Νευρο-Ασαφής Υπολογιστική
Neuro-Fuzzy Computing

Διδάσκων –
Δημήτριος Κατσαρός

@ Τμ. ΗΜΜΥ
Πανεπιστήμιο Θεσσαλίας

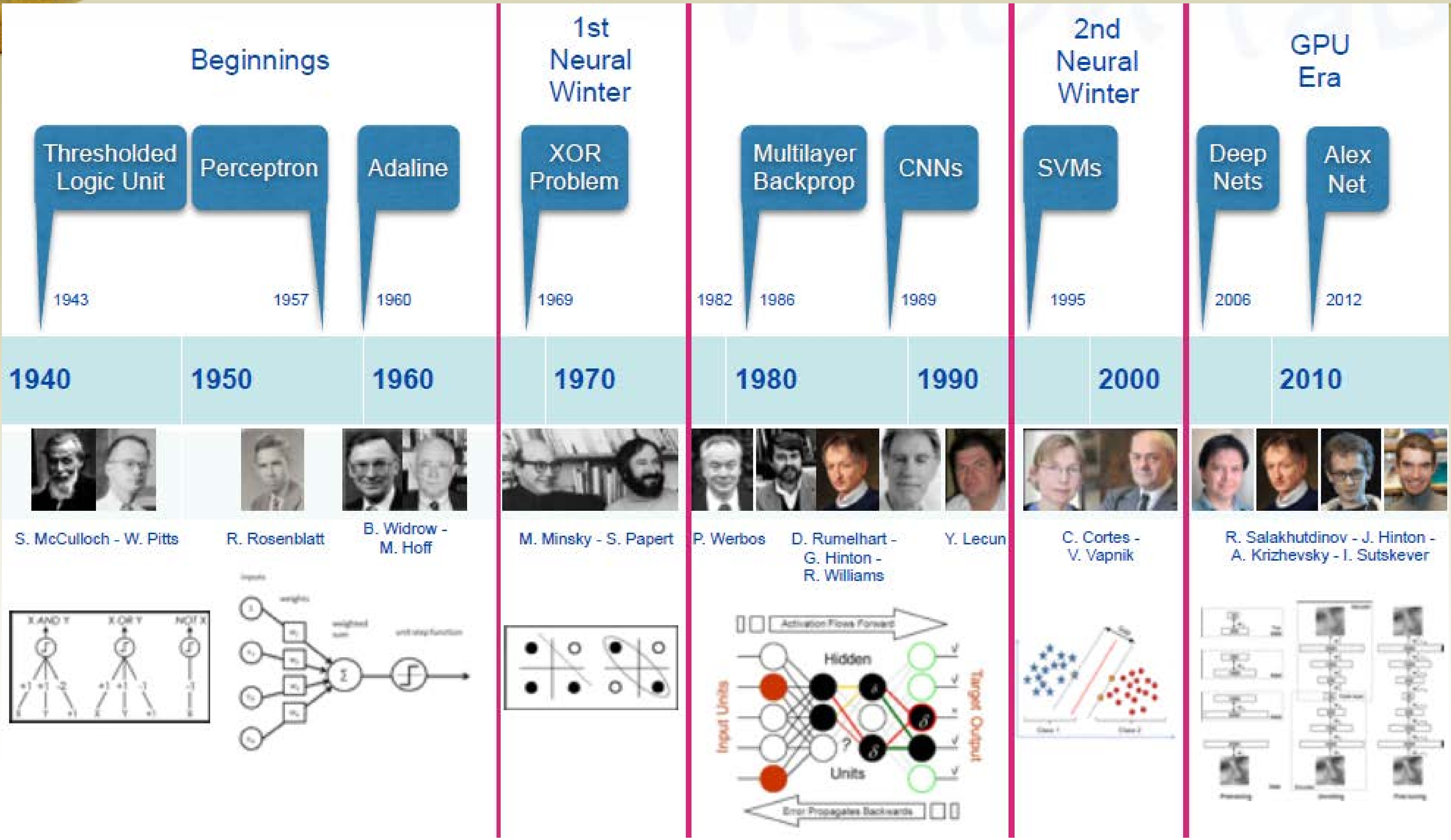
Διάλεξη 1η

Introduction to Neural Networks

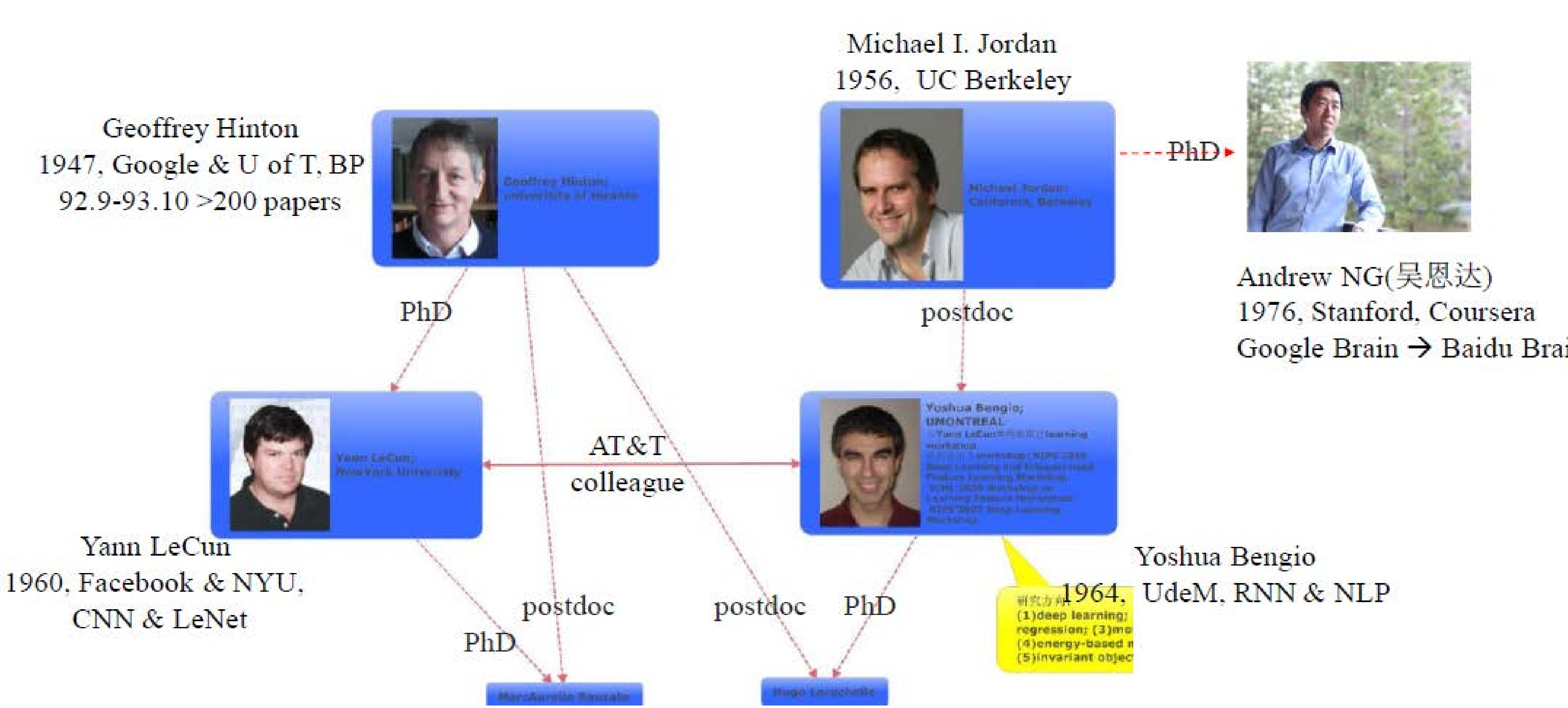
Introduction

- Neural networks (NN),
  - beautiful biologically-inspired programming paradigm which enables a computer to learn from observational data
- Deep Learning (DL)
  - powerful set of techniques for learning in neural networks
- Neural networks and deep learning currently provide the best solutions to many problems in
  - image recognition
  - speech recognition
  - natural language processing

