Artificial Neural Networks

(Machine Learning and Artificial Intelligence)

Morgado Dias

University of Madeira

BIESA at the Interactive Technologies Institute/Larsys/ARDITI

UNIVERSIDADE da MADEIRA | Interactive Technologies Institute LARSyS

Artificial Intelligence

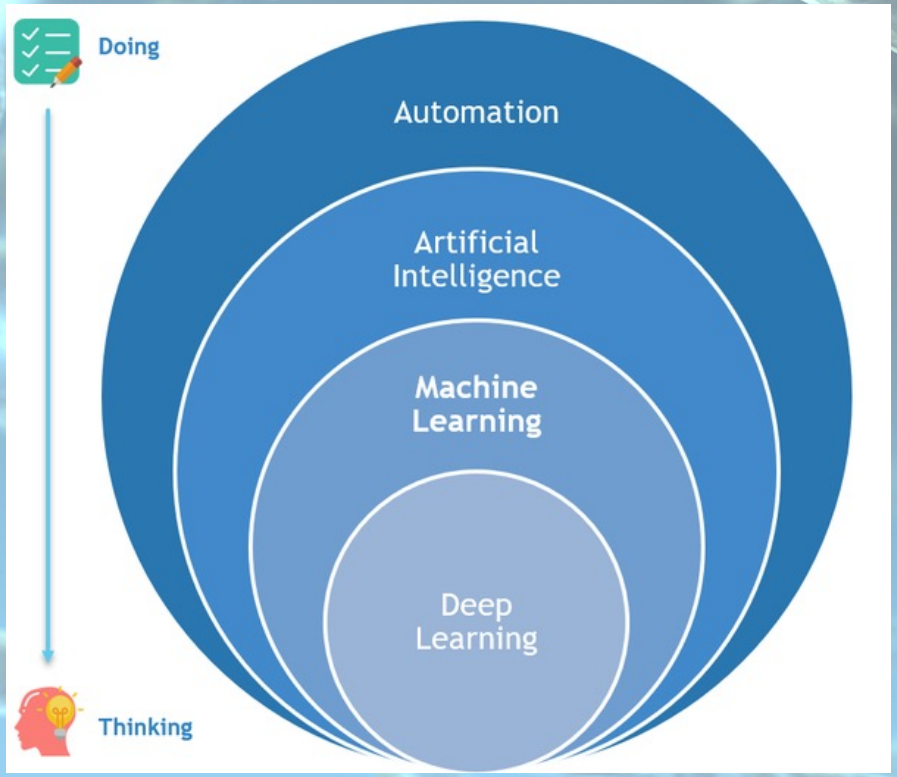
The name Artificial Intelligence corresponds to the capacity of a device to perform actions that are usually connected to human intelligence, such as analysis and decision or reasoning and optimization through experience.

- The many attempts to equip machines (mainly computers) with some sort of intelligence were responsible for the creation of a great diversity of AI tools.
- Some of these tools are part of Machine Learning since they extract knowledge from data.
- But many others forms of Artificial Intelligence exist.

Artificial Intelligence

Artificial Intelligence starts from a desire of automation of processes and resulted in Machine Learning, which corresponds to extracting knowledge from data:

