Recent NLP Highlights

AM11/AM12

Week 2

David Schlangen
david.schlangen@uni-potsdam.de

A Glossary of Machine Learning Terms and Concepts

Overview

- Learning & Machine Learning
- Everything's a Function
- Experience
- Types of Machine Learning
- Supervised Learning: X, Y, h, and θ
- Loss and Cost (as a function of θ), and Gradient Descent
- Representations
- The Perceptron and Its Learning Rule; What it Can't Do
- Weights & Activation Functions
- Backpropagation
- The Computation Graph
- Fully Connected Neural Networks
- Other Types of NN Architectures

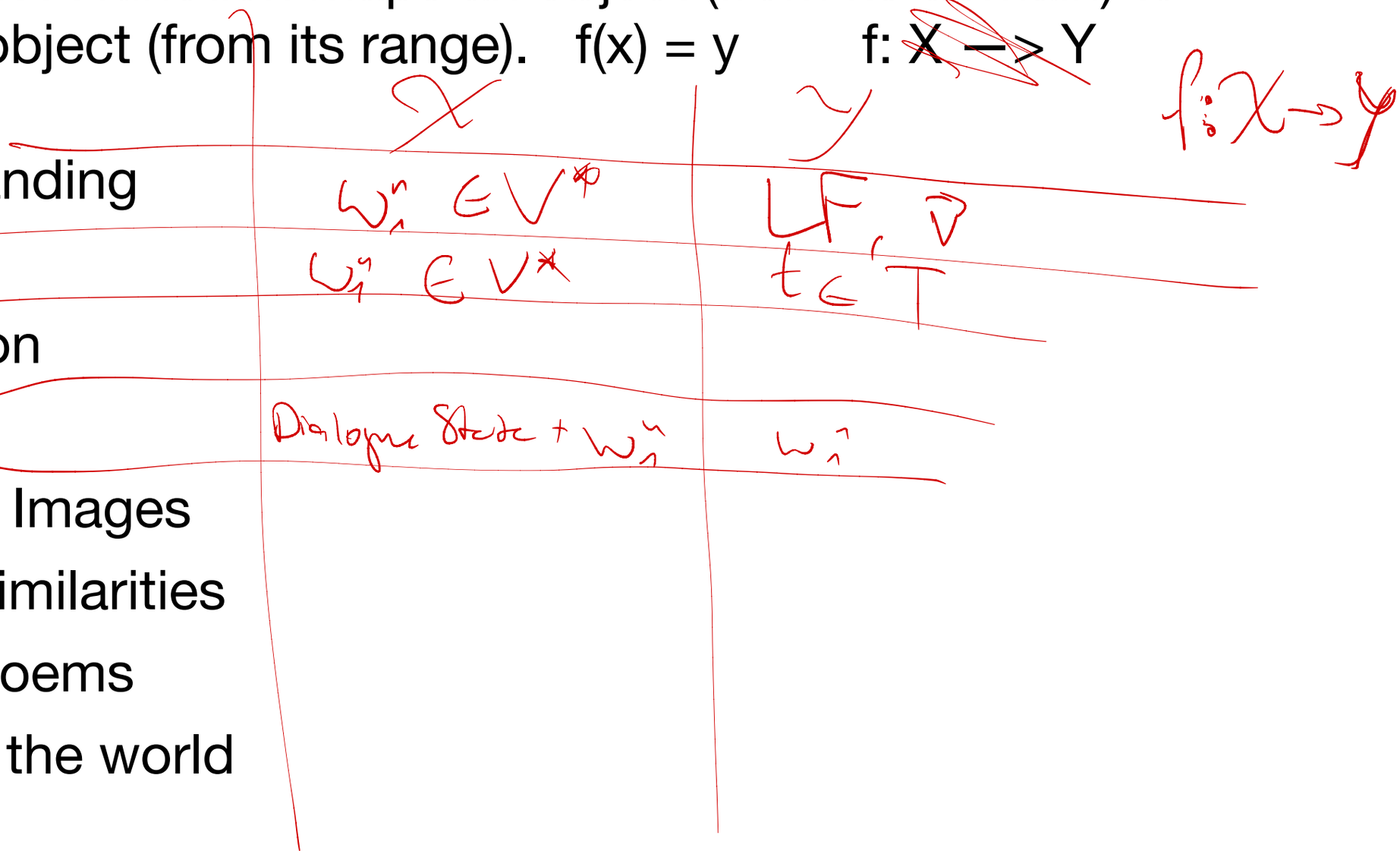
Learning

- Using experience to get better at something.
- Herb Simon: “Learning is any process by which a system improves performance from experience”
- Tom Mitchell: “A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.”
- Note: “Normal” programs don’t improve through experience. But ML models are still programs. ML is a form of *programming*. (Through spec of model + provision of data.)

Everything's a function

- Reminder: A function maps an object (from its domain) to another object (from its range). f(x) = y f: X —> Y

- Understanding
- Parsing
- Translation
- Dialogue
- Labelling Images
- Seeing Similarities
- Writing Poems
- Acting in the world
- …



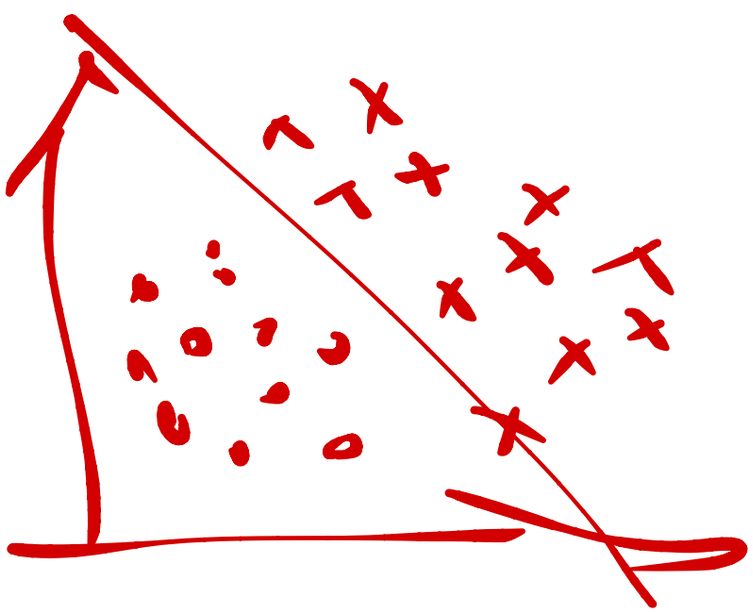
Experience

- = Data

$$X = \{ x^{(1)}, x^{(2)}, \dots, x^{(m)} \} \qquad x^{(i)} \in \mathcal{X}$$

Types of Machine Learning

- supervised learning: data + intended response
  - “supervision signal”
- unsupervised learning: just data, figure out structure