Milestones in NN evolution

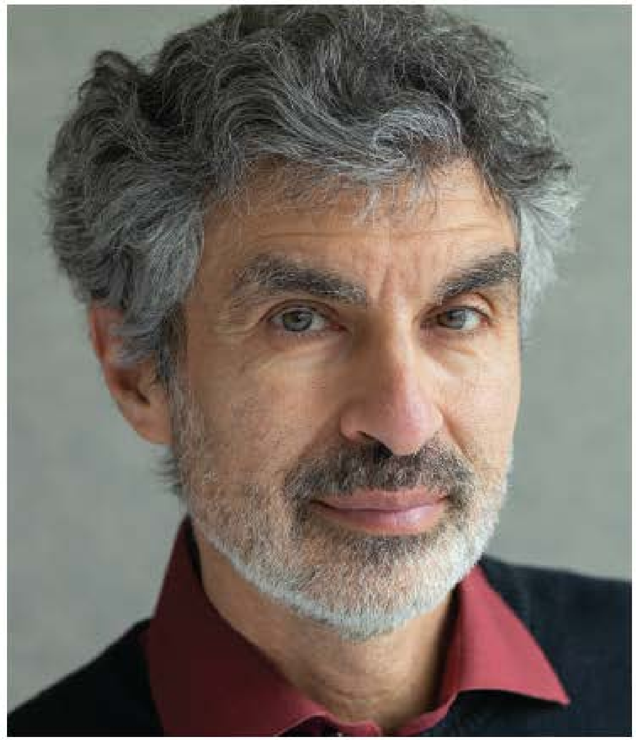


Some contemporary NN persons



Turing Award (called the Nobel Prize of computing) 2018

Geoffrey Hinton



Yoshua Bengio

Yann LeCun

Historical waves of Artificial NN

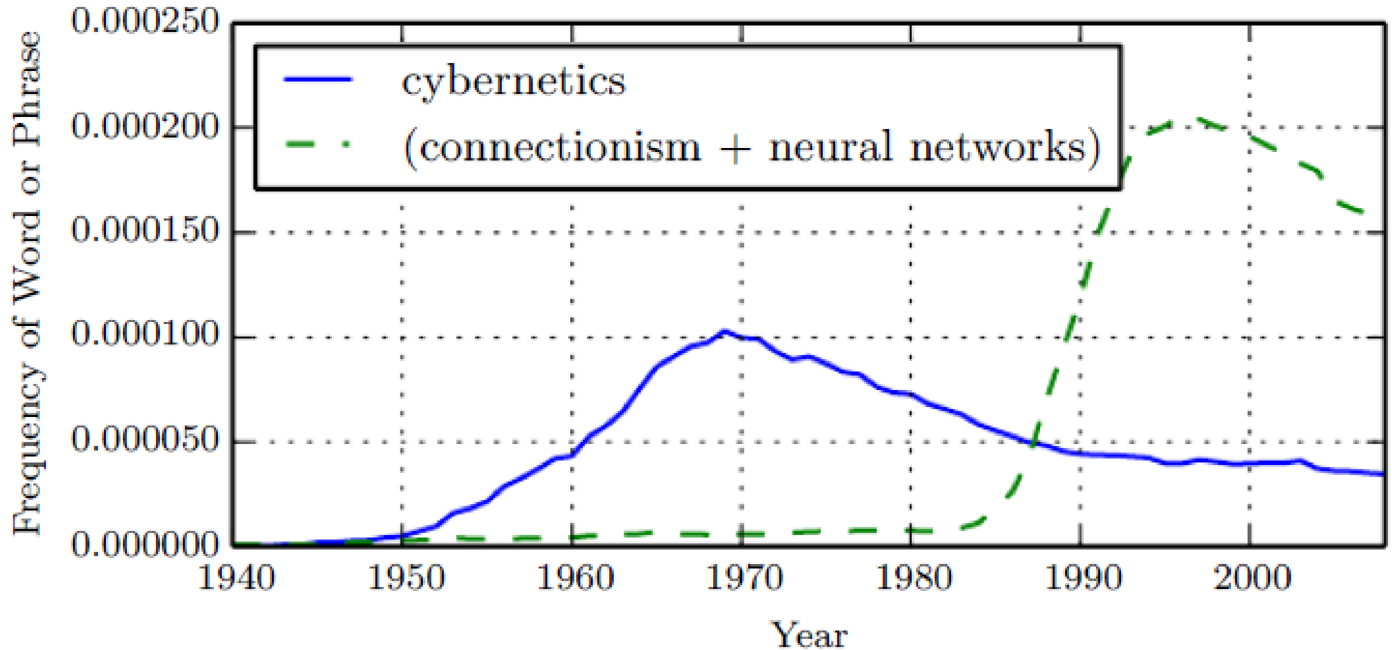
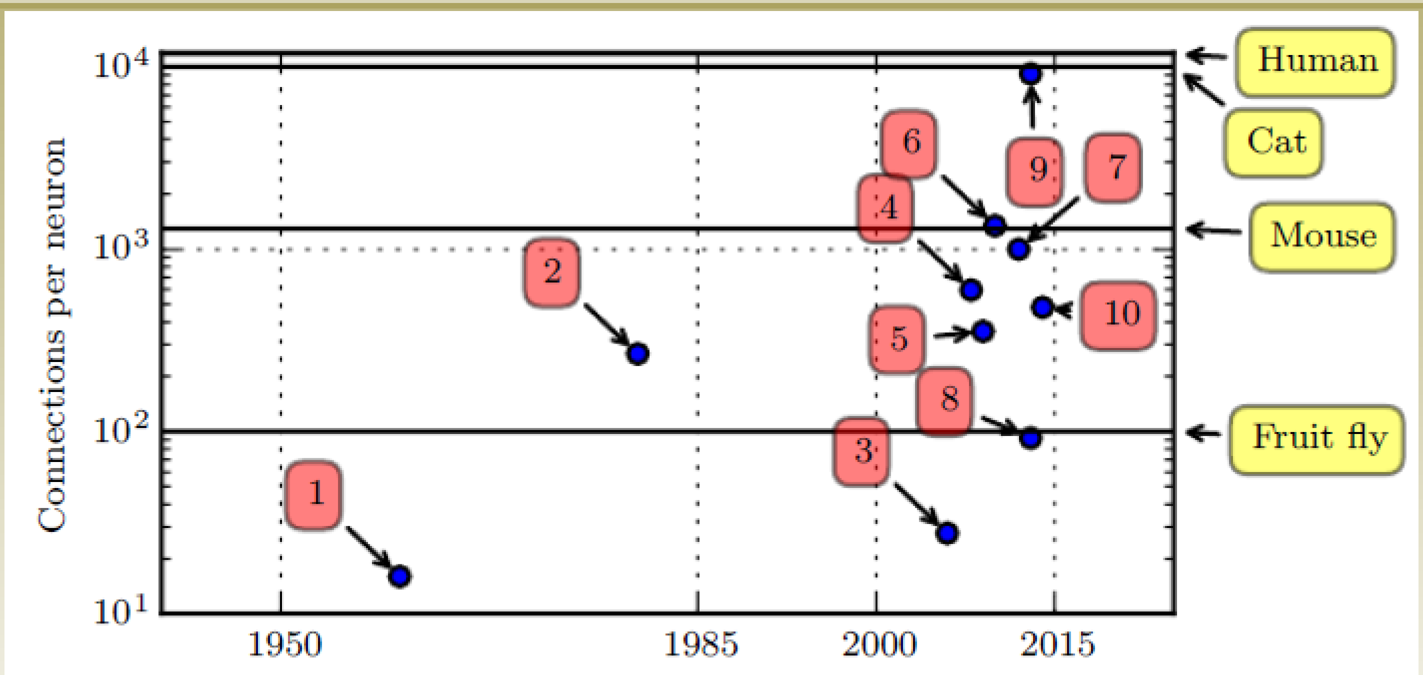