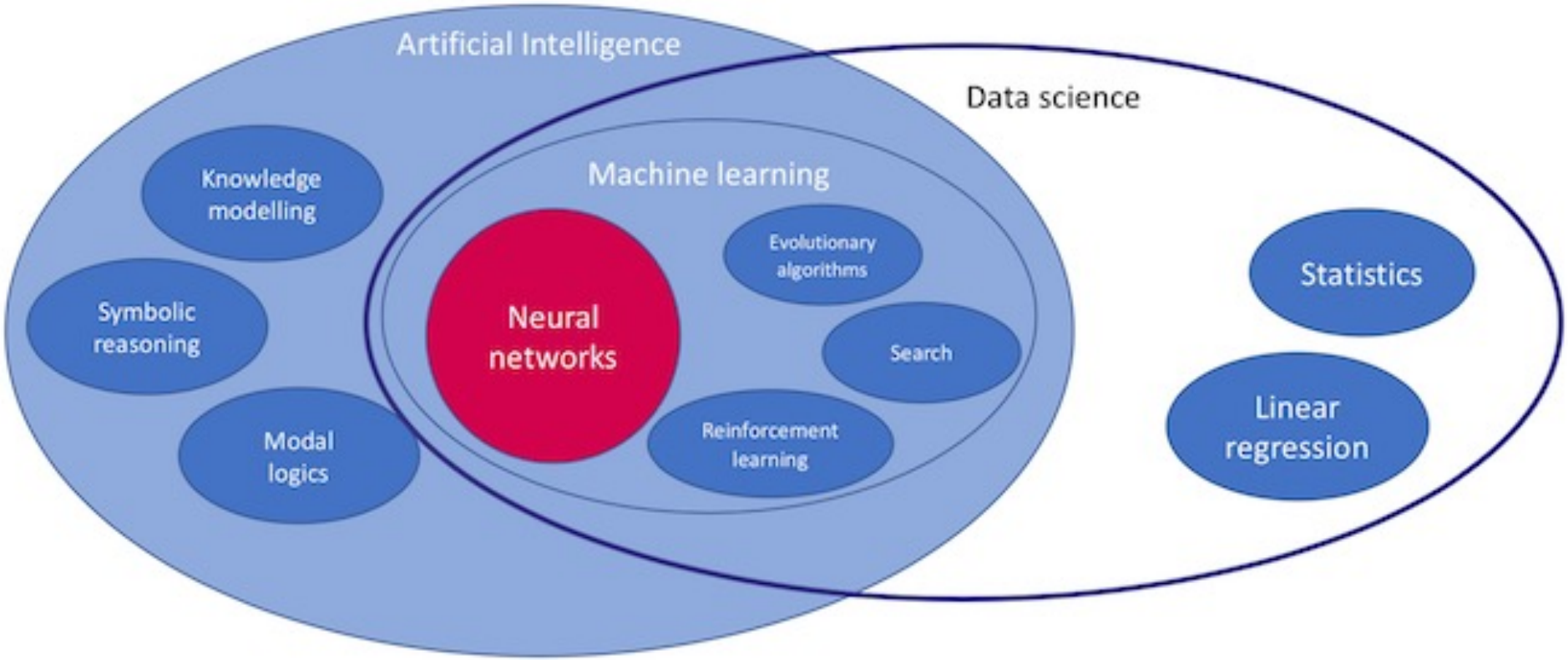
"Machine learning is a subfield of artificial intelligence, which enables machines to learn from past data or experiences without being explicitly programmed"



Artificial Intelligence

And Artificial Intelligence correlates to many other areas that can also cooperate to solve a particular problem.

The focus of this presentation is Neural Networks.



Summary

Introduction

Neural Networks

Artificial Neural Networks

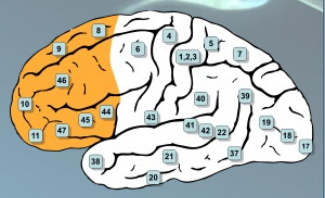
Conclusion

Introduction

Brain study is quite recent in spite of the development of other medical areas. Some relevant facts for ANN development are:

- Egas Moniz, neurologist and diplomat; first Portuguese Nobel prize(1949). Developed leucotomy (or lobotomy), which consists of cutting the connections to and from the prefrontal cortex. Lobotomy is today considered inhumane and is no longer used but it was an advance at the time (showed the use of certain brain areas).

-António Damásio describes in “Descartes´ Error” one example of behavior change as a result of a work accident that severed the brain of a railway worker in the XIXth century.



- Alan Hodgkin and Andrew Huxley, medicine Nobel prize winners in 1963, conducted several tests with the giant axon of squids by applying electric pulses. These test led to the Hodgkin-Huxley model for natural neurons.

Introduction

•Ivan Pavlov studied the conditioned reflex. From his work it became famous the aledged experience in which Pavlov's dogs learned to associate a neutral stimulus (a bell ring) to food.

•Recent tests shown that it is possible to comand rabit's actions through the implantation of electrodes in their brains and applying electrical stimulus with the appropriate duration and intensity in the relevant areas of the brain.

•Animals present a larger neural development in childhood. One test with a young cat whose eye was covered during early development showed that the cat never fully developed vision from that eye.

Introduction

•In 1989 a DARPA report estimated that the largest capacity computer at the time would have, at most, the capacity to simulate the functioning of the brain of a fly.

•Recently it was possible to recover partial vision for patients which were not born blind through the use of a CMOS sensors. The sensor is similar to a camera but it can only be used if the brain was trained to process image. The level of vision recovered was of 10%.

•The processing speed of a natural neuron is 1000 times smaller than the one obtained in electronic circuits.

•In general, the brain is still pretty much **UNKNOWN**!

Introduction

•With more knowledge of the human brain and the fact that there are tasks that are not easy to perform by a PC, many attempts to mimic the brain were developped;

•Tasks difficult for a PC: face recognition, state of mind decoding;

•Tasks that a PC performs better than a human: repetitive calculations;

•Many attempts to mimic nature resulted in failure or are still being pursued:
-moving wings fly modes;
-dragon fly flying type devices;
-dumping systems based in tail movement (such as the cats do).