Supervised Learning

- supervised learning: data + intended response
  - "supervision signal" $X = \{x^{(1)}, \dots x^{(m)}\}$ $y = \{y^{(1)}, \dots y^{(m)}\}$
- Hypothesis, Loss, Cost, Objective Function

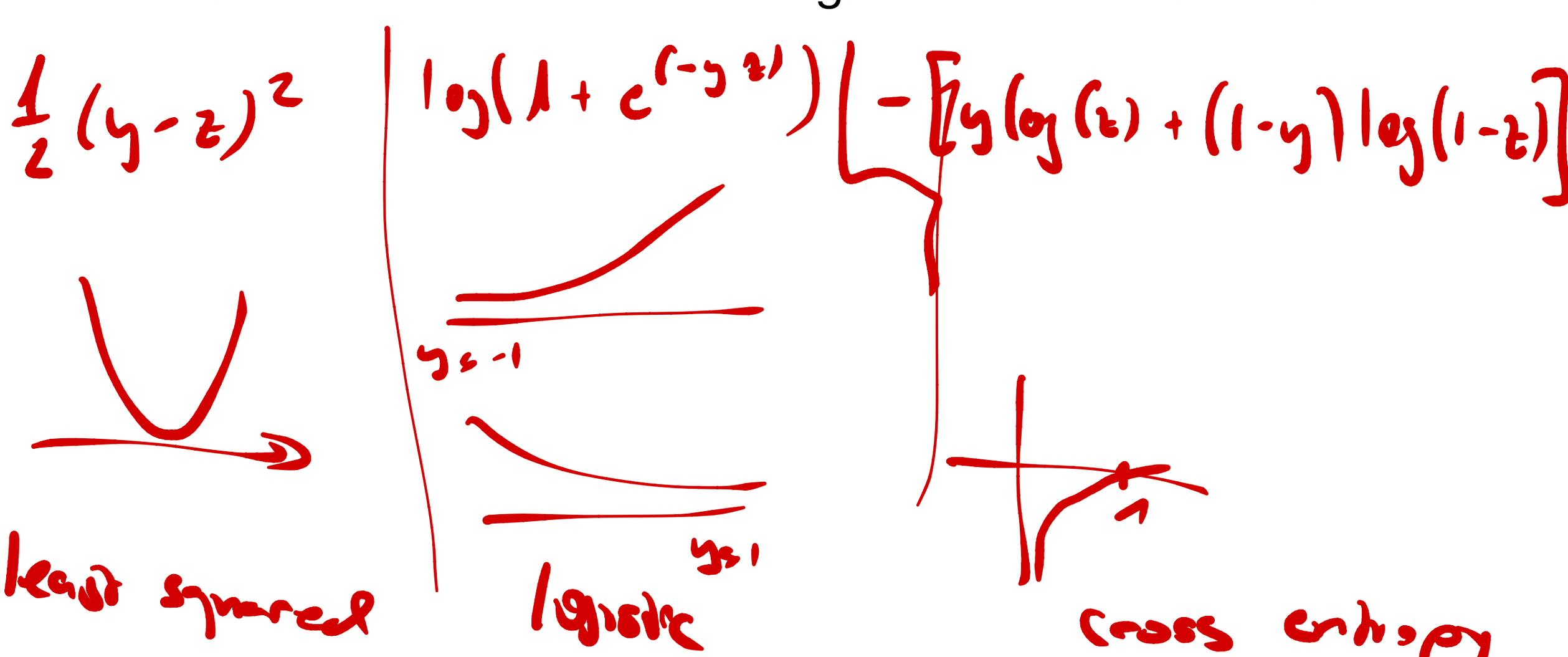
$$h_\theta(x^{(i)}) = z^{(i)}$$

$$L : (z, y) \in \mathbb{R} \times Y \longmapsto L(z, y) \in \mathbb{R}$$

$$J(\theta) = \sum_{i=1}^{m} L(h_\theta(x^{(i)}), y^{(i)})$$

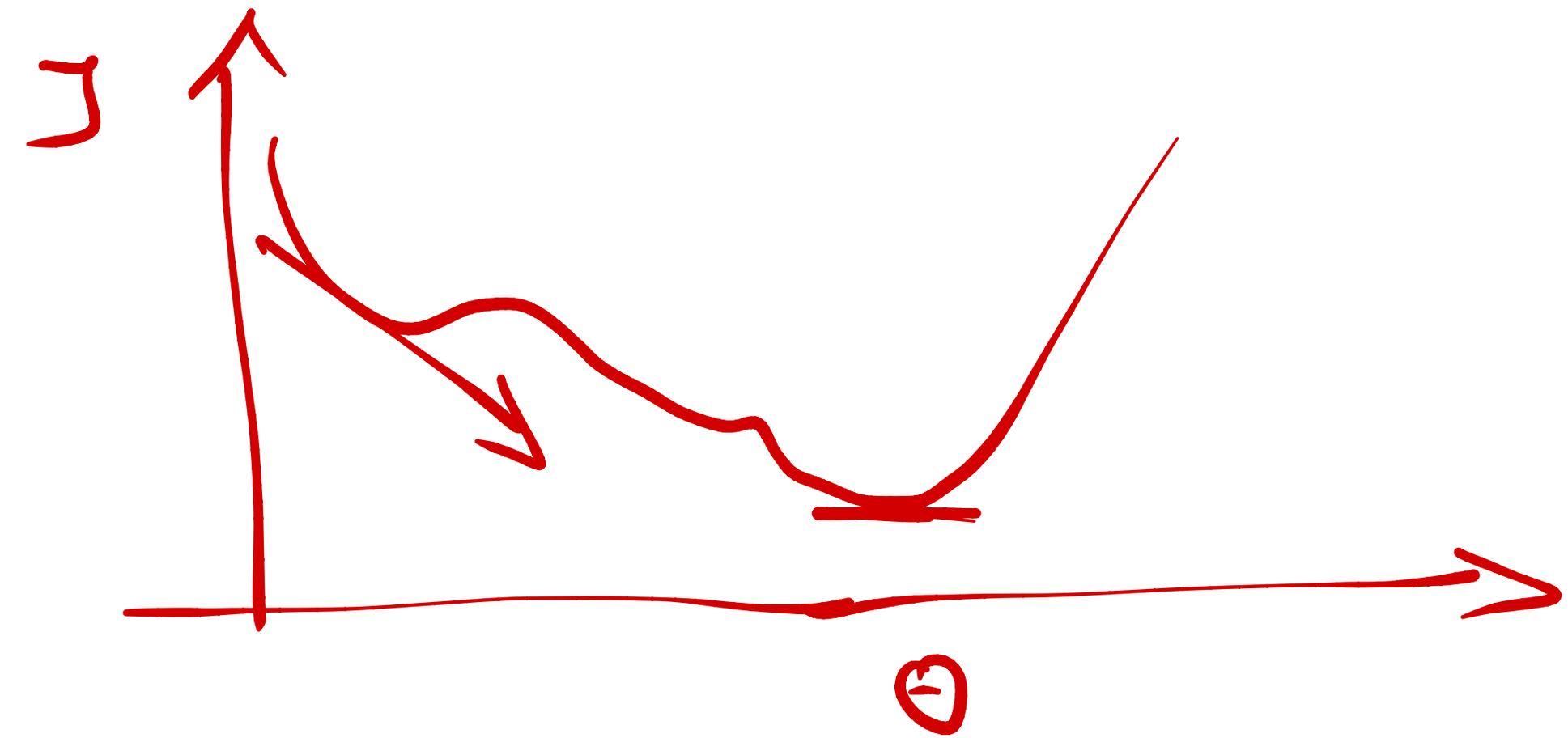
Loss & Cost

- main concept: Loss and cost as function of parameters; data is kept constant.
- loss function direct the learning. different kinds exist.