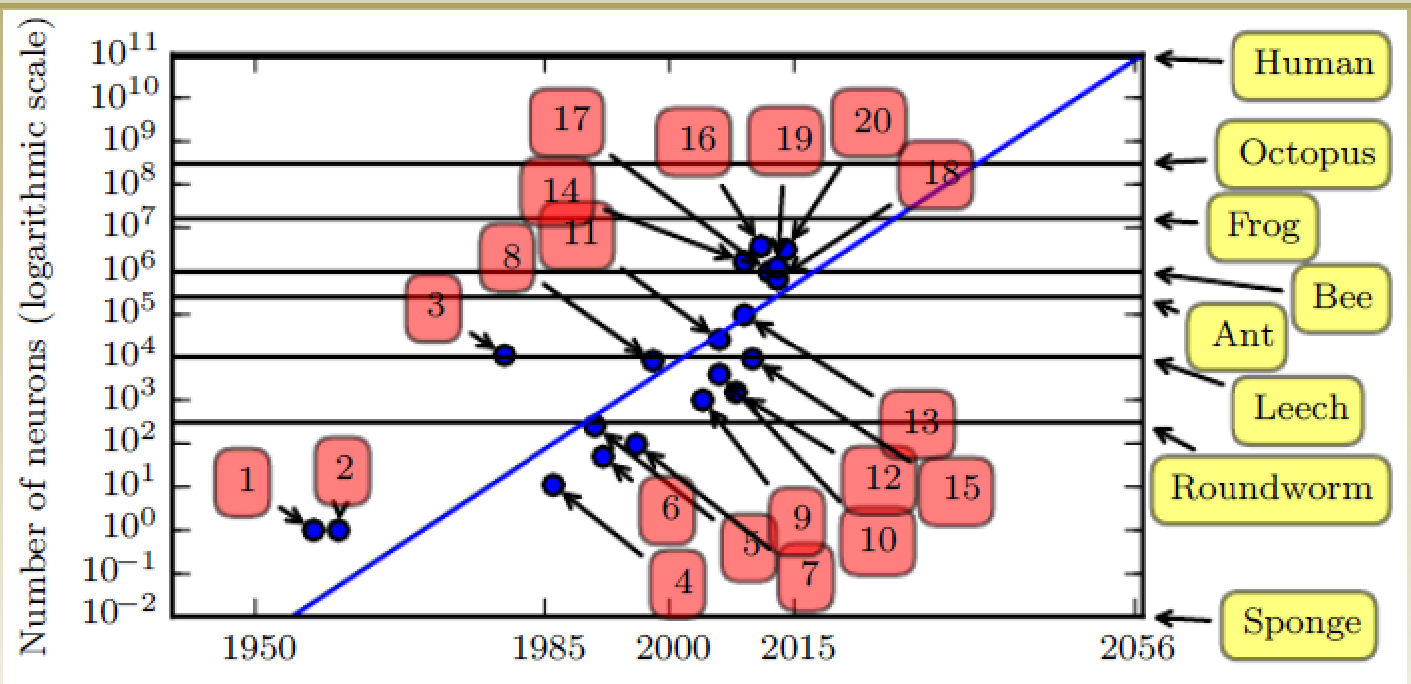
Figure 1.7: The figure shows two of the three historical waves of artificial neural nets research, as measured by the frequency of the phrases "cybernetics" and "connectionism" or "neural networks" according to Google Books (the third wave is too recent to appear). The

ANN Neuron connectivity over time



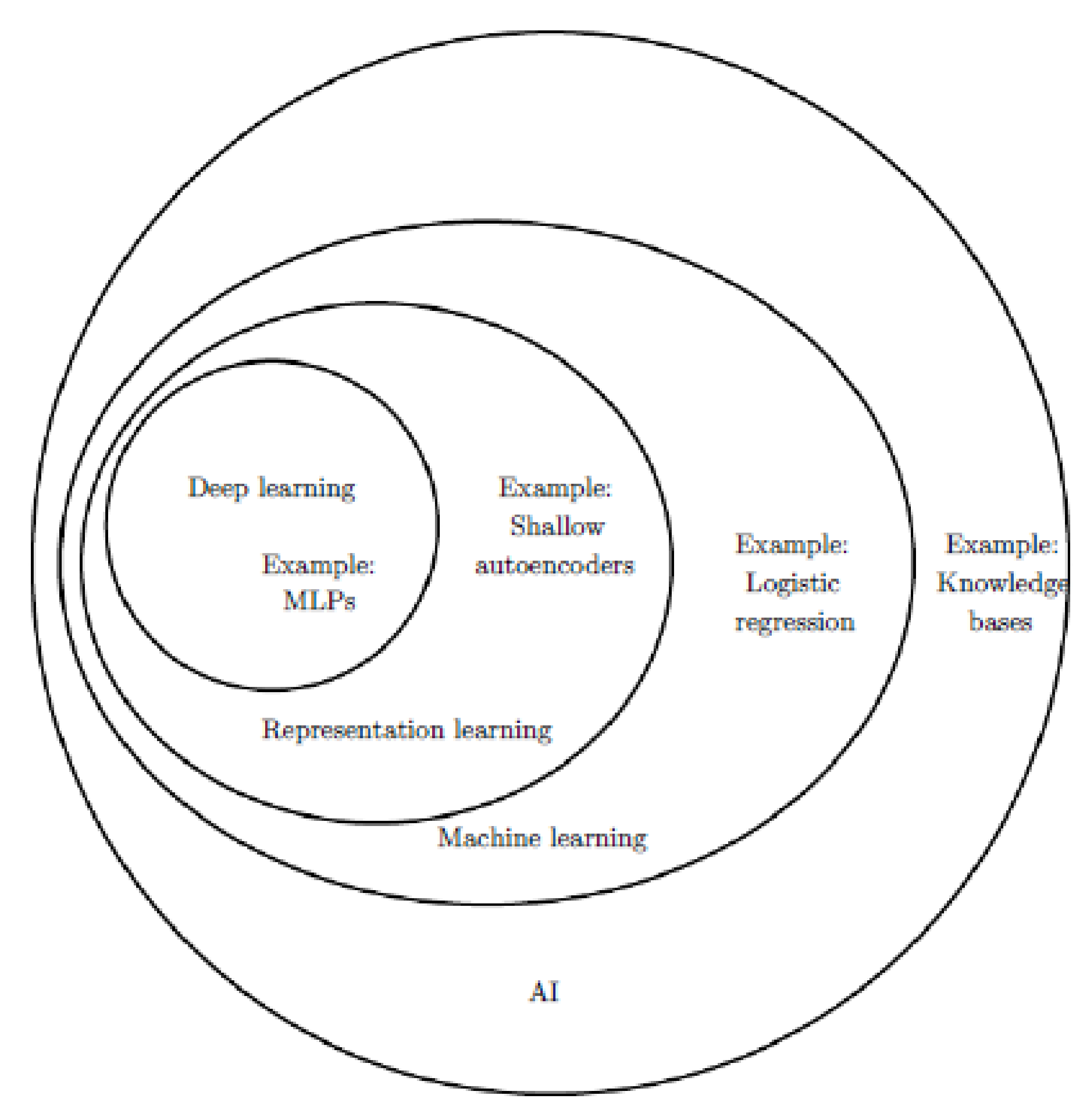
1. Adaptive linear element (Widrow and Hoff, 1960)
2. Neocognitron (Fukushima, 1980)
3. GPU-accelerated convolutional network (Chellapilla et al., 2006)
4. Deep Boltzmann machine (Salakhutdinov and Hinton, 2009a)
5. Unsupervised convolutional network (Jarrett et al., 2009)
6. GPU-accelerated multilayer perceptron (Ciresan et al., 2010)
7. Distributed autoencoder (Le et al., 2012)
8. Multi-GPU convolutional network (Krizhevsky et al., 2012)
9. COTS HPC unsupervised convolutional network (Coates et al., 2013)
10. GoogLeNet (Szegedy et al., 2014a)