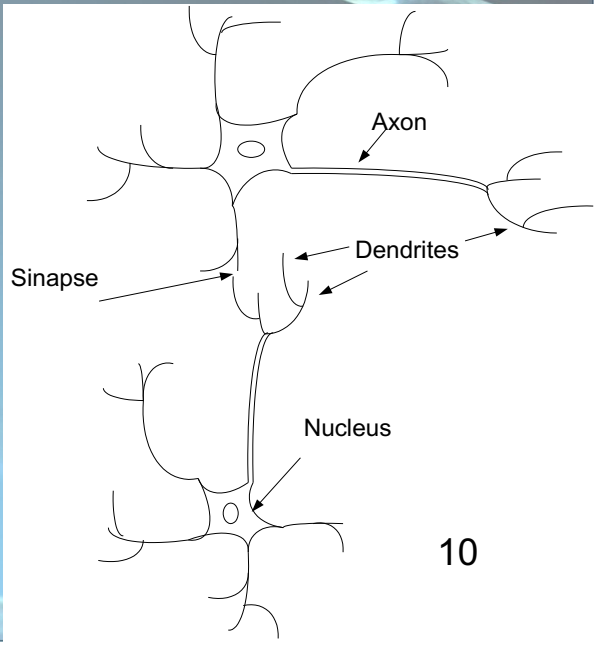
Neural Networks

• The brain is composed of very simple processing units.

• It has about 10 billion neurons with about 10.000 billion synapses. Each neuron has, on average, 1000 synapses (connections) though some might have up to 6000.

• The neurons, functionally, are grouped in neural networks and it is common that they share neurons with other networks.

• Neurons communicate through spikes as a result of certain events.



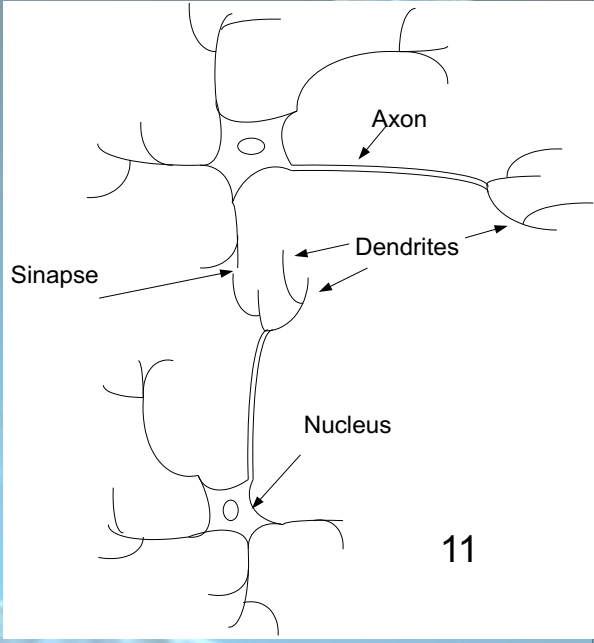
Neural Networks

• The dendrites form a ramified network of nervous fibers with the capacity to transport electric signals to the nucleus.

• The nucleus performs the functions of adding up the inputs affected by their weights and applies a threshold function.

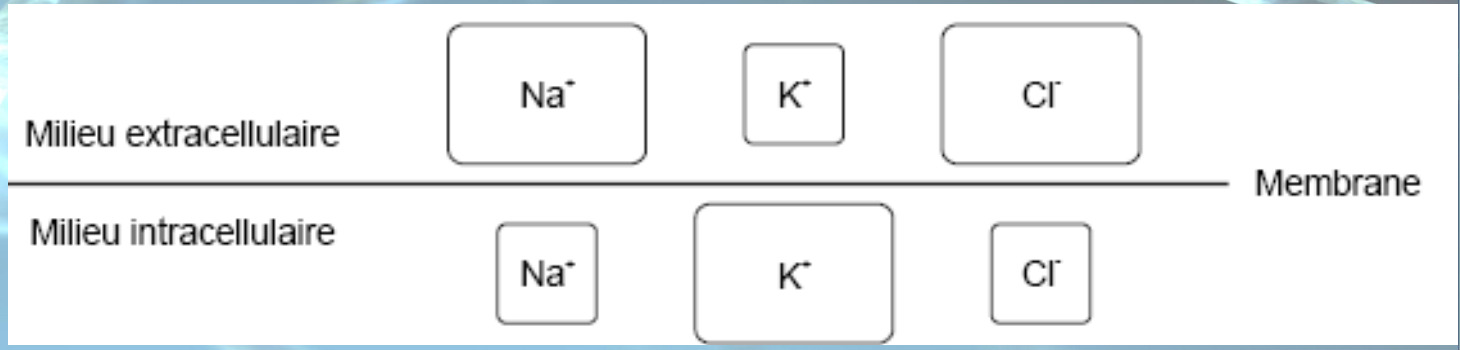
• The axon is a nervous fiber that takes the output signal to other neurons.

