SUMMARY – ALGORITHM

Algorithm

Step 1: new input

receive $\underline{u} \in \{0,1\}^N$ — binary input vector

build: $\underline{U} = \underline{1}^M \cdot \underline{u}^T$ — binary input matrix $\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & \\ 1 & 0 & 0 & 1 & 1 & \cdots \\ 1 & 0 & 0 & 1 & 1 & \end{bmatrix}$

Step 2: update output matrix

$\underline{P} = \underline{X}'$ — represents previous state matrix

$\underline{B} = 1 - \underline{1}^m \max_{col}(\underline{P})$ — burst matrix (with bursting col's)

$\underline{C} = \underline{P} + \underline{B}$ — candidates for active outputs

$\underline{Y} = \underline{C} \cdot \underline{U}$ — active outputs

or: $$\boxed{\underline{Y} = (\underline{X}' + (1 - \max_{col} \underline{X}')) \circ (\underline{1}^m \cdot \underline{U})}$$

Step 3: state transition = calculate for all (i,j)

$\underline{S}_{ij} = \sigma(\underline{D}_{ij} - \Theta)$ — binary synaptic matrix of cell_{ij}

$\underline{Z}_{ij} = [\underline{y}(\underline{k}_{ij1}), \underline{y}(\underline{k}_{ij2}), \ldots, \underline{y}(k_{ijd})]$ — projected output to synaptic connections

$\underline{Q}_{ij} = \underline{S}_{ij} \circ \underline{Z}_{ij}$ *) — distal activation matrix

$\underline{X} = [x_{ij}] = [\| \underline{1}^{dT} \underline{Q}_{ij} \|_\infty \geq \Theta]$ — new state after transition

Step 4: learning

$\underline{\Delta D}_{ij} = (\delta^+ \underline{Q}'_{ij} - \delta^-)\, y'_{ij}$ $\quad (\delta^+ > \delta^-)$

$\underline{D}_{ij} = \underline{D}'_{ij} + \underline{\Delta D}_{ij}$ — update distal weights

*) „$\circ$“ means element wise multiplication

HTM SEQUENCE MEMORY

HTM-Layer

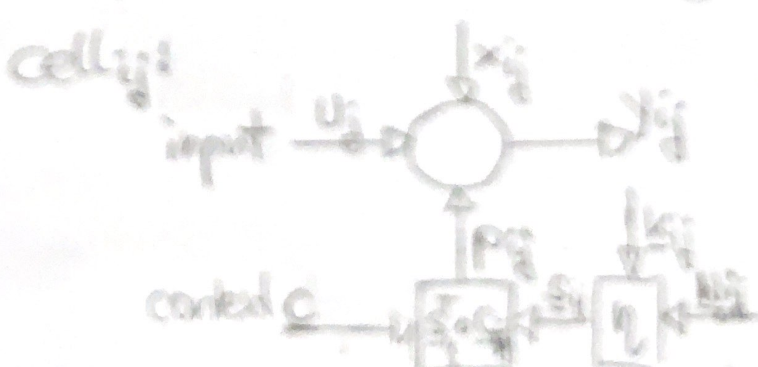
An HTM layer consists of n (cell) columns* each containing m cells**

Definitions

We use the notation $\mathbb{B} = \{0,1\}$ for the set of binary values, and $[0,1] := \{x \in \mathbb{R} \mid 0 \le x \le 1\}$. We denote [illegible] the set $\{x \in \mathbb{B}^{n} \mid$ [illegible] $\}$

Cells

Cells (also called HTM-neurons) are the elementary processing units with the following structure



$u_j \in [0,1]$
$x_{ij} \in [0,1]$
$y_{ij} \in [0,1]$
$\underline{c} \in [0,1]^{N}$
$p_{ij} \in [0:N]$
$\underline{k}_{ij} \in [1:N]^{M}$
$\underline{w}_{ij} \in [0,1]^{M}$

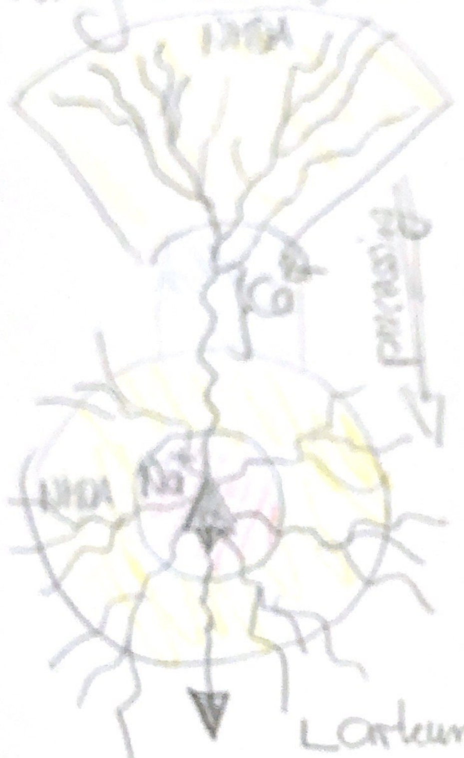
There are $N = n \cdot m$ cells in an HTM-layer.
s is the number of synapses per distal segment and d is the number of distal segments with $M = sd$

*) a cell column is frequently called minicolumn.