NN size over time

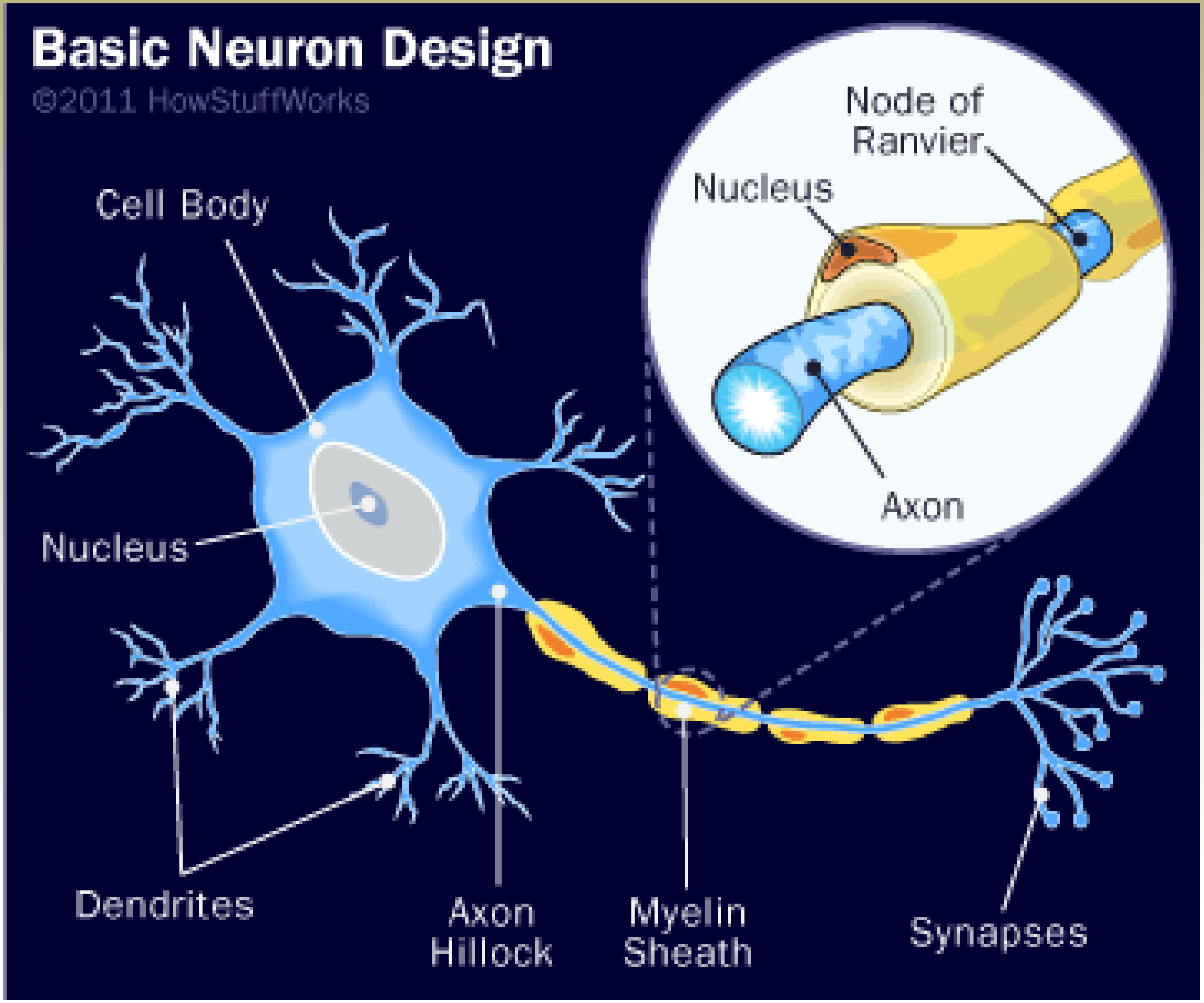


1. Perceptron (Rosenblatt, 1958, 1962)
2. Adaptive linear element (Widrow and Hoff, 1960)
3. Neocognitron (Fukushima, 1980)
4. Early back-propagation network (Rumelhart et al., 1986b)
5. Recurrent neural network for speech recognition (Robinson and Fallside, 1991)
6. Multilayer perceptron for speech recognition (Bengio et al., 1991)
7. Mean field sigmoid belief network (Saul et al., 1996)
8. LeNet-5 (LeCun et al., 1998b)
9. Echo state network (Jaeger and Haas, 2004)
10. Deep belief network (Hinton et al., 2006)
11. GPU-accelerated convolutional network (Chellapilla et al., 2006)
12. Deep Boltzmann machine (Salakhutdinov and Hinton, 2009a)
13. GPU-accelerated deep belief network (Raina et al., 2009)
14. Unsupervised convolutional network (Jarrett et al., 2009)
15. GPU-accelerated multilayer perceptron (Ciresan et al., 2010)
16. OMP-1 network (Coates and Ng, 2011)
17. Distributed autoencoder (Le et al., 2012)
18. Multi-GPU convolutional network (Krizhevsky et al., 2012)
19. COTS HPC unsupervised convolutional network (Coates et al., 2013)
20. GoogLeNet (Szegedy et al., 2014a)