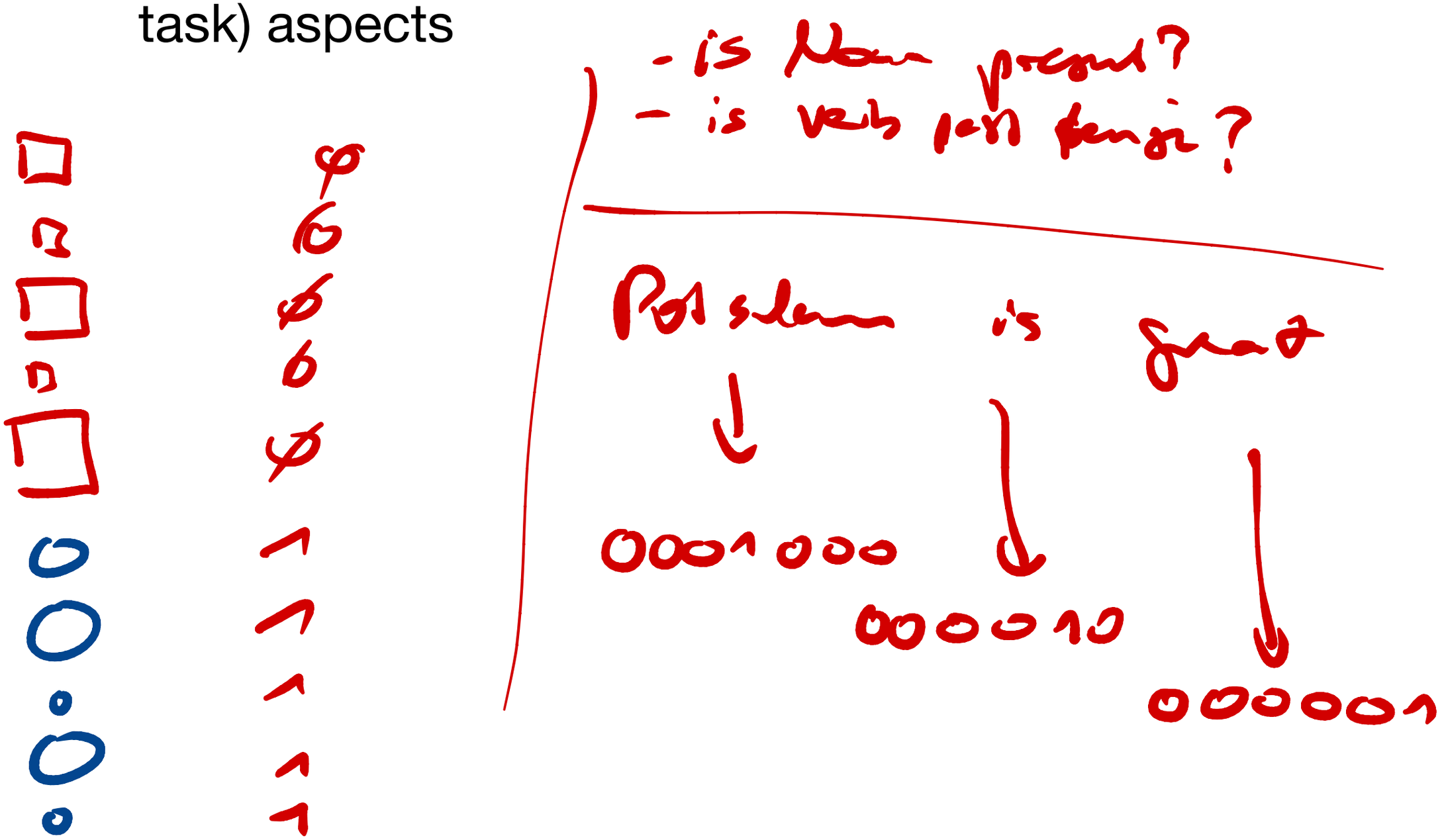
Minimise Cost

- differential calculus, finding minima

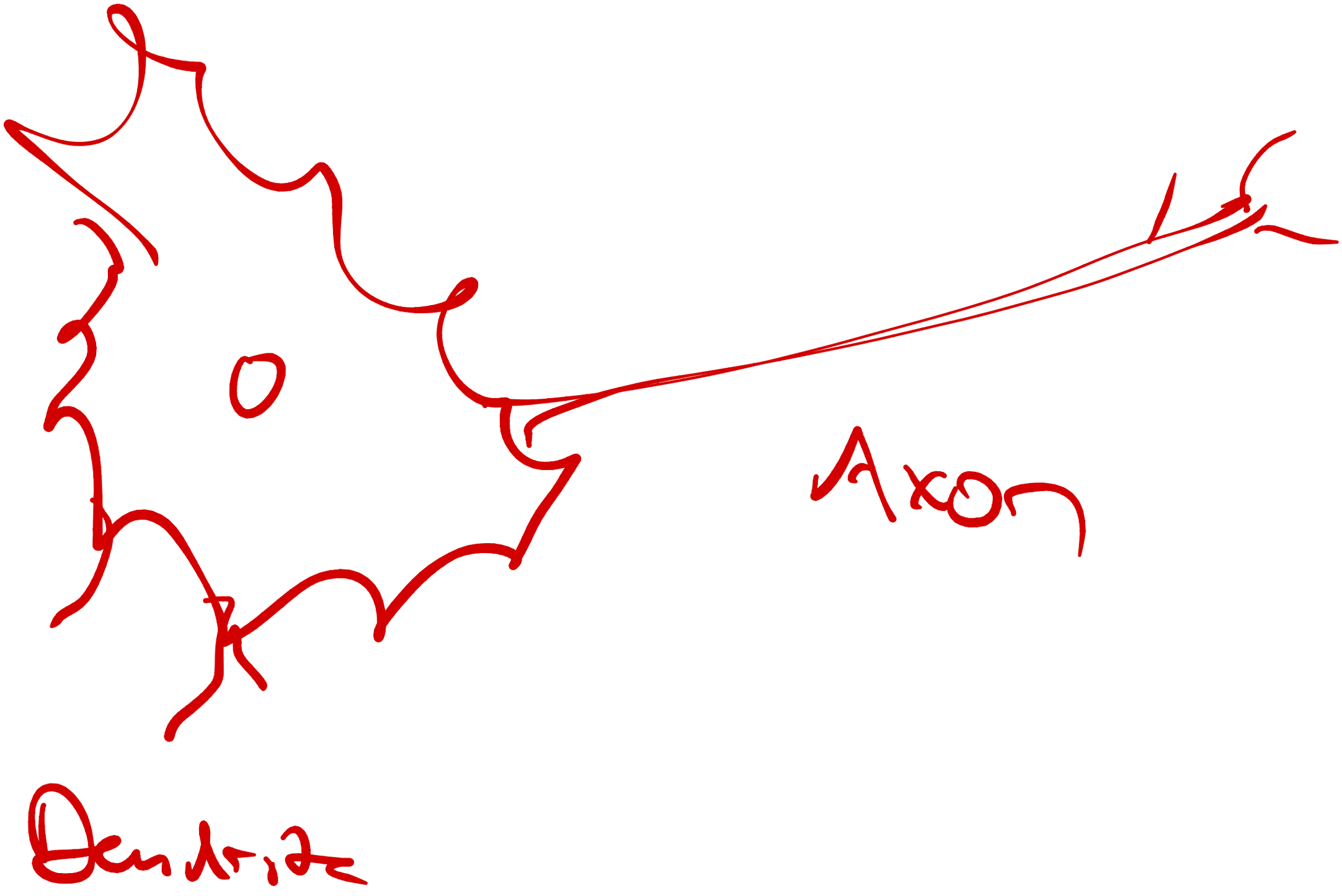


Representations

- Description of object, presenting the important (for the task) aspects