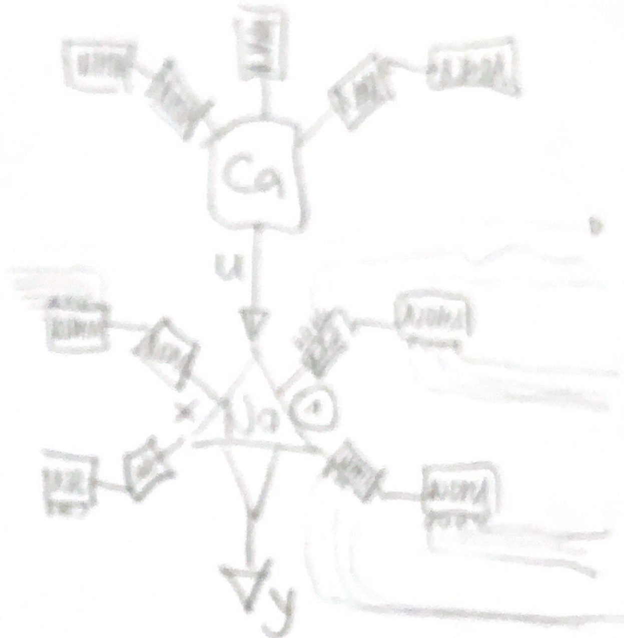
**) a cell is also called HTM-neuron

NEURAL MODEL

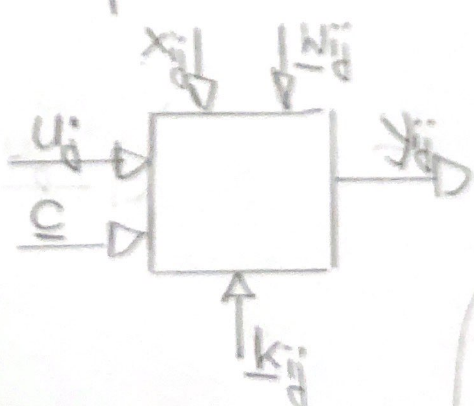
Our computational model refers to the following neural model

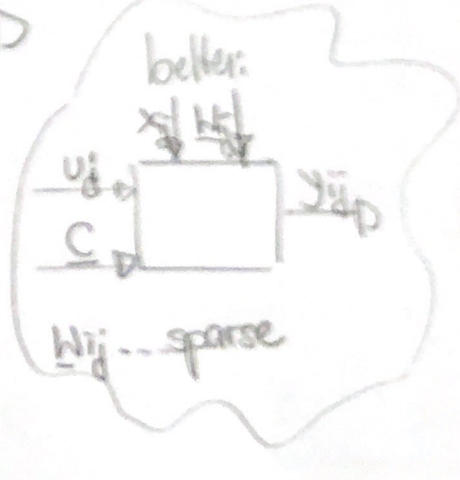


Larkum et. al. (2009)



From this neural model our HTM neuron is inspired as a computational block:





Units (minicolumns)

The $m \times n$ cells of an HTM layer are organized as a set of n units (or minicolumns). Each cell$_j$ comprises m cells, indexed by (i,j). It is important to realize that all cells of a unit$_j$ have the same common input $u_j \in [0,1]$, and the same context $\underline{c}$.

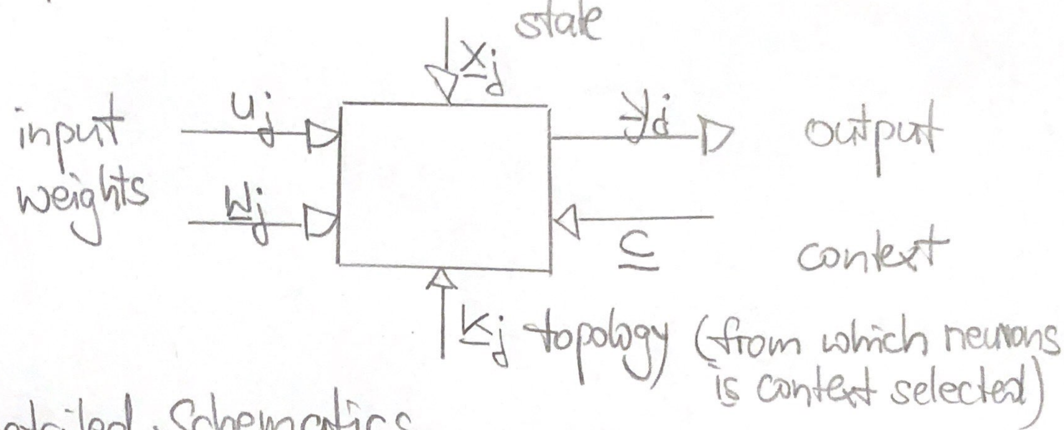
Vector-Signals/States

For a compact formalism it makes sense to combine scalar signal values to vectors. In this sense we define