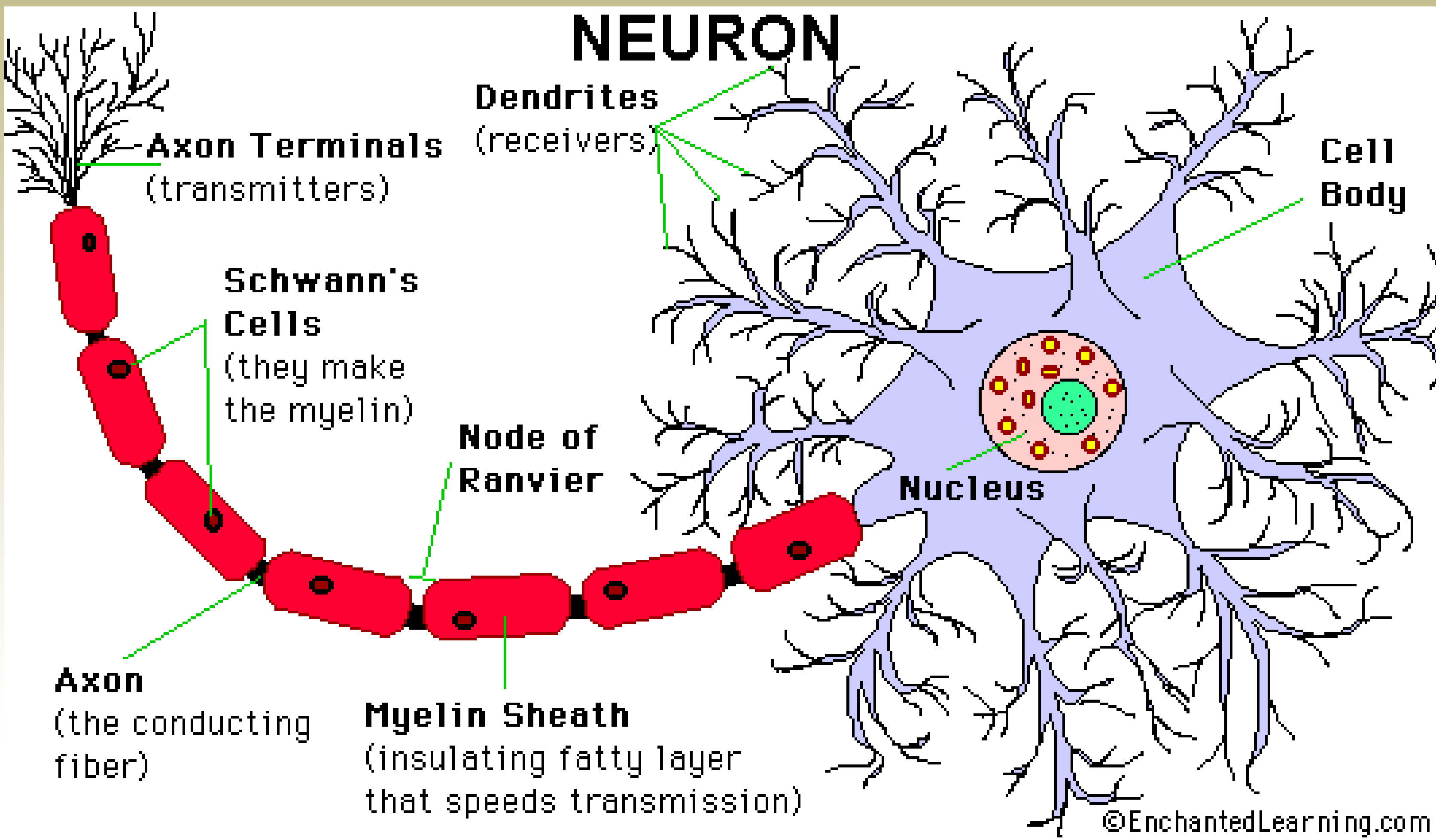
Neural Networks their relation to AI



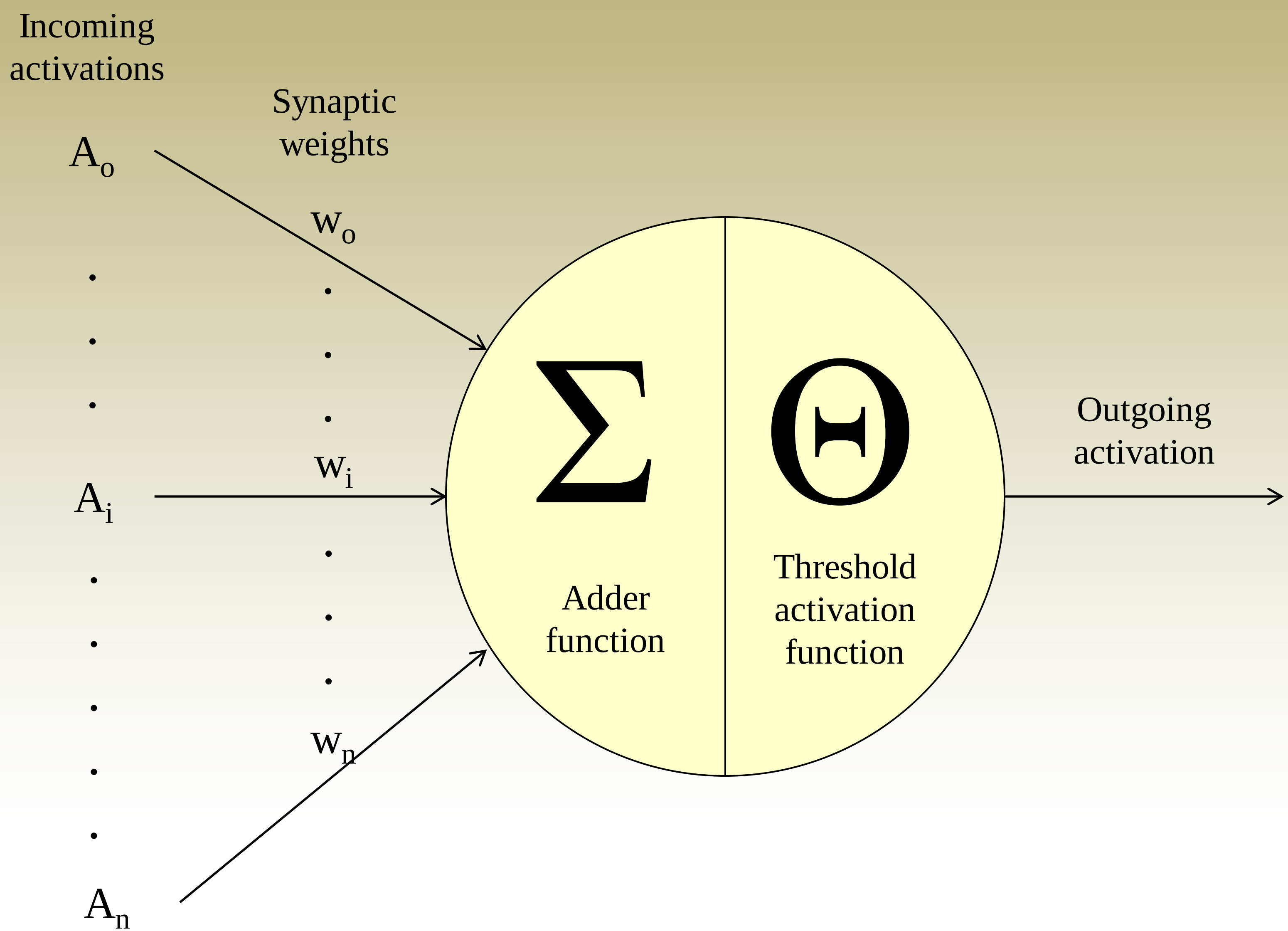
How does a neuron look like?



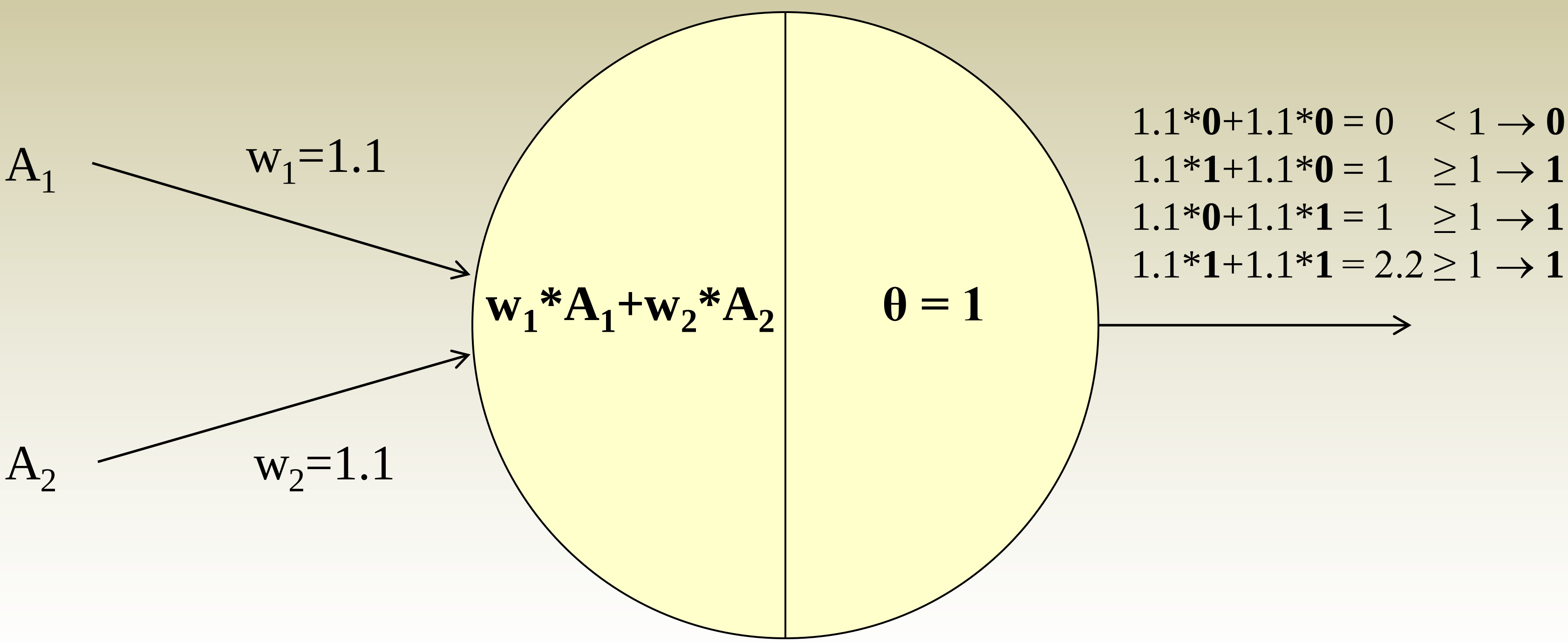
How does a neuron look like?



