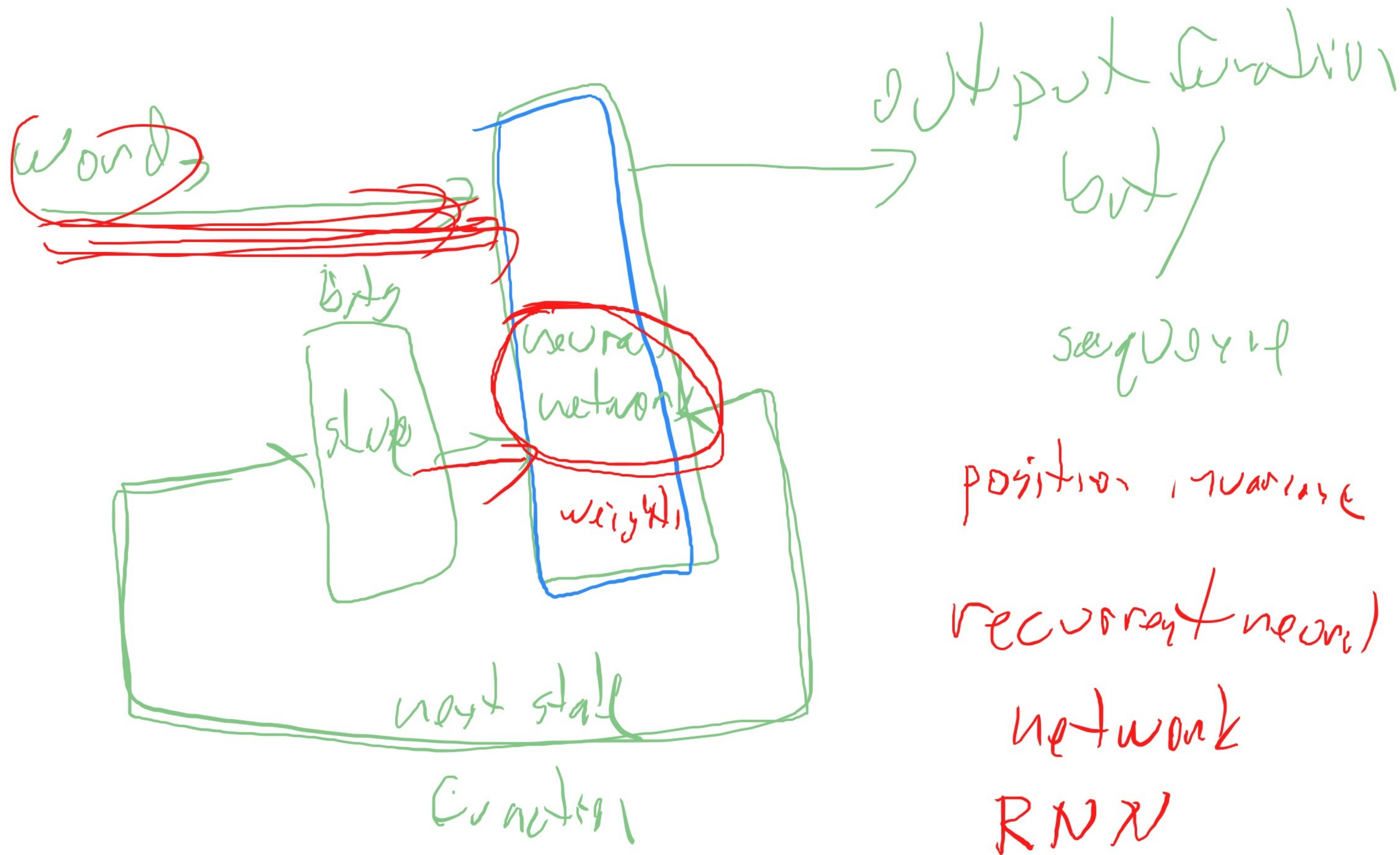
translation invariance

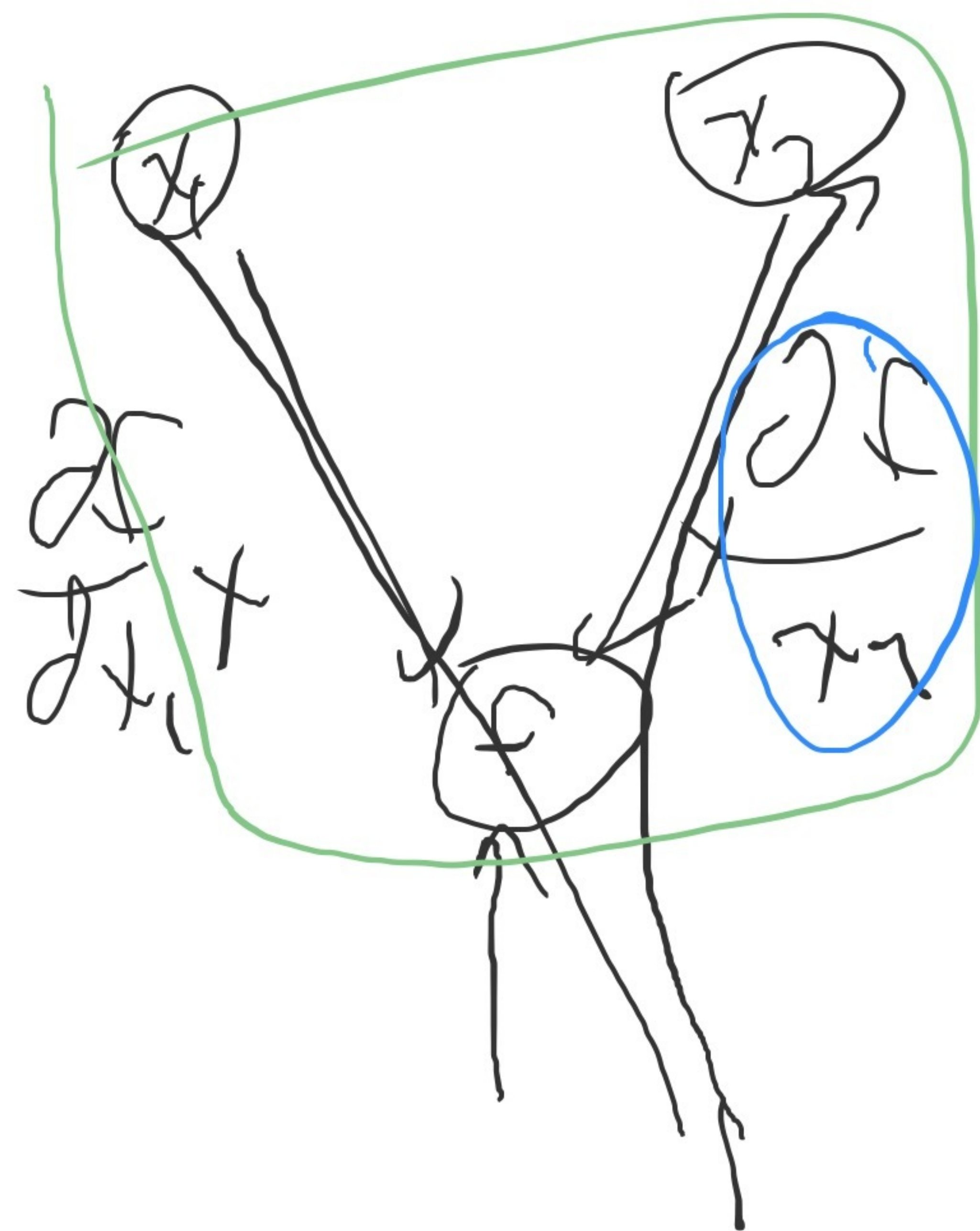
2D CNN

3D CNN









01