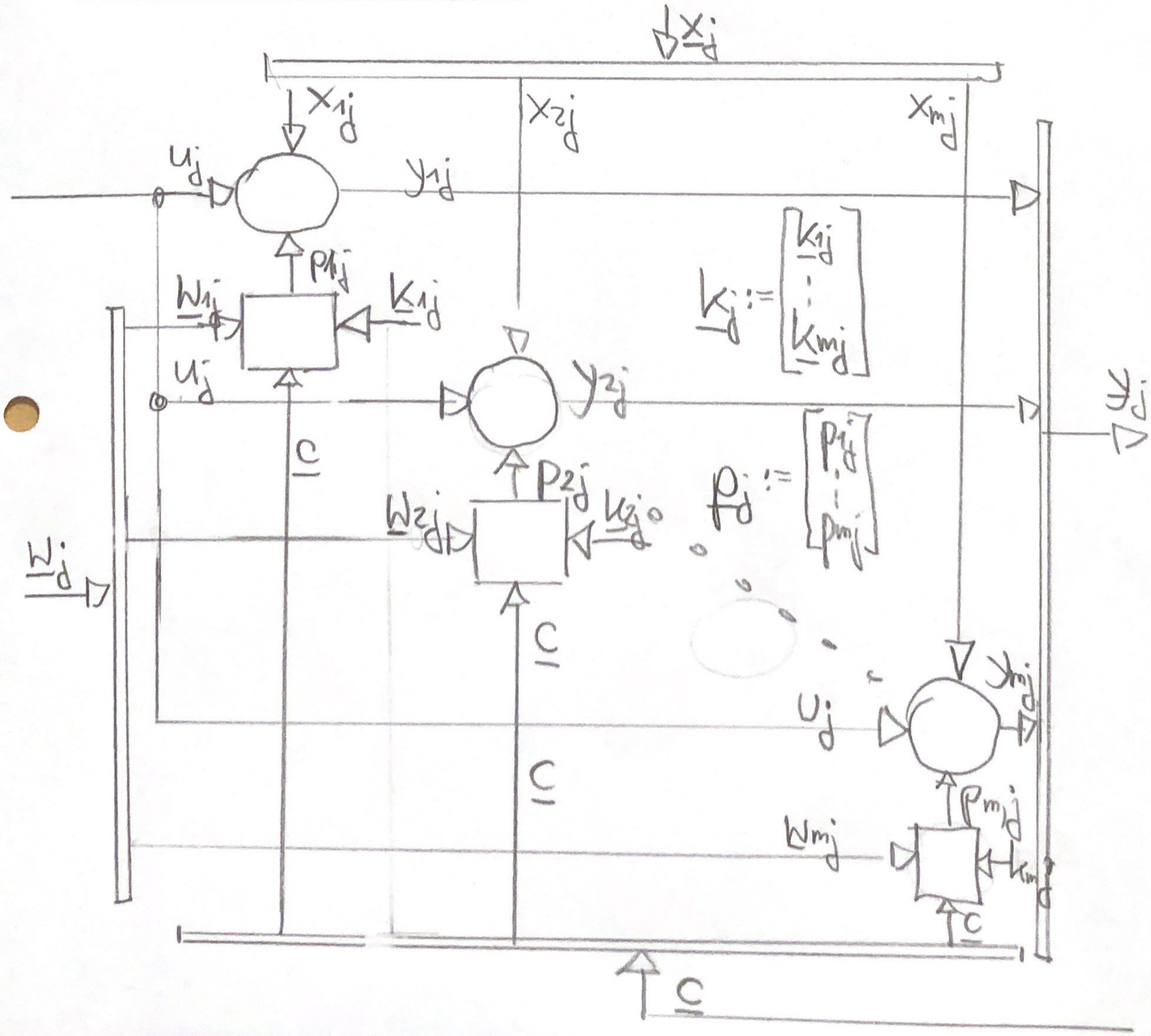
$$\underline{x}_j := [x_{1j}, \dots, x_{mj}]^T \in [0,1]^m \quad \text{(predictive) unit state}$$
$$\underline{y}_j := [y_{1j}, \dots, y_{mj}]^T \in [0,1]^m \quad \text{unit}_j \text{ output}$$
$$\underline{p}_j := [p_{1j}, \dots, p_{mj}]^T \in \mathbb{N}_0^m \quad \text{prediction vector}$$
$$\underline{S}_j := [\underline{s}_{1j}, \dots, \underline{s}_{mj}] \in [0,1]^{N \times m} \quad \text{synaptic matrix}$$
$$\underline{K}_j := [\underline{k}_{1j}, \dots, \underline{k}_{mj}] \in [1..N]^{M \times m} \quad \text{synaptic index matrix}_j$$
$$\underline{W}_j := [\underline{w}_{1j}, \dots, \underline{w}_{mj}] \in [0,1]^{M \times m} \quad \text{synaptic weight matrix}$$