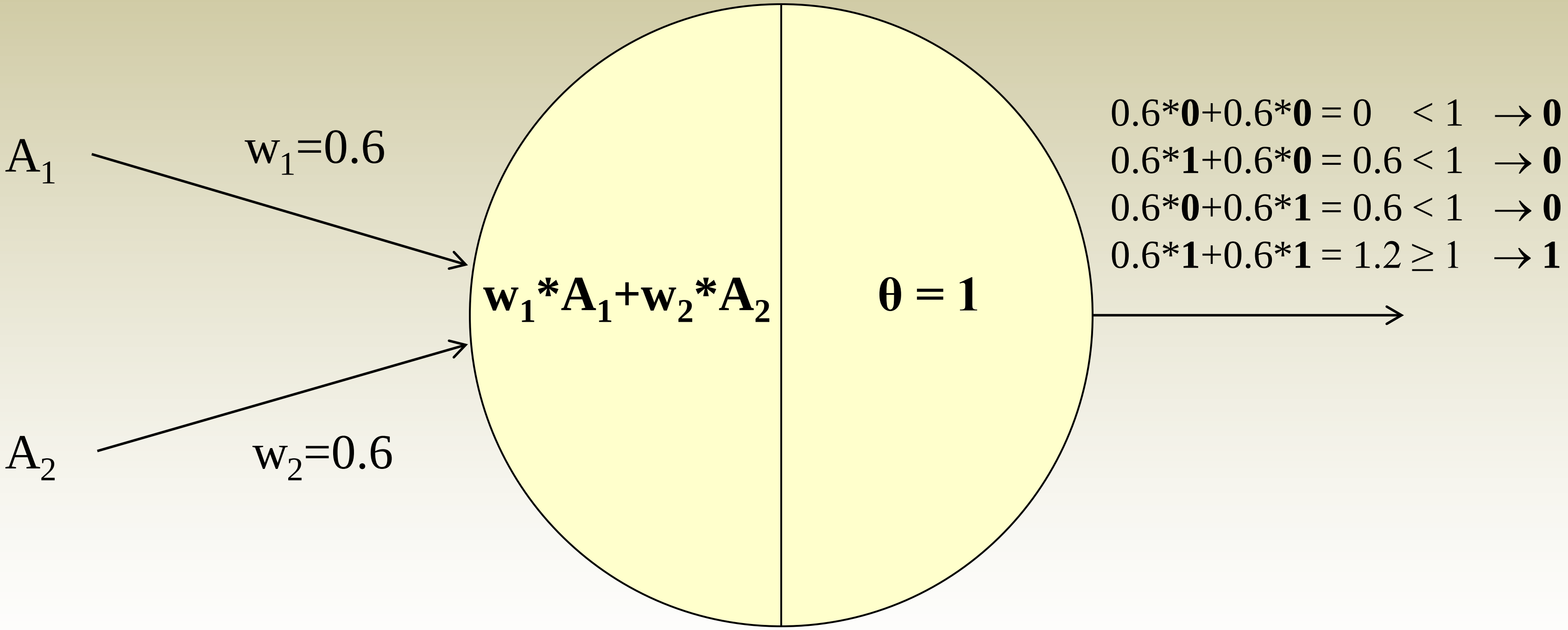
How does a computation neuron look like? The McCulloch-Pitts neuron