• The contact between the axon of one neuron and the dendrite of another neuron is the synapse.



Neural Networks

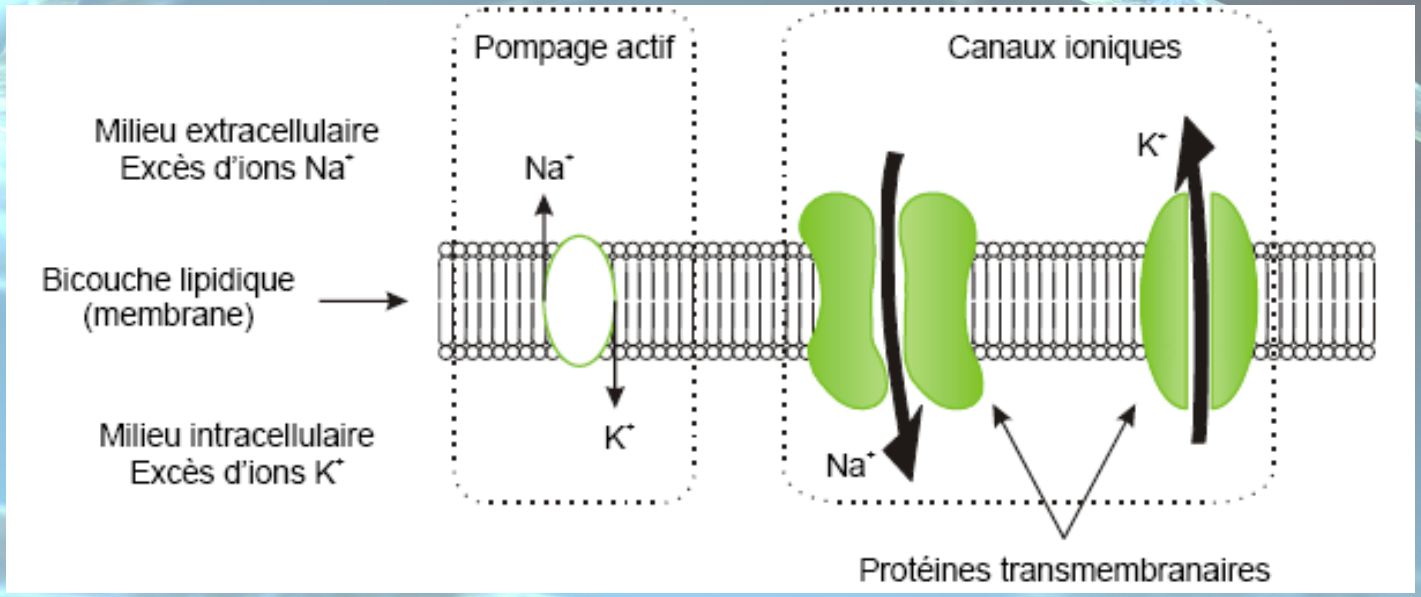
- The neural membrane that limits the cell, separates the intracellular from the extracellular part of the nucleus.
- Both sides have the same ions, but their concentrations are different.



Neural Networks

• The ions can pass the membrane through ionic channels. These channels work like proteins that allow the selective transfer of ions.