SCHEMATICS OF A (CELL) UNIT

Simplified Schematics



Detailed Schematics



HTM-Layer / Recap

Since for (cell) unit j we have the following scalar/vector quantities

1 $u_j \in [0,1]$ input (common for all cells per unit)

4 $\underline{x}_j, \underline{y}_j \in [0,1]^m$ state and output

4 $\underline{p}_j \in [0:N]^m$ prediction strength

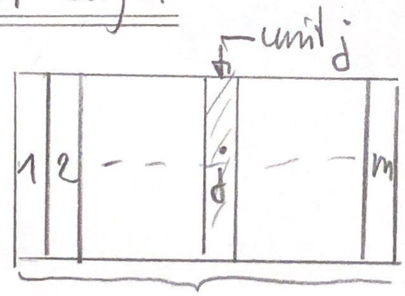
12 $\underline{c} \in [0,1]^{N^*}$ context (N^* might be greater than N)

40 $\underline{S}_j \in [0,1]^{m \times N^*}$ synaptic connection matrix

12 $\underline{k}_j \in [0,1]^M$ topology indizes

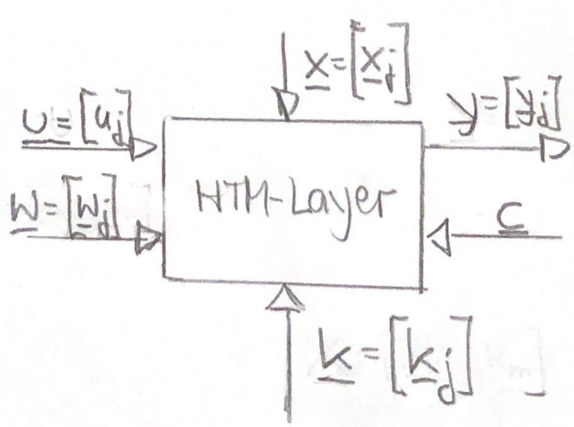
48 $\underline{W}_j \in [0,1]^{Mm}$

HTM-Layer



$$\underline{x} := [\underline{x}_j] = \begin{bmatrix} \underline{x}_1 \\ \vdots \\ \underline{x}_n \end{bmatrix}$$

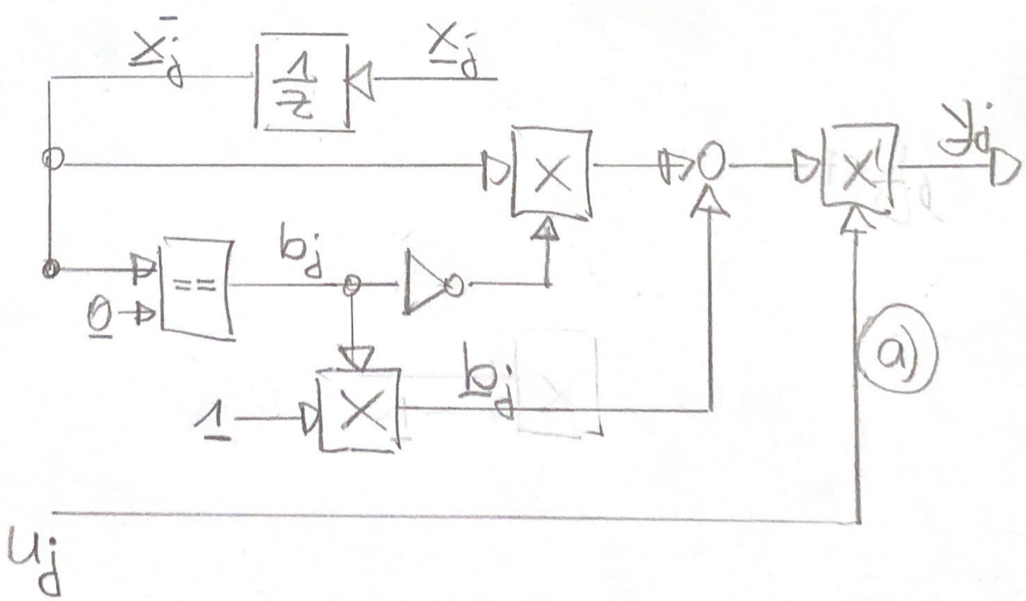
$$\underline{y} := [\underline{y}_j] = \begin{bmatrix} \underline{y}_1 \\ \vdots \\ \underline{y}_n \end{bmatrix}$$



Output Activation

a) Outputs are only active for active input signals
b) if a cell was predictive in the previous step, then the output is active (given $u_j = 1$)
c) if no cell in the unit was predictive in the previous step, then the output gets active (given $u_j = 1$)