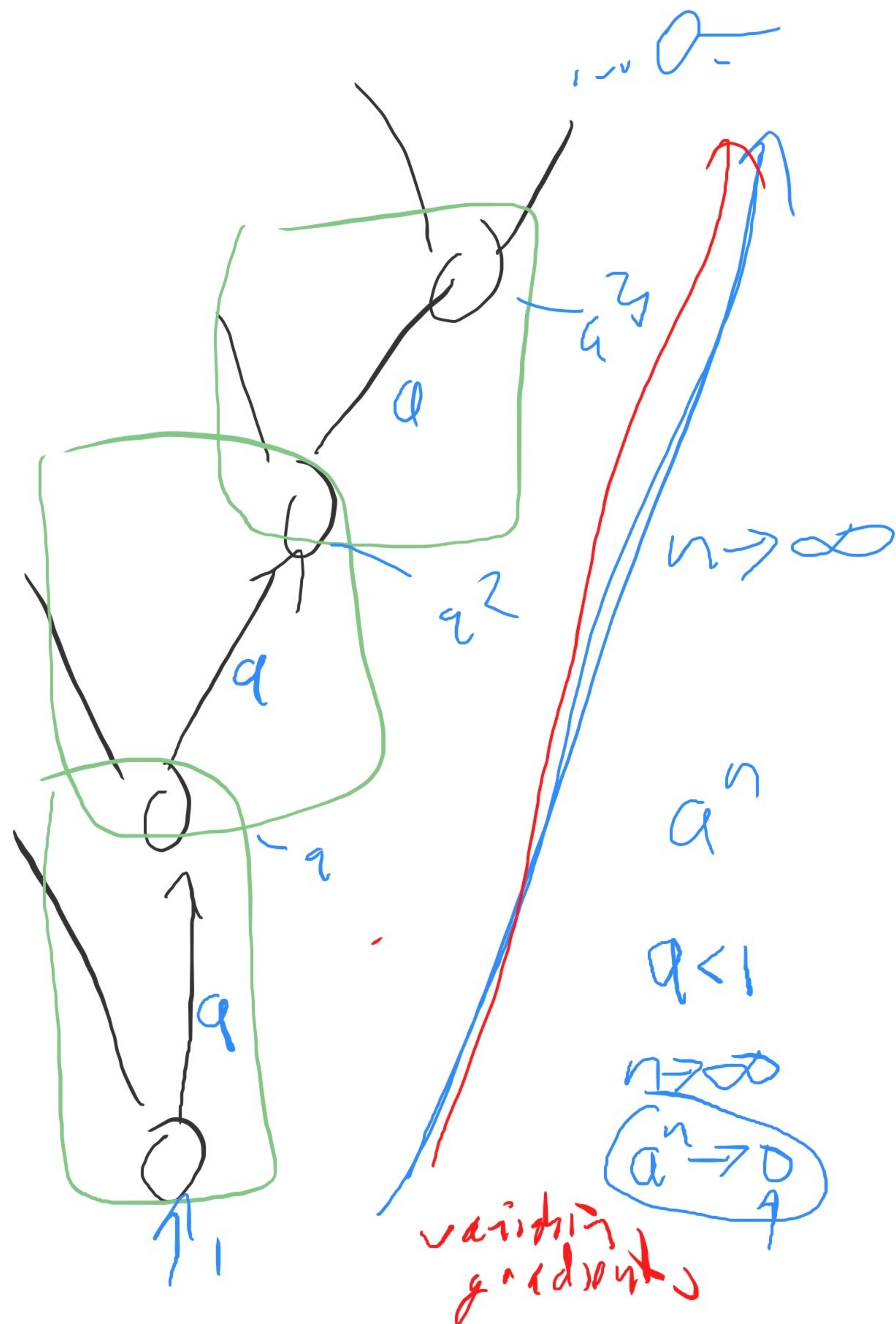
v

0

$$d > 1$$

$$n \rightarrow \infty$$

$$a^n \rightarrow \infty$$



$$n \rightarrow \infty$$

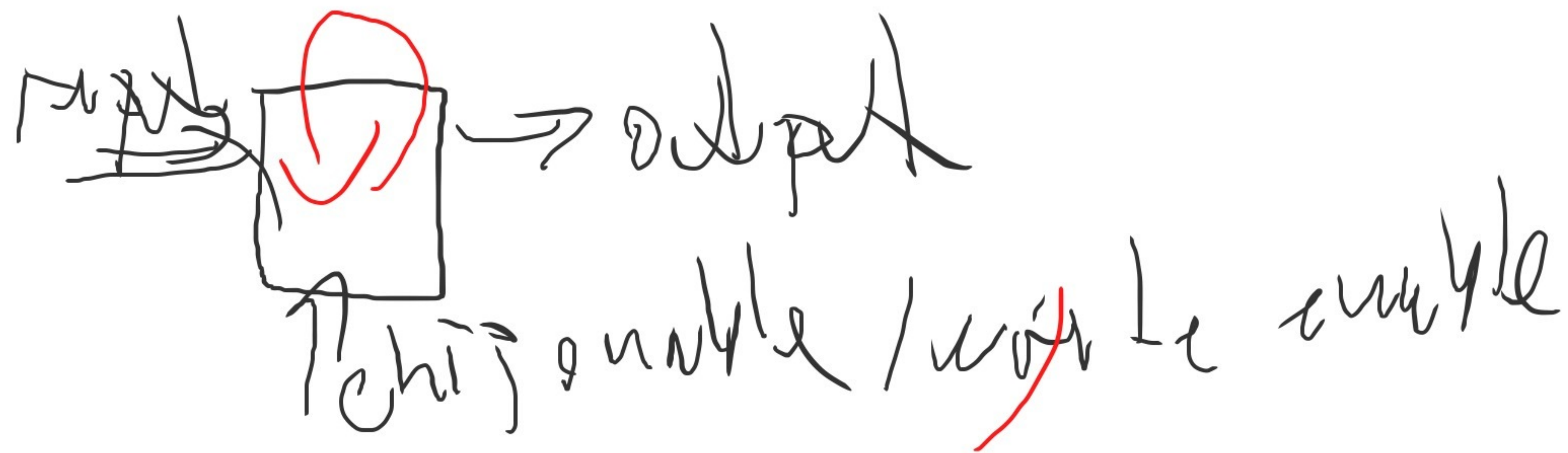
$$a^n$$

$$a < 1$$

$$n \rightarrow \infty$$

$$a^n \rightarrow 0$$

variation gradient



LSTM
long short-term memory
gate
forget gate