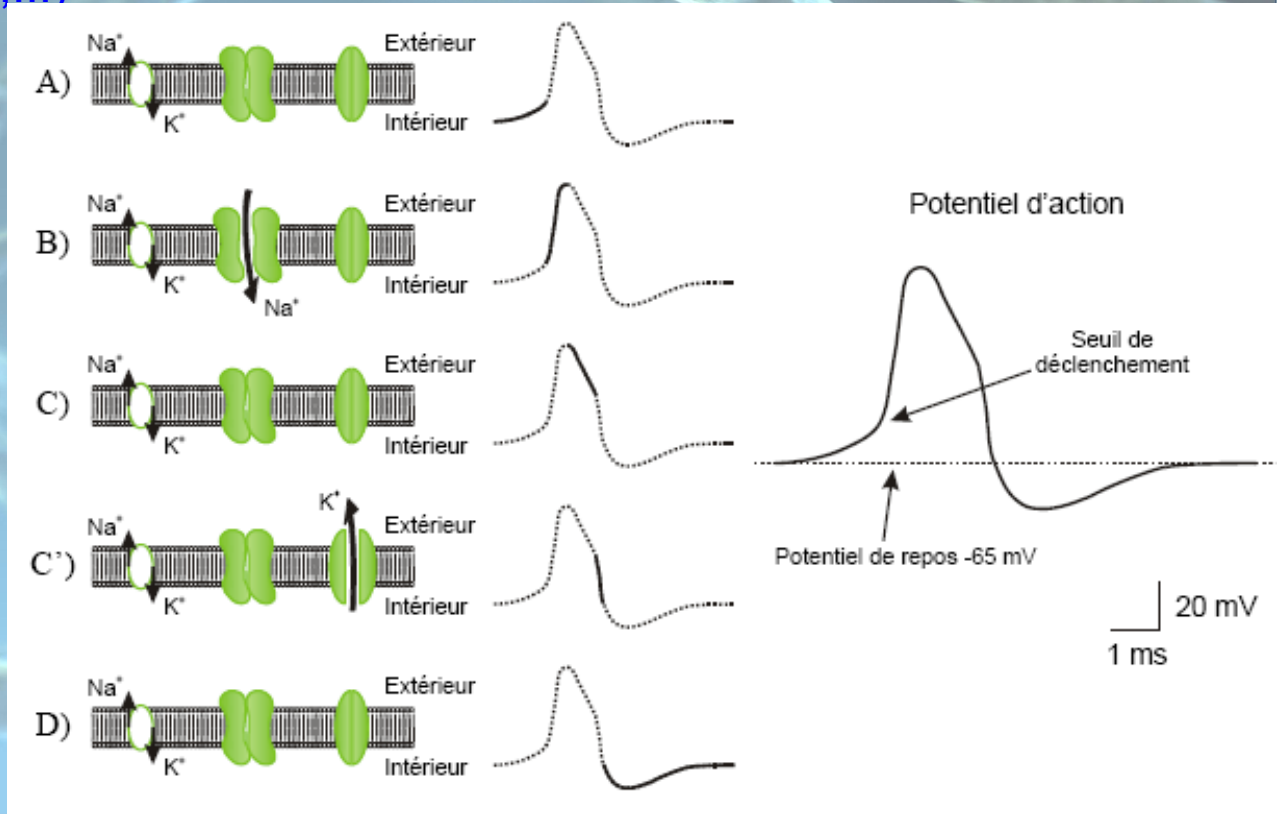
• Since we are talking about ions there is a potential difference (voltage) between both means. Under steady state this difference is about 65mV.



Neural Networks

• After a certain set of external stimulus the voltage difference can rise above a reference value causing the neuron to emit a voltage associated with the ion exchange.

• These spikes also serve to stimulate other neurons that can also be activated by voltage sources of other types (biological sensors such as the nervous terminations, eyes, ears,...)