Signal Processing Chart



Remarks

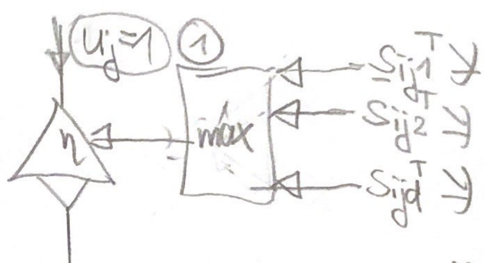
(i) given $\underline{x}_j^-$ (previous predictive state) equals $\underline{0}$, we have $b_j = 1$, and $\underline{b}_j = [1, 1, \ldots, 1]$ (burst). Then (given $u_j = 1$) the output becomes $\underline{y}_j = [1, 1, 1, 1]$
(ii) given $\underline{x}_j^- \neq \underline{0}$ we have $b_j = 0$ and $\underline{y}_j$ gets $\underline{x}_j^-$ (given $u_j = 1$)

STATE TRANSITION

In every cell (HTM neuron) the updated state is 1, iff the cell has at least one distal segment with ...

...active input ($u_j=1$) ①
and a sufficient number of neurons
② firing to this dendritic segment. ("sufficient ~~support~~")

"active sufficient contextual ~~support~~ activation"

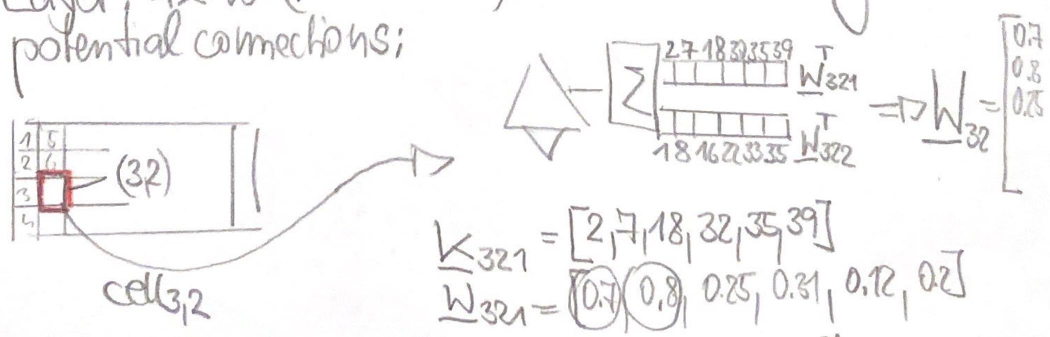


What means "sufficient support"

=> it means that the total sum of synaptic connections to activated cells exceed threshold η

Example

Layer: 4x10 (N=40 cells) with 2 distal segments à 6 potential connections:



$K_{321} = [2, 7, 18, 32, 35, 39]$

$W_{321} = [0.7, 0.8, 0.25, 0.31, 0.12, 0.2]$

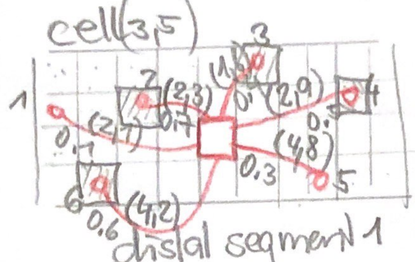
$K_{322} = [18, 16, 22, 33, 35]$

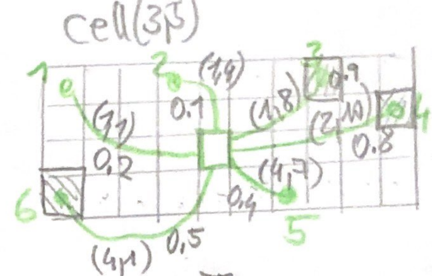
$W_{322} = [\quad]$

DISTAL ACTIVATION

the distal segments are characterized by a weight matrix $\underline{W}_{ij}$ and a connection index matrix $\underline{K}_{ij}$

Let $\varkappa(\cdot,\cdot)$ be the index mapping function from $(i,j) \mapsto k$, given by $\varkappa(i,j) \mapsto (m-1)j + i$ we could have for example ($s=6$, $d=2$)





$$\underline{K}_{351} = [\varkappa(2,1), \varkappa(2,3), \varkappa(1,6), \varkappa(2,9), \varkappa(4,8), \varkappa(4,2)]^T =$$
$$= [2, 10, 21, 34, 32, 8]^T$$

$$\underline{K}_{352} = [\varkappa(1,1), \varkappa(1,4), \varkappa(1,8), \varkappa(2,10), \varkappa(4,7), \varkappa(4,1)]^T =$$
$$= [1, 13, 25, 38, 28, 4]^T$$

$$\underline{W}_{351} = [0{,}1, 0{,}7, 0{,}9, 0{,}5, 0{,}3, 0{,}6]^T$$
$$\underline{W}_{352} = [0{,}2, 0{,}1, 0{,}9, 0{,}8, 0{,}4, 0{,}5]^T$$

This leads us to the synaptic connection matrices

$$\underline{S}_{351} = \begin{bmatrix} 0&0&0&0&0&1&0&0&0&0 \\ 0&0&1&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0&0 \end{bmatrix} = \sigma(\underline{K}_{351} \otimes \underline{W}_{351} - \eta)$$

$$\underline{S}_{352} = \begin{bmatrix} 0&0&0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0&0&0 \end{bmatrix} = \sigma(\underline{K}_{352} \otimes \underline{W}_{351} - \eta)$$

with $\eta = 0{,}5$ (in our example)