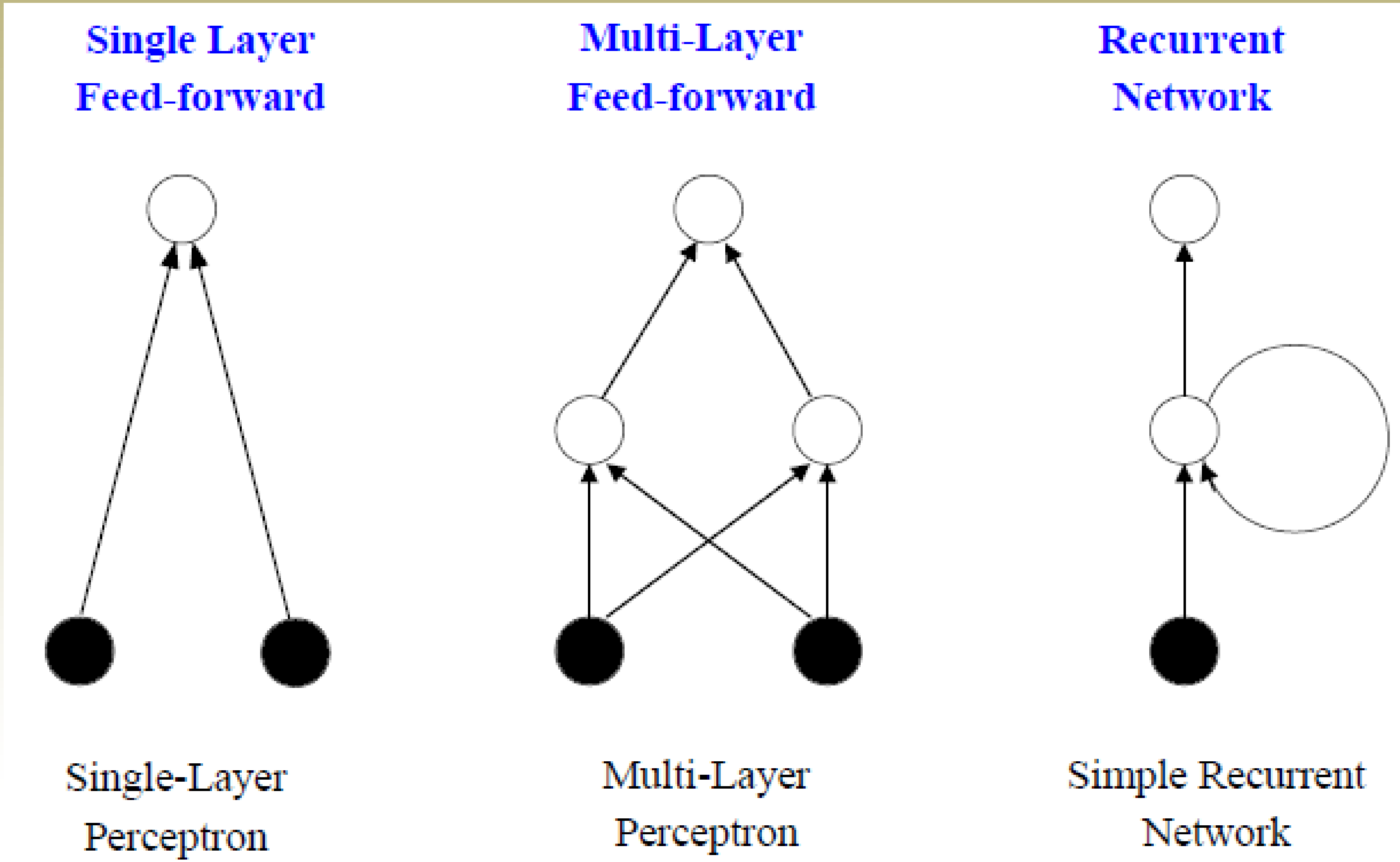
Neural computation of logical OR



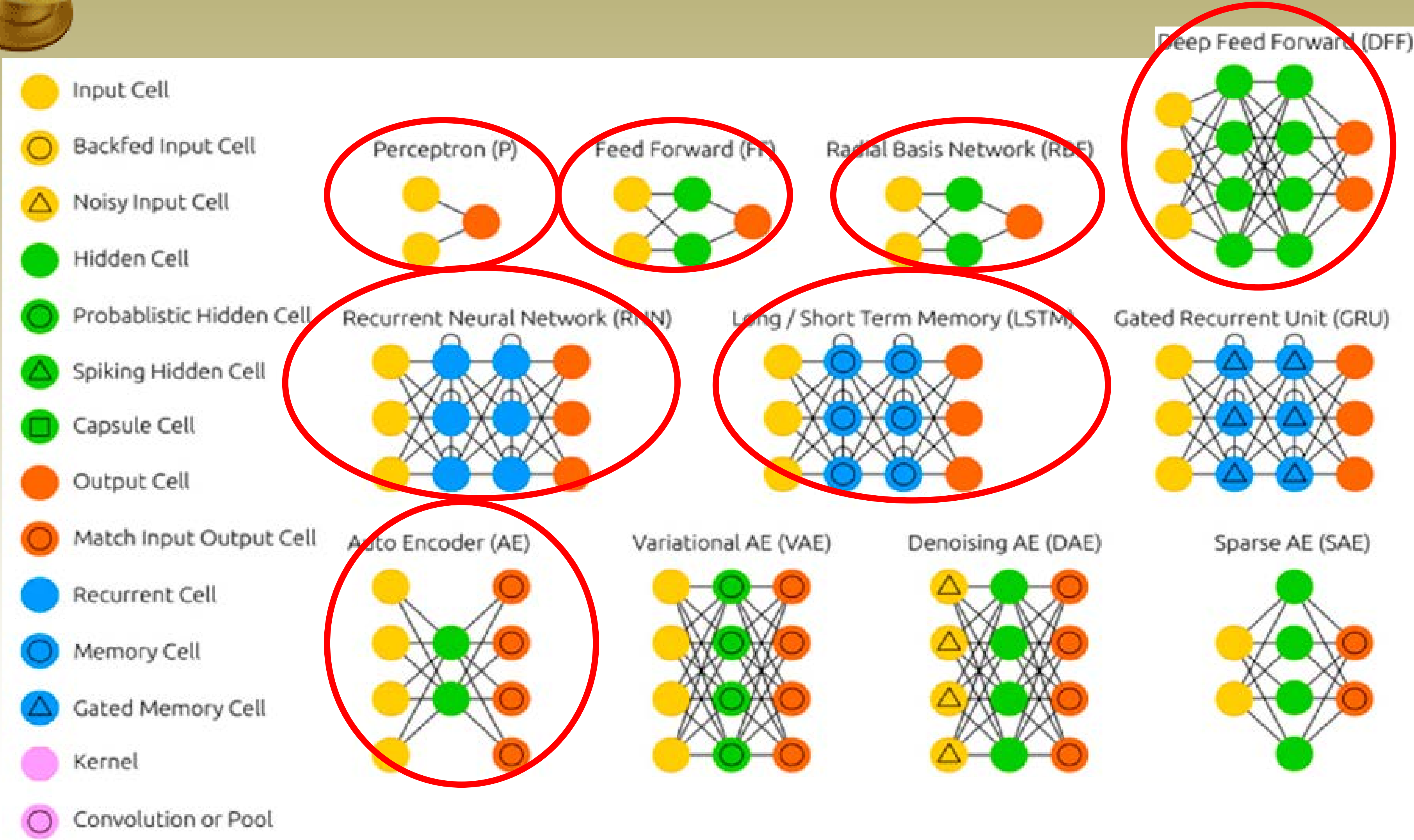
Neural computation of logical AND



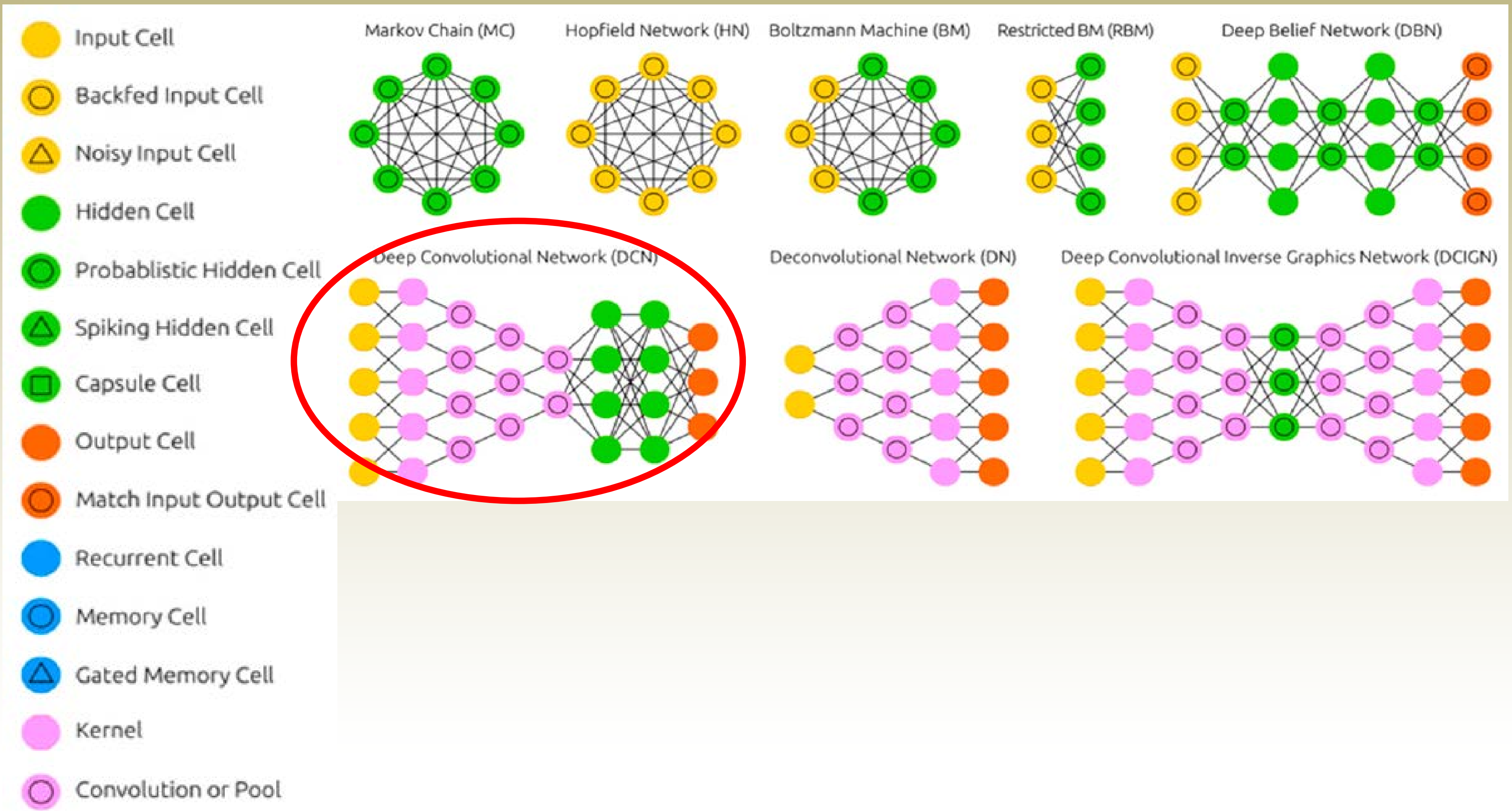
Types of neural networks



Which architectures we will deal with?