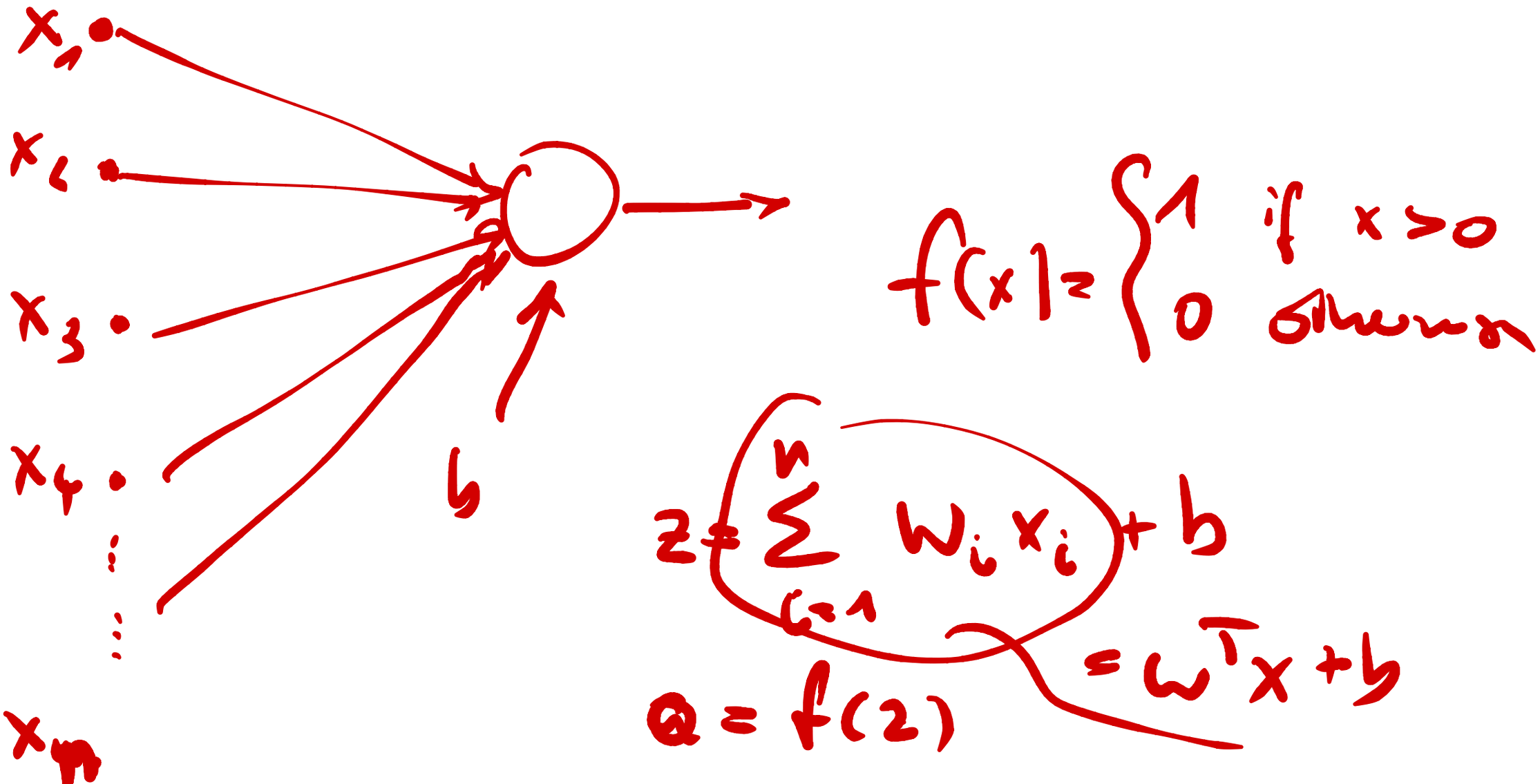


A Neuron



The Perceptron

- Rosenblatt 1957

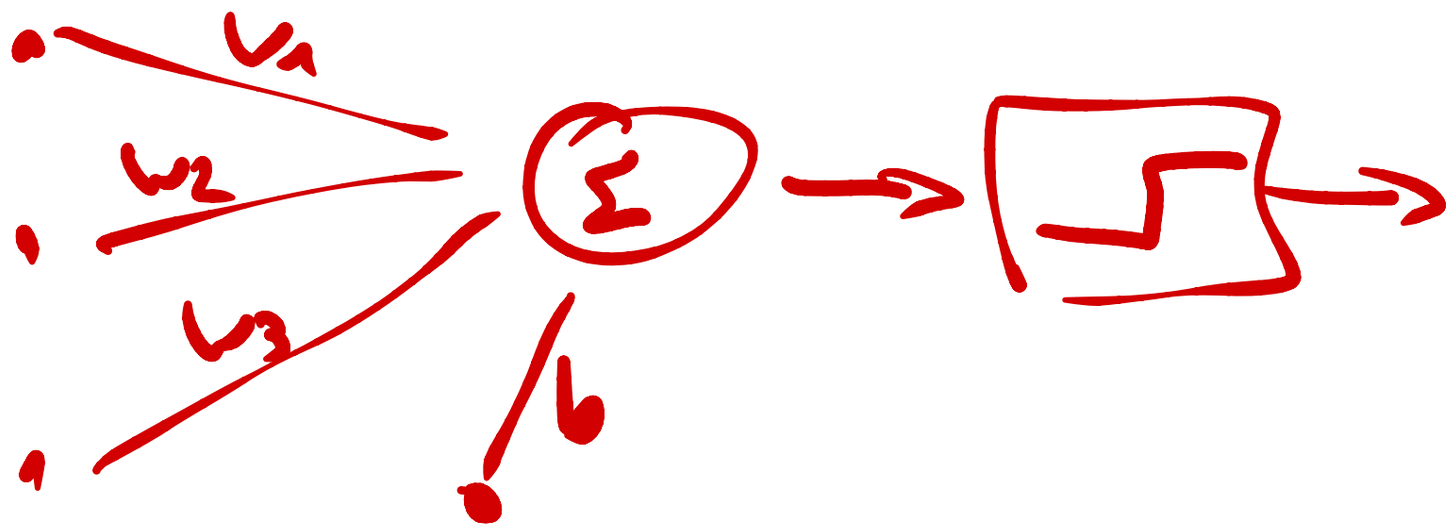


The Perceptron

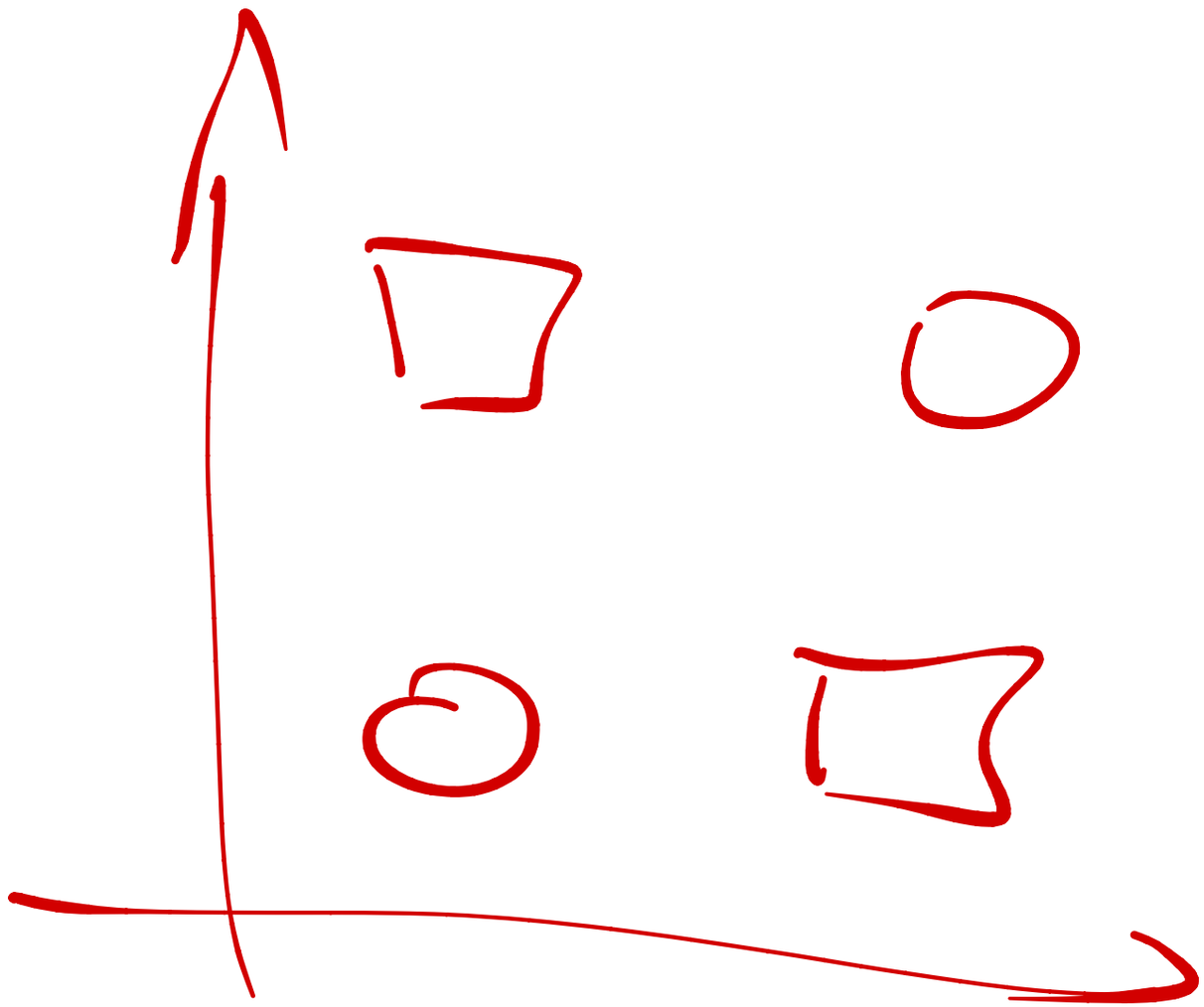
$$z^{(i)} = f(w_t x'^{(i)})$$