DISTAL ACTIVATION VECTOR

Given a cell (i,j) with d distal segments given by $\underline{K}_{ij} = [\underline{k}_{ij1}, \ldots, \underline{k}_{ijd}]$ and $\underline{W}_{ij} = [\underline{W}_{ij1}, \ldots, \underline{W}_{ijd}]$ the distal activation is defined as

$$\boxed{\underline{d}_{ij} = \underline{S}_{ij}^T \cdot \underline{y}} \qquad \underline{S}_{ij} := [\underline{S}_{ij1}, \ldots, \underline{S}_{ijd}]$$

$$\underline{S}_{ijk} := \underline{S}_{ijk}(:)$$

Example

Given $\underline{Y} = \underline{y}(m \times n) = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$

We get:

$$d_{351} = \underline{S}_{ij1}^T \underline{y} = 3, \quad d_{352} = \underline{S}_{ij2}^T \underline{y} = 1$$

$$\Rightarrow \underline{d}_{35} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

which means

For distal segment 1 there are 3 feeding cells to synaptic connections, and for distal segment 2 there is only 1 feeding cell, given the activated output vector $\underline{y}$

DISTAL ACTIVATION MATRIX AND STATE TRANSITION

Distal Activation Matrix

$$\underline{D} := [\underline{d}_1, \underline{d}_2, \dots, \underline{d}_N] \quad \text{with } \underline{d}_k := \underline{d}_j \text{ given } k = \alpha(j)$$

The distal activation $\underline{p} \in [0,1]^N$ is defined as

$$\boxed{\underline{p} = \max_{row} \underline{D}^T}$$

$\max_{row}(\ )$ means find the maximum element of each row has to be taken

Example

$$\underline{D} = \begin{bmatrix} \cdots & 3 & \cdots \\ \cdots & \underbrace{1}_{13 = \alpha(33)} & \cdots \end{bmatrix} \Rightarrow \max_{row} \underline{D}^T = \begin{bmatrix} \vdots \\ 3 \\ \vdots \end{bmatrix} \Rightarrow \left(\max \underline{D}^T \geq 2\right) = \begin{bmatrix} \vdots \\ 1 \\ \vdots \end{bmatrix}$$

State transition

The new state $\underline{x}$ is calculated

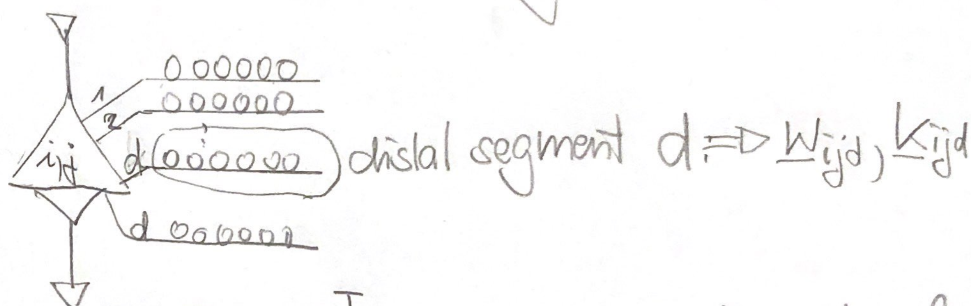
$$\boxed{\underline{x} = (\underline{p} \geq \theta)} \qquad \text{precisely: } \underline{y}^t = (\underline{p}^t \geq \theta)$$

Example: $\theta = 2$

$$\underline{D} = \begin{bmatrix} \cdots & 3 & \cdots \\ & 1 & \end{bmatrix}, \quad \underline{p} = \max_{row} \begin{bmatrix} \vdots & \vdots \\ 3 & 1 \\ \vdots & \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ 3 \\ \vdots \end{bmatrix} \Rightarrow \underline{y} = \begin{bmatrix} \vdots \\ 3 \geq 2 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

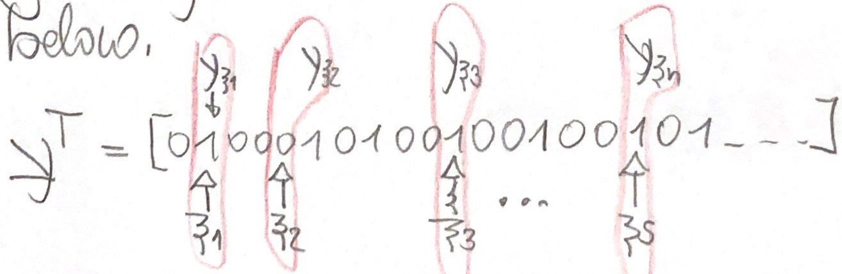
PROJECTED OUTPUT VECTOR

Let us focus on cell (i,j) with n_d distal segments indexed from $1..n_d$, each having s synapses



Let $\underline{K}_{ijd} = \underline{\xi} = [\xi_1, \ldots, \xi_s]^T$ $\underline{\xi}$ .. index vector under focus

Then $\underline{\xi}$ sets a focus on the cell output vector $\underline{y}$ given by the s indices $\xi_1, \ldots, \xi_s$. This is visualized below.