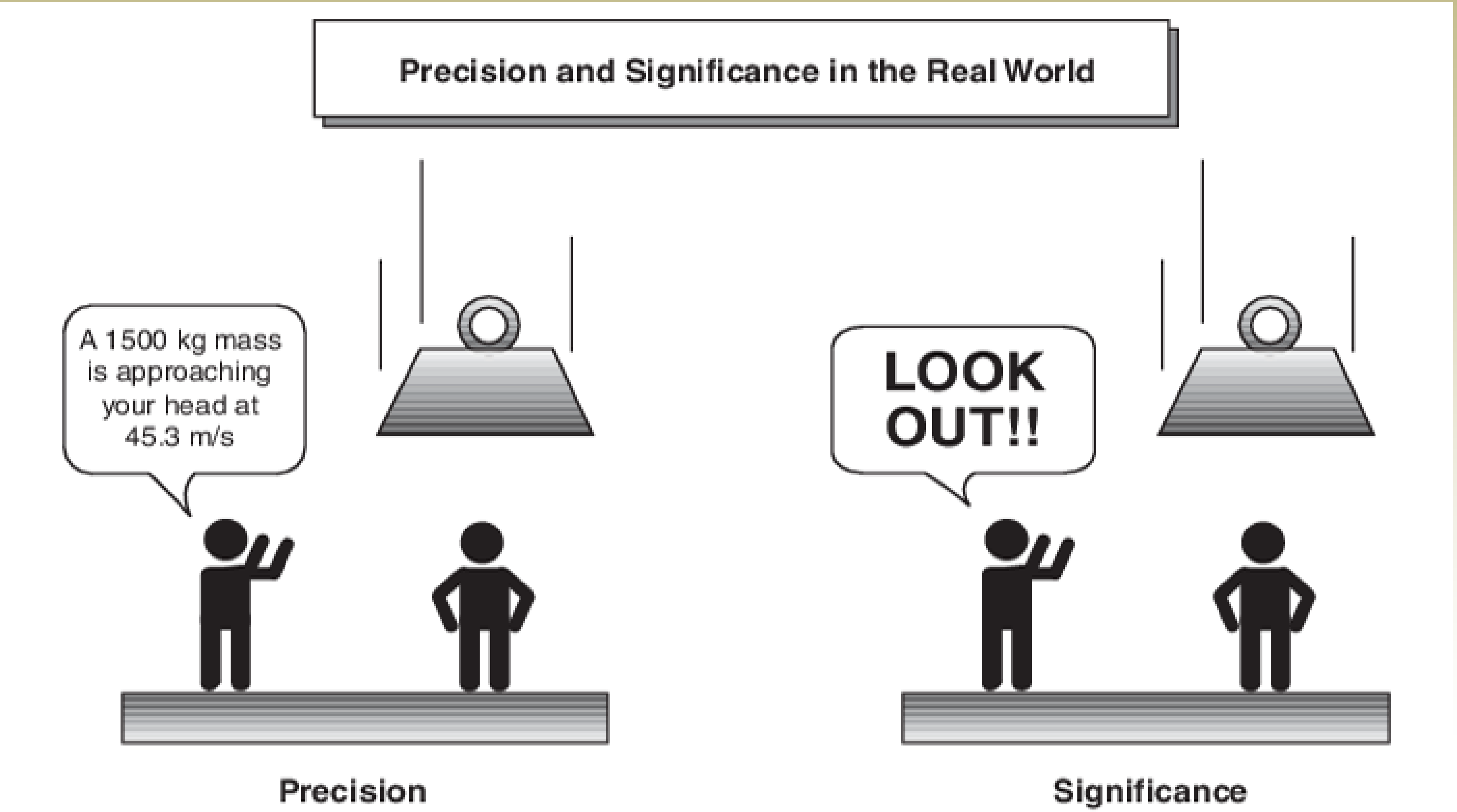
Which architectures we will deal with?



Which architectures we will deal with?

Introduction to Fuzzy Logic

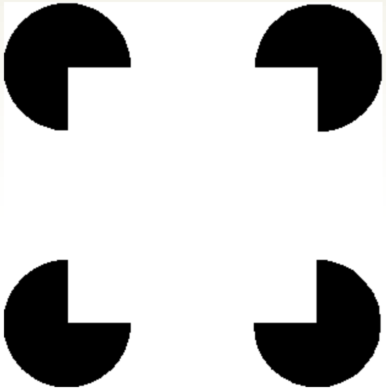
Precision and Significance



Neural pre-attentive processing

- The human visual systems behaves as an adaptive system: The Kanizsa square exists in our brain, not 'out there'
- The four symmetric ink patterns are *noumena*, "things in themselves", according to Immanuel Kant
- The noumena-induced sensation produces the Kanizsa-square *phenomenon* or perception in our brain
- The real-time interaction of millions of competing and cooperating neurons produces the Kanizsa-square illusion, and everything we "see"

The Kanizsa square



Bivalent paradoxes: Where classic logic fails

- Does the liar from Crete lies when he says that all Cretans are liars?
- Russell's barber is a man in a town who advertises his services with the logo: "I shave all, and only, those men who do not shave themselves"
  - Who shaves the barber?
- Consider the card that says on one side "The sentence on the other side is true", and says on the other side "The sentence on the other side is false"

- These paradoxes have the same form: A statement *S* and its negation *not-S* have the same *truth-value t(S)*:

$$t(S) = t(not\text{-}S) \qquad (1\text{-}1)$$

Bivalent paradoxes: Where classic logic fails

- The two statements are both TRUE or both FALSE
- This violates the laws of noncontradiction [*not-(A and not-A)*] and excluded middle [*either A OR not-A*]
- For bivalent truth tables, it holds that:

$$t(not\text{-}S) = 1 - t(S) \tag{1-2}$$

- So, it reduces to

$$t(S) = 1 - t(S) \tag{1-3}$$

- If S is true, if $t(S)$=1, then 1=0. $t(S)$=0 also implies the contradiction 0=1.
- The fuzzy or multivalued interpretation accepts the logical relation (1-3), and instead of insisting that $t(S)$=0 or $t(S)$=1, simply solves for t(S) in (1-3):

$$\mathbf{t(S) = ½} \tag{1-5}$$

Bivalent paradoxes: Where classic logic fails

- Multivaluedness also resolves the classical *sorites* paradoxes.
- Consider a heap of sand.
- Is it still a heap if we remove one grain of sand?
- How about two grains? Three?
- If we argue bivalently by induction, we eventually remove all grains and still conclude that a heap remains, or that it has suddenly vanished
- No single grains takes us from heap to nonheap

Bivalent paradoxes: Where classic logic fails

- Suppose there are *n* grains of sand in the heap. Removing one grain leaves *n-1* grains and a truth value $t(S_{n-1})$ of the statement S_{n-1} that the *n-1* sand grains are a heap
- In general, $t(S_{n-1}) < 1$, and $t(S_{n-1})$ may be close to unity but we have a non-zero doubt d_{n-1} about the truth of the matter
- So, $t(S_n) = 1 – d_n$ where $0 \leq d_n \leq \ldots\ d_{n-m} \leq 1$
- Inductively, $t(S_n \rightarrow S_{n-m}) = (1-d_{n-k})$
- If we interpret the conjunction operator as the minimum operator, we have:

$$t(S_n \rightarrow S_{n-m}) = \min(1-d_n, \ldots, 1-d_{n-m}) = 1 - max(d_n, \ldots, d_{n-m})$$

Fuzzy logic

- Polish Jan Lukasiewicz in 1930 first introduced a three-value logic (inspired by Heisenberg uncertainty principle – quantum theory)
- Lotfi Zadeh introduced (instead of the bivalent indicator function) the membership function $m_A : X \rightarrow [0\ldots1]$
- Re-defined union and intersection
  - $I_{A\cap B}(x) = \min(I_A(x), I_B(x))$
  - $I_{A\cup B}(x) = \max(I_A(x), I_B(x))$
  - $I_A c(x) = 1 - I_A(x)$
  - $A \subset B$ iff $I_A(x) \leq I_B(x)$ for all x in X

Fuzzy logic

- The membership value $m_A(x)$ measures the elementhood or **degree** to which element x belongs to set A