$$= f(w_{1t} \cdot x_1 + \dots + w_{nt} \cdot x_n + b)$$

$$w_{i(t+1)} = w_{it} + r \cdot (y^{(i)} - z^{(i)}) x_i^{(i)}$$



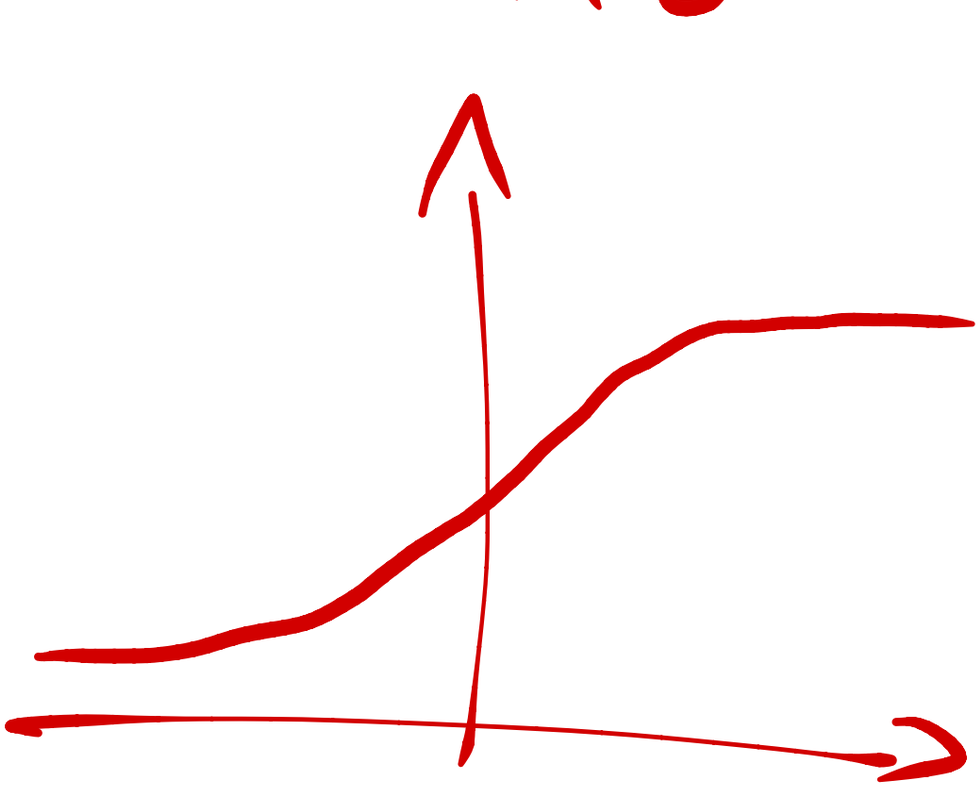
What the Perceptron Can't Learn



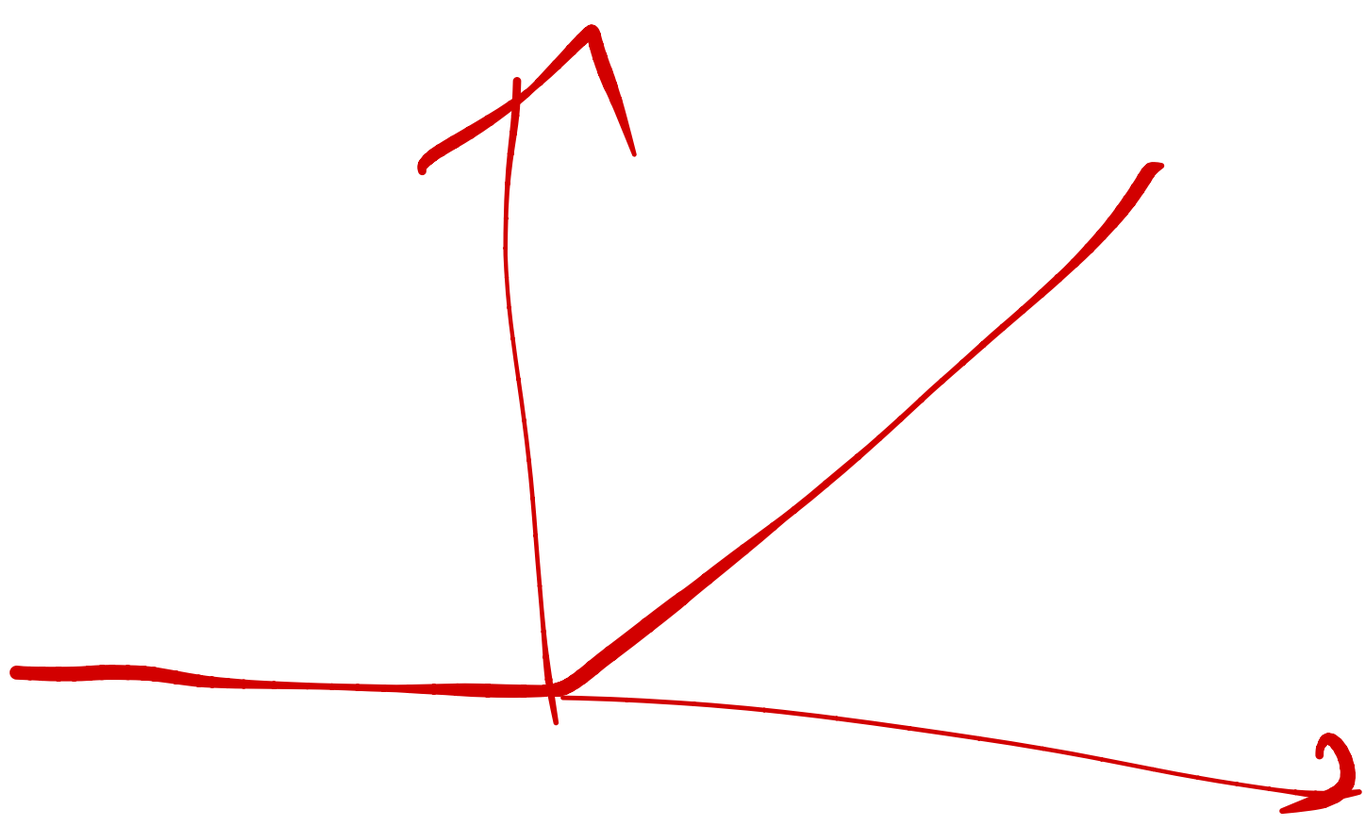
Other Activation Functions