The projected vector $\underline{p}_{\xi} = [y_{\xi 1}, y_{\xi 2}, y_{\xi 3}, \ldots, y_{\xi s}]^T$ is the compact summary of the focused elements of $\underline{y}$ given by $\underline{\xi}$, and can be calculated as a projection

$$\underline{p}_{\xi} = \underline{P}(\underline{\xi}) \cdot \underline{y} \qquad \underline{P} \in \{0,1\}^{s \times N}$$

where

$$\underline{P}(\underline{\xi}) := \begin{bmatrix} \underline{e}^T_{\xi 1} \\ \underline{e}^T_{\xi 2} \\ \vdots \\ \underline{e}^T_{\xi s} \end{bmatrix}$$

and $\underline{e}_i^T$ being the i-th row of the $N \times N$ unit matrix $\underline{I}^{N \times N}$

DISTAL ACTIVATION $\underline{Q}_{ij}$

In each distal segment μ of cell (ij) we have

$$\underline{q}_{ij\mu} := \underbrace{\vartheta(\underline{W}_{ij\mu})}_{=\underline{s}_{ij\mu}} \circ \underbrace{\left[\underline{P}(\underline{k}_{ij\mu})\underline{y}\right]}_{=\underline{z}_{ij\mu} = \underline{y}(\underline{k}_{ij\mu})}, \quad \vartheta(x) := \begin{cases} 1 \ldots x \geq \Theta \\ 0 \ldots x < \Theta \end{cases}$$

as the activation vector for the distal segment, and

$$\underline{Q}_{ij} := [\underline{q}_{ij1}, \ldots, \underline{q}_{ijd}]$$

as the "distal activation matrix".

$\underline{Q}_{ij}$ gives an overview over the activation potentials arriving at the synapses

Example

consider: $\underline{W}_{ij} = \begin{bmatrix} 1 & 2 \\ 5 & 8 \\ 8 & 7 \\ 3 & 4 \\ 7 & 0 \\ 2 & 6 \end{bmatrix} \cdot 10^{-1}$, $\underline{K}_{ij} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 6 & 7 \\ 8 & 8 \\ 9 & 9 \end{bmatrix} = [\underline{k}_{ij1}, \underline{k}_{ij2}]$

Let: $\underline{y} = [0\,1\,1\,0\,0\,1\,0\,1\,1\,0\,0\,1\,0 \ldots]^T$

Then:

$$\underline{p}_{ij1} = [\underline{e}_1, \underline{e}_3, \underline{e}_5, \underline{e}_6, \underline{e}_8, \underline{e}_9]^T \underline{y} = [0\,1\,0\,1\,1\,1]^T$$

$$\underline{p}_{ij2} = [\underline{e}_2, \underline{e}_4, \underline{e}_6, \underline{e}_7, \underline{e}_8, \underline{e}_9]^T \underline{y} = [1\,0\,1\,0\,1\,1]^T$$

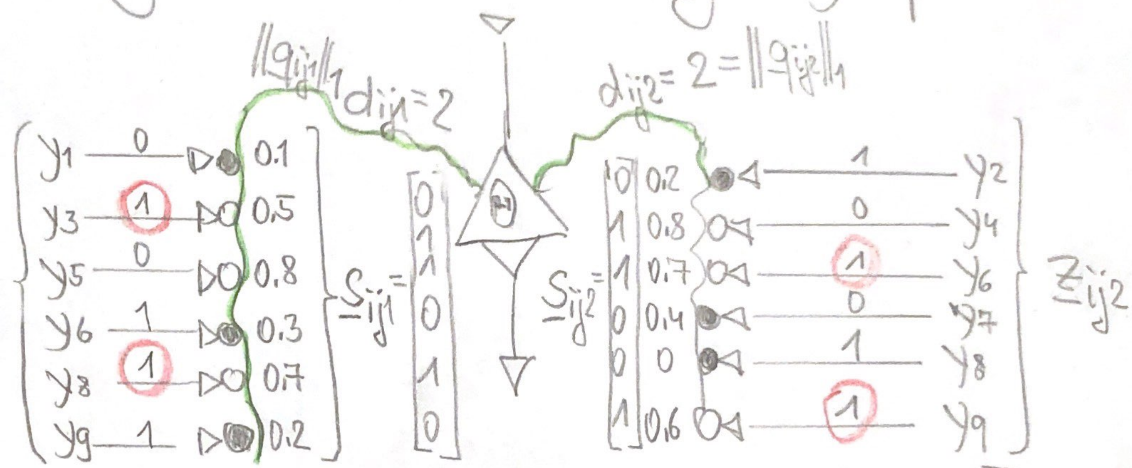
$$\underline{q}_{ij1} = \vartheta([1\,5\,8\,3\,7\,2]/10) \circ \underline{p}_{ij1} = [0\,1\,0\,0\,1\,0]^T$$

$$\underline{q}_{ij2} = \vartheta([2\,8\,7\,4\,0\,6]/10) \circ \underline{p}_{ij2} = [0\,0\,1\,0\,0\,1]^T$$

$$\Rightarrow \underline{Q}_{ij} = \begin{bmatrix} 0\,1\,0\,0\,1\,0 \\ 0\,0\,1\,0\,0\,1 \end{bmatrix}^T$$