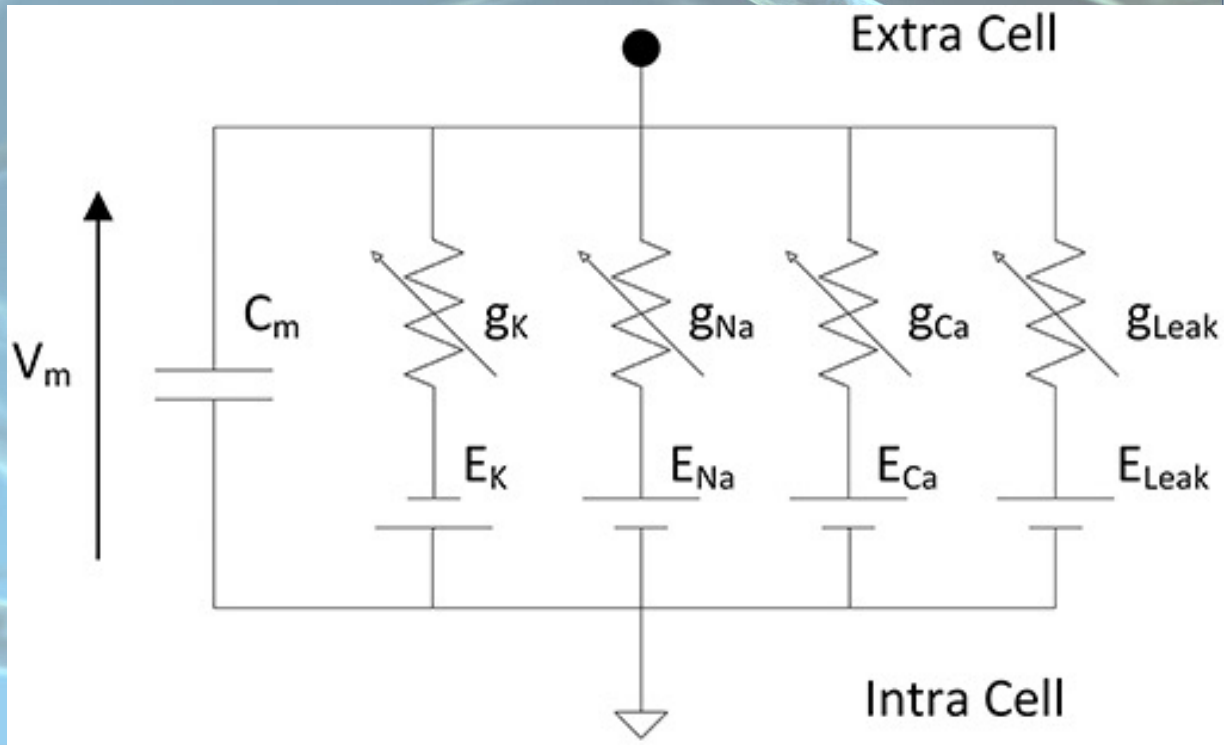


Neural Networks

• The interactions between the intra and extracellular mean are difficult to explore in living being.

• To minimize this problem there are electrical models, such as the Hodgkin and Huxley model that allow testing the neuron properties.

• These models were extracted based on the giant squid neuron that is big enough to test without miniaturization.



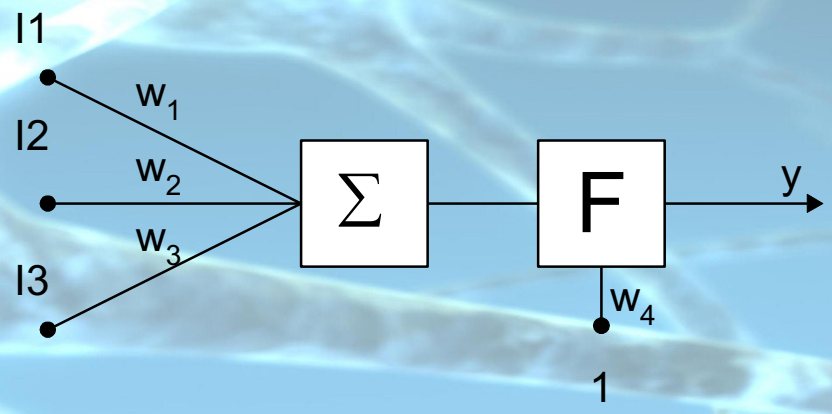
Artificial Neural Networks

• There are many different types of ANN, with different characteristics, all biologically inspired, though some remotely.

• The origin of so many different models is related with the reduced knowledge of the brain functioning.

• Probably the most well known model is the Perceptron or Feedforward Neural Network. This model holds no lateral or feedback connections.

• The neuron performs a very simple function:



$$y = F(\sum_{i=1}^{n} Ii.wi)$$

Artificial Neural Networks

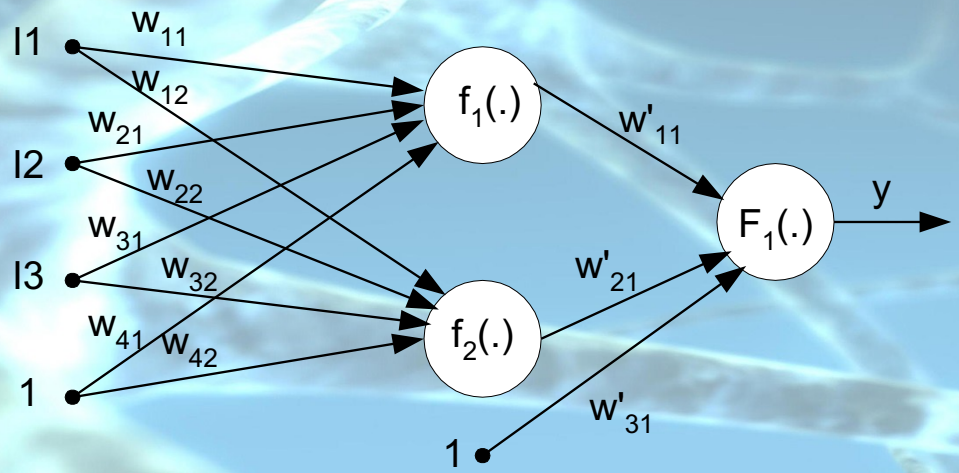
• A network of such neurons implements the following function:

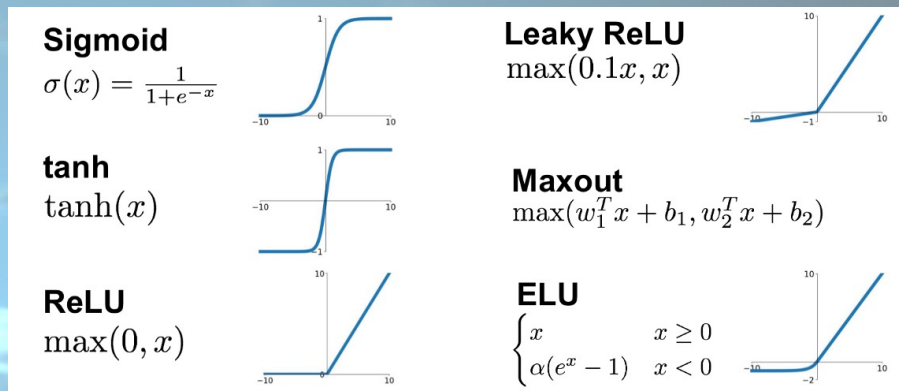
$$y = F_1\left(\sum_{j=1}^{nh} w'_{j1} f_j\left(\sum_{l=1}^{nI} w_{lj} I_l\right)\right)$$

• One of the advantages of this function is the possibility to choose different activation functions, some of which are nonlinear.

• Although one can use ANN with any number of layers, it has been proved that one hidden layer ANN is an universal approximator.

(several layers means more capacity though)





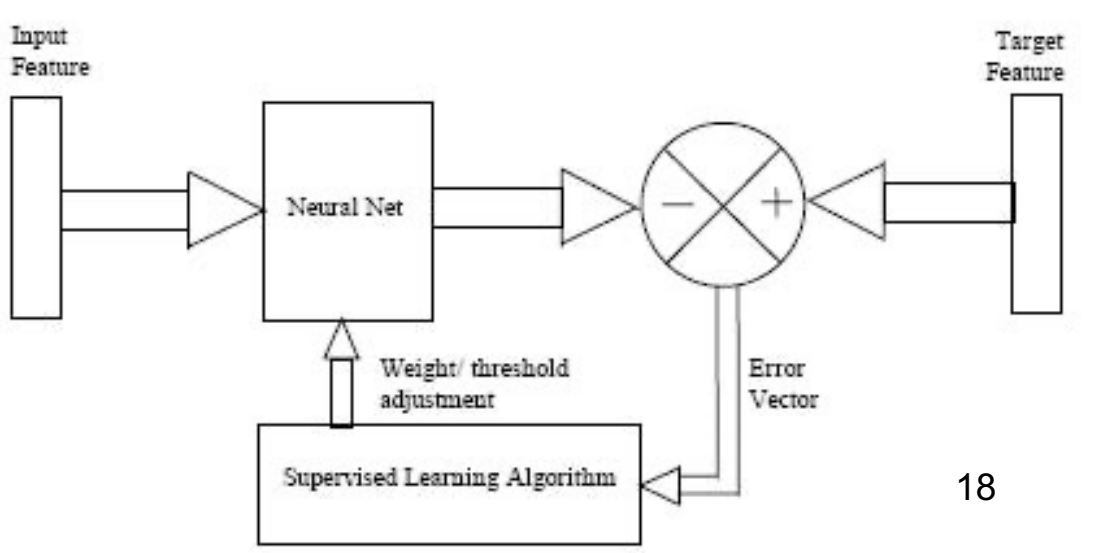
Artificial Neural Networks

- How do they learn?
- Just as with their natural counterpart, they learn through examples that lead to connection (weight) changes.
- Although there are many different algorithms, the common part is to use the output error to adjust the weights throughout the ANN to correct its answer.
- ANN weights are updated in an iterative way. Starting with an initial guess, they are, usually, updated in the following way:

$$w_{k+1} = w_k + \alpha_k . p_k$$

Where pk stands for the search direction and αk represents the learning rate.

Different algorithms choose different ways to select αk and pk.