sigmoid

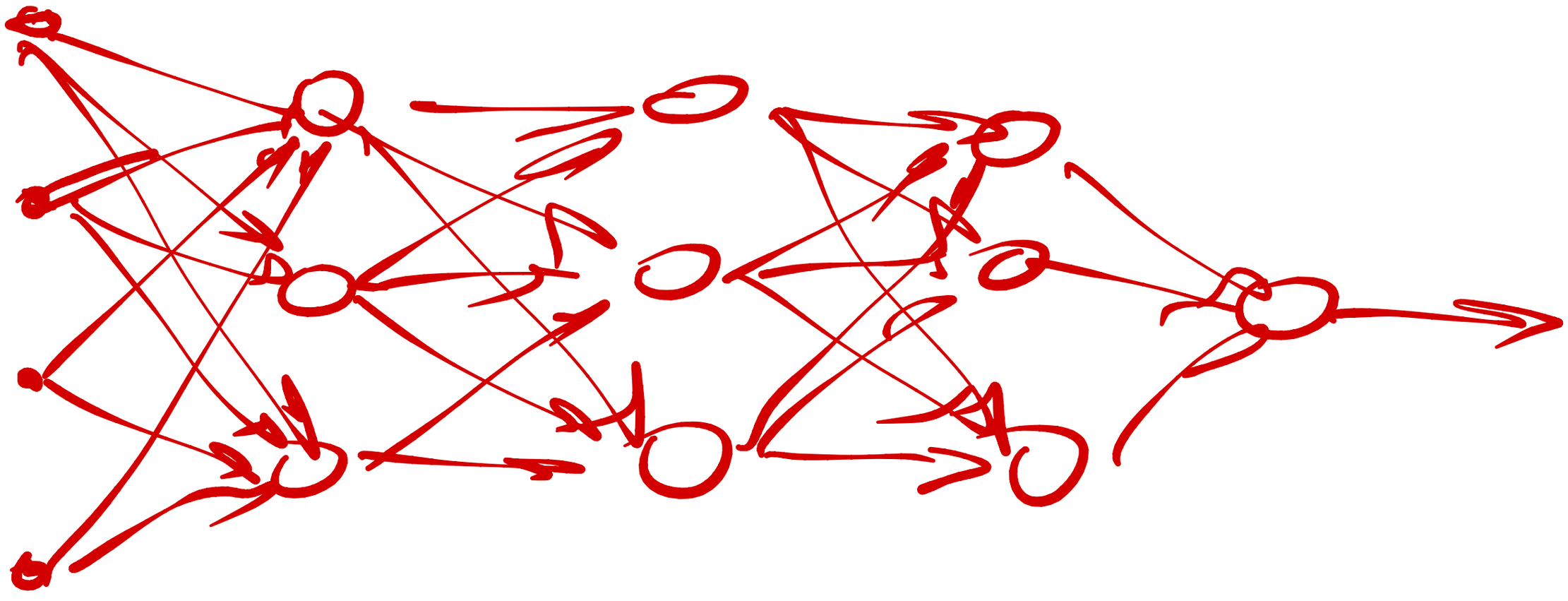
$$\sigma(z) = \frac{1}{1+e^{-z}}$$



ReLU



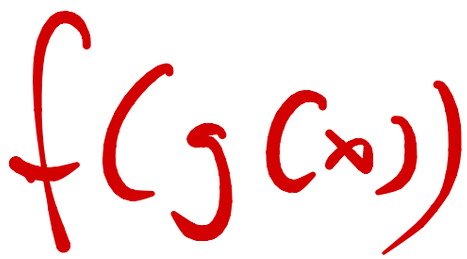
Deep (= Multilayer) Networks



$g(Wx+b)$

How does the supervision signal get to each weight?