What Is Distal Activation

Consider cell (i,j), which has d distal segments, each containing s synapses



$$y = [0\ 1\ 1\ 0\ 0\ 1\ 0\ 1\ 1\ 0\ 0\ 1\ 0 \dots]^T$$

Note

- Synapses with weight $< \eta$ (η -- synaptic threshold) are considered as ~~not~~ connected (black circles) and have as such no contribution

The contribution according to connected synapses is represented bei the distal activation vectors

distal segment 1:

$$q_{ij1} = S_{ij1} \circ Z_{ij1} = \begin{bmatrix}0\\1\\1\\0\\1\\0\end{bmatrix} \circ \begin{bmatrix}0\\1\\0\\1\\1\\1\end{bmatrix} = \begin{bmatrix}0\\1\\0\\0\\1\\0\end{bmatrix}$$

distal segment 2:

$$q_{ij2} = S_{ij2} \circ Z_{ij2} = \begin{bmatrix}0\\1\\1\\0\\0\\1\end{bmatrix} \circ \begin{bmatrix}1\\0\\1\\0\\1\\1\end{bmatrix} = \begin{bmatrix}0\\0\\1\\0\\0\\1\end{bmatrix}$$

thus $\underline{Q}_{ij} = [q_{ij1}\ q_{ij2}] = \begin{bmatrix}0 & 0\\1 & 0\\0 & 1\\0 & 0\\1 & 0\\0 & 1\end{bmatrix}$

distal activation {in segment 1, in segment 2

LEARNING

If a cell is correctly predicted, it means that it has at least one distal segment which has sufficient distal activation, i.e.

$$\boxed{\|q_{ijd}\|_1 \geq \Theta}$$

$$\|q_{ijd}\|_1 = \underline{s}_{ijd}^T \underline{z}_{ijd} \geq \Theta$$

Thus we calculate

$$\underline{d}_{ij} := \begin{bmatrix} \|q_{ij1}\|_1 \\ \|q_{ij2}\|_1 \\ \vdots \\ \|q_{ijd}\|_1 \end{bmatrix} = \underline{Q}_{ij}^T \underline{1}$$

Example

$$\underline{d}_{ij} = \underbrace{\begin{bmatrix} 0&1&0&0&1&0 \\ 0&0&1&0&0&1 \end{bmatrix}}_{\underline{Q}_{ij}^T} \begin{bmatrix} 1\\1\\1\\1\\1\\1 \end{bmatrix} = \begin{bmatrix} 2\\2 \end{bmatrix}$$

Learning Rule:

Any distal segment of an active cell ($y_{ij}=1$) is reinforced by δ^+

$$\boxed{\Delta \underline{W}_{ij} = \delta^+ \underline{Q}_{ij}^T - \delta^- \underline{1}^{s \times d}, \quad \delta^+ > \delta^-}$$

AGAIN: OUTPUT Calculation

Let $\underline{P} = \underline{X}^{-} = [\underline{p}_1, \dots, \underline{p}_m] \in \{0,1\}^{m \times n}$

$\Rightarrow$ define a burst matrix $\underline{B} = [\underline{b}_1, \dots, \underline{b}_m]$
with

$$\underline{b}_j = \begin{cases} \underline{0} & \text{for } \underline{p}_j \neq \underline{0} \\ \underline{1} & \text{for } \underline{p}_j = \underline{0} \end{cases}$$

$$\boxed{\underline{B} = 1 - \underline{1}^{m} \max_{col}(\underline{P})}$$

„bursts" at those columns which are zero in $\underline{P}$ (so it complements)

Then the output matrix $\underline{Y}$ calculates:

$$\boxed{\underline{Y} = (\underline{P} + \underline{B}) \circ (\underline{1}^{m} \cdot \underline{u}^{T})}$$

Example

$$\underline{P} = \begin{array}{c} \begin{array}{cccccccccc} & & & & & \text{column 6} & & & & \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

mind: the principle is that we use $\underline{P}$ as a basis with zero columns replaced by bursting columns

$$\underline{u} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}^{T}$$

$$\underline{B} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} = 1 - \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \underbrace{\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}}_{\max_{col}(\underline{P})}$$

$$\underline{P} + \underline{B} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

$$\underline{Y} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix} = (\underline{P} + \underline{B}) \circ \underbrace{(\underline{1}^{m} \cdot \underline{u}^{T})}_{\text{mask according to input } \underline{u}}$$