- Backpropagation



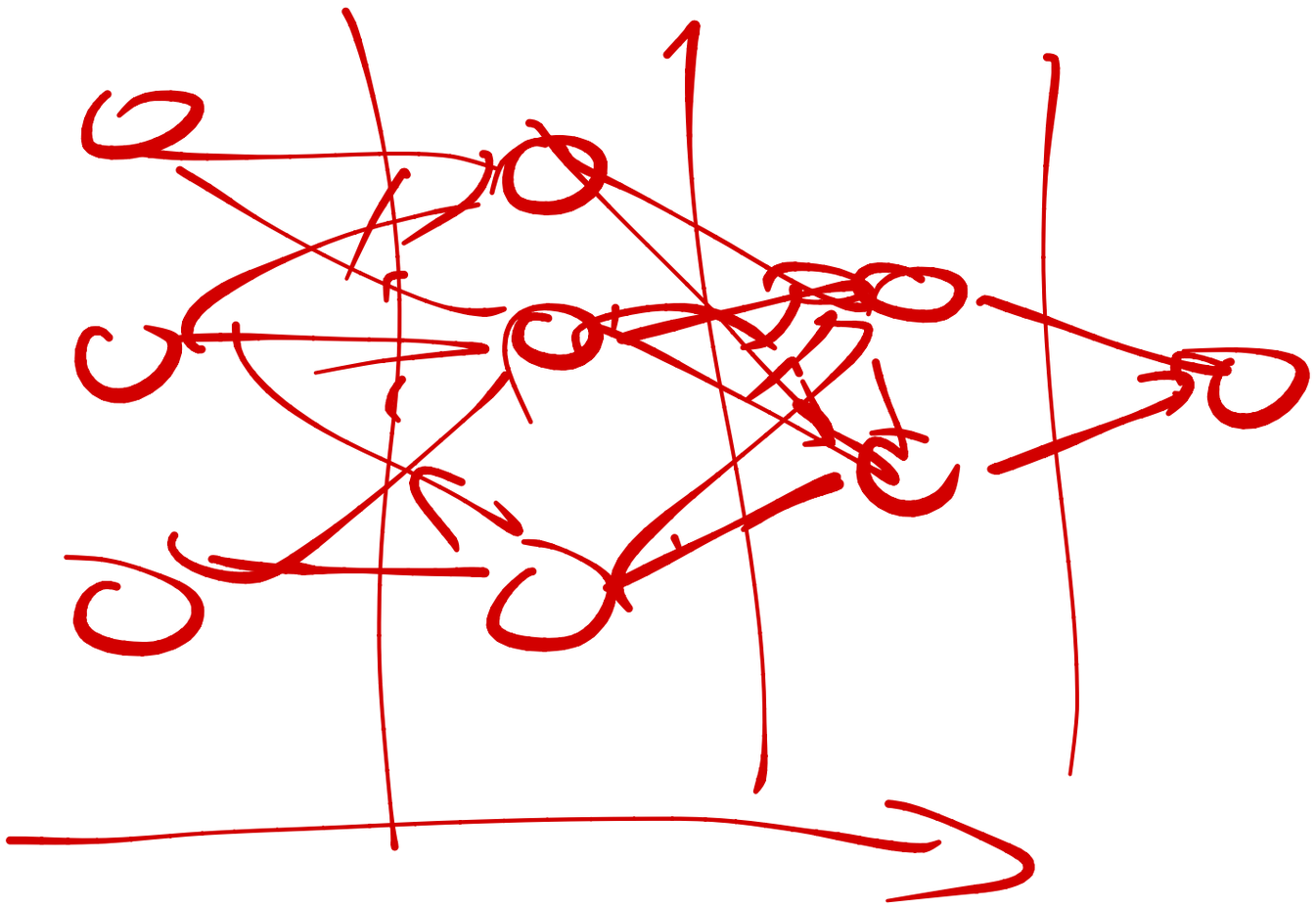
The Computation Graph

Theano

Tensorflow

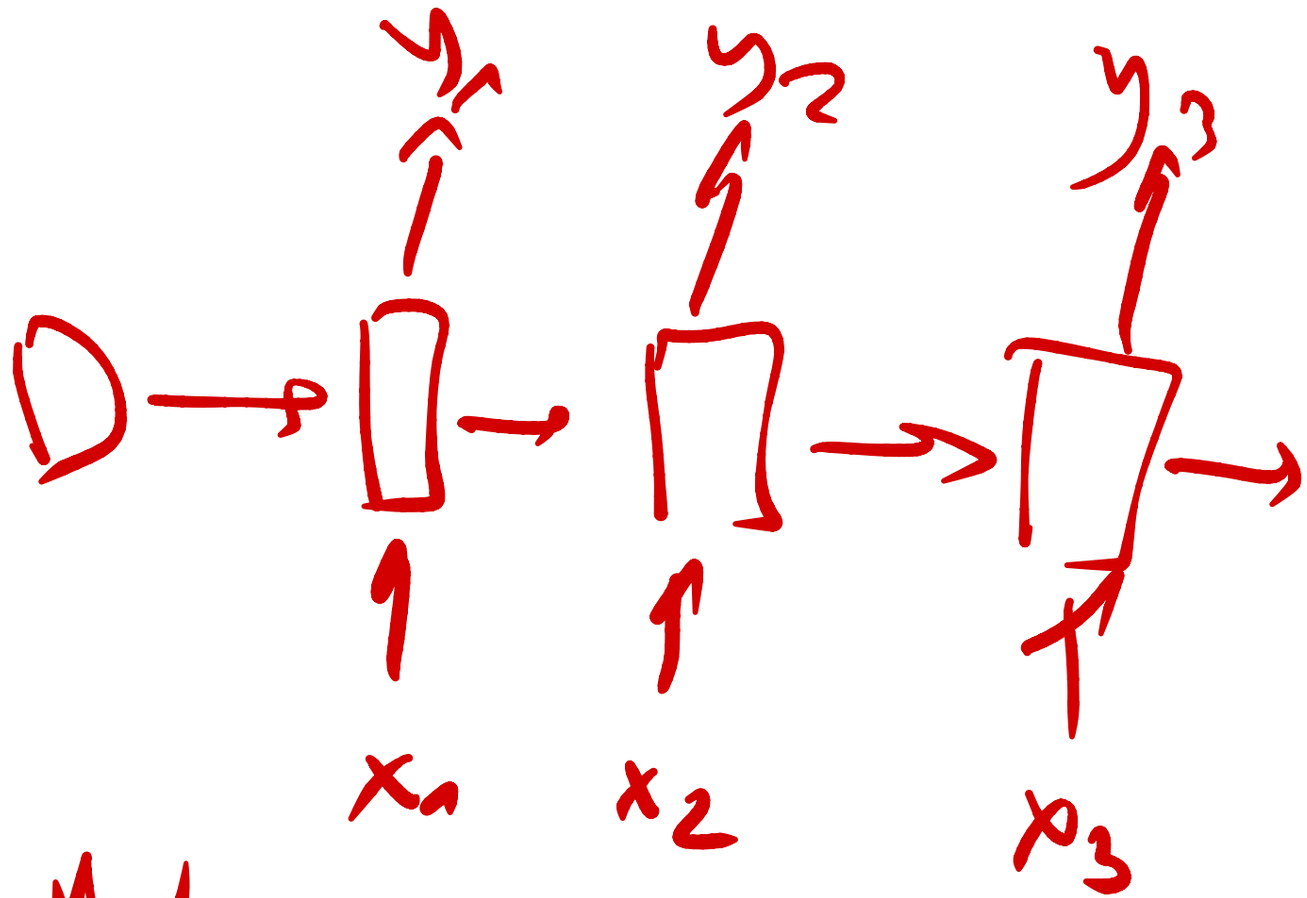
PyTorch

Fully Connected (Feed Forward) Network



Convolutional Networks

Recurrent Networks



RNN

LSTM

Overview

- Learning & Machine Learning
- Everything's a Function
- Experience
- Types of Machine Learning
- Supervised Learning: X, Y, h, and θ
- Loss and Cost (as a function of θ), and Gradient Descent
- Representations
- The Perceptron and Its Learning Rule; What it Can't Do
- Weights & Activation Functions
- Backpropagation
- The Computation Graph
- Fully Connected Neural Networks
- Other Types of NN Architectures

advertisement

- The CL Colloquium! Wednesdays, 16-18h. https://github.com/compling-potsdam/sose19-cl-colloquium

Week	Date	Presenter(s)	Title / Abstract
01	2019-04-10	*no talk scheduled*	
02	2019-04-17	*David Schlangen*	*Information on getting credit for this course*
03	2019-04-24	CL students	Poster Slam
04	2019-05-01	*public holiday*	
05	2019-05-08	Alan Nichols (RASA)	tba
06	2019-05-15	Shlomi Hod (Potsdam)	tba
07	2019-05-22	Alan Akbik (Zalando)	tba
08	2019-05-29	*no talk scheduled*	
09	2019-06-05	*no talk scheduled*	
10	2019-06-12	Staffan Larsson (Gothenburg)	tba
11	2019-06-19	Feiyu Xu (Lenovo) *to be confirmed*	tba
12	2019-06-26	Milica Gasic (Düsseldorf)	tba
13	2019-07-03	Raquel Fernández (Amsterdam)	tba
14	2019-07-10	*no talk scheduled*	
15	2019-07-17	*no talk scheduled*	