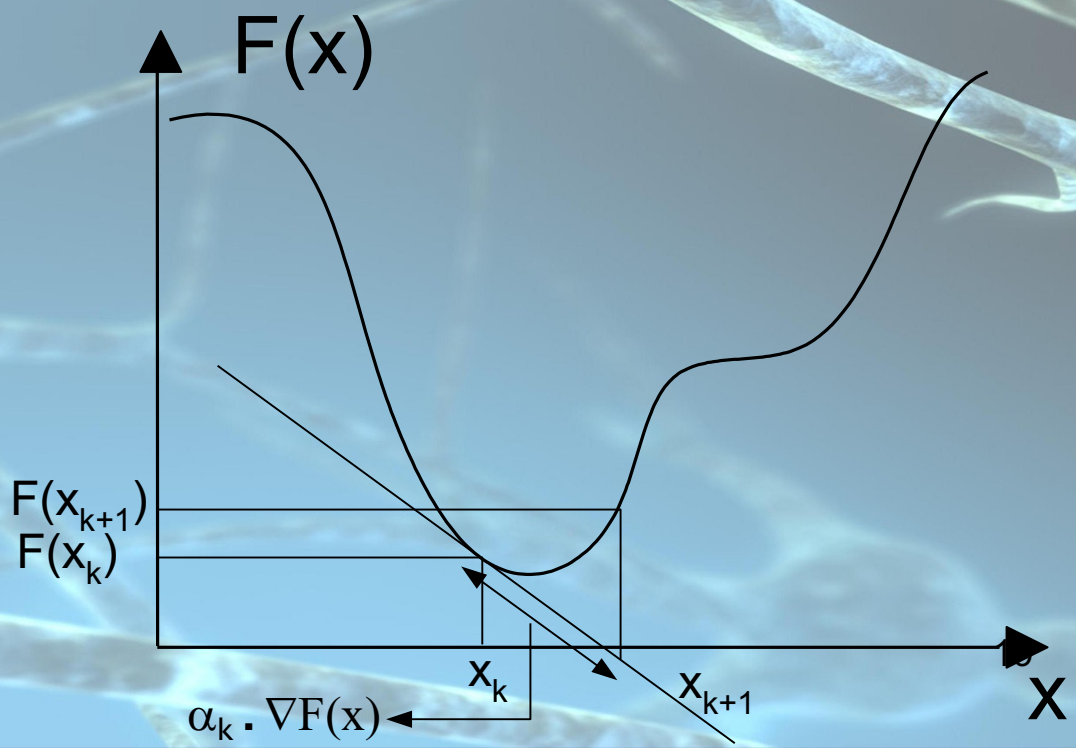
Artificial Neural Networks

• Since the functions to minimize are usually quadratic (squared error), we use derivatives to find the minimum value:

$$w_{k+1} = w_k - \alpha_k . G(x,k)$$

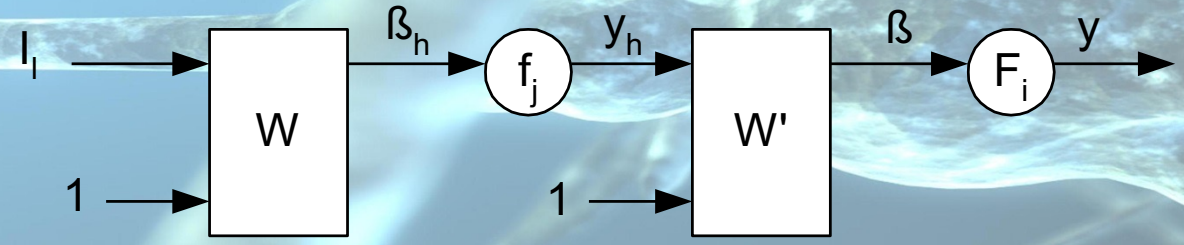
where G(x,k) is the gradient of the function to minimize at iteration k.

• The choice of learning rate is also an important factor:



Artificial Neural Networks

• The previous transparencies explain the principle of updating the weights but not how to distinguish them inside the ANN. This is done through Backpropagation which is based on the derivative chain rule:



$$w(k+1) = w(k) - \alpha . \frac{\partial EQ}{\partial w}$$

$$y = F(\sum_{j=1}^{nn} w'_{jl} . f_j(\sum_{l=1}^{ne} w_{lj} . I_l))$$

$$\frac{\partial EQ}{\partial W'} = \frac{\partial EQ}{\partial y} . \frac{\partial y}{\partial W'} = \frac{\partial EQ}{\partial y} . \frac{\partial y}{\partial \beta} . \frac{\partial \beta}{\partial W'}$$

$$\frac{\partial EQ}{\partial W} = \frac{\partial EQ}{\partial y} . \frac{\partial y}{\partial \beta} . \frac{\partial \beta}{\partial y_h} . \frac{\partial y_h}{\partial \beta_h} . \frac{\partial \beta_h}{\partial W}$$

Artificial Neural Networks

•These algorithms are based on the availability of examples (or a tutor that classifies situations as good or bad). This is similar to human being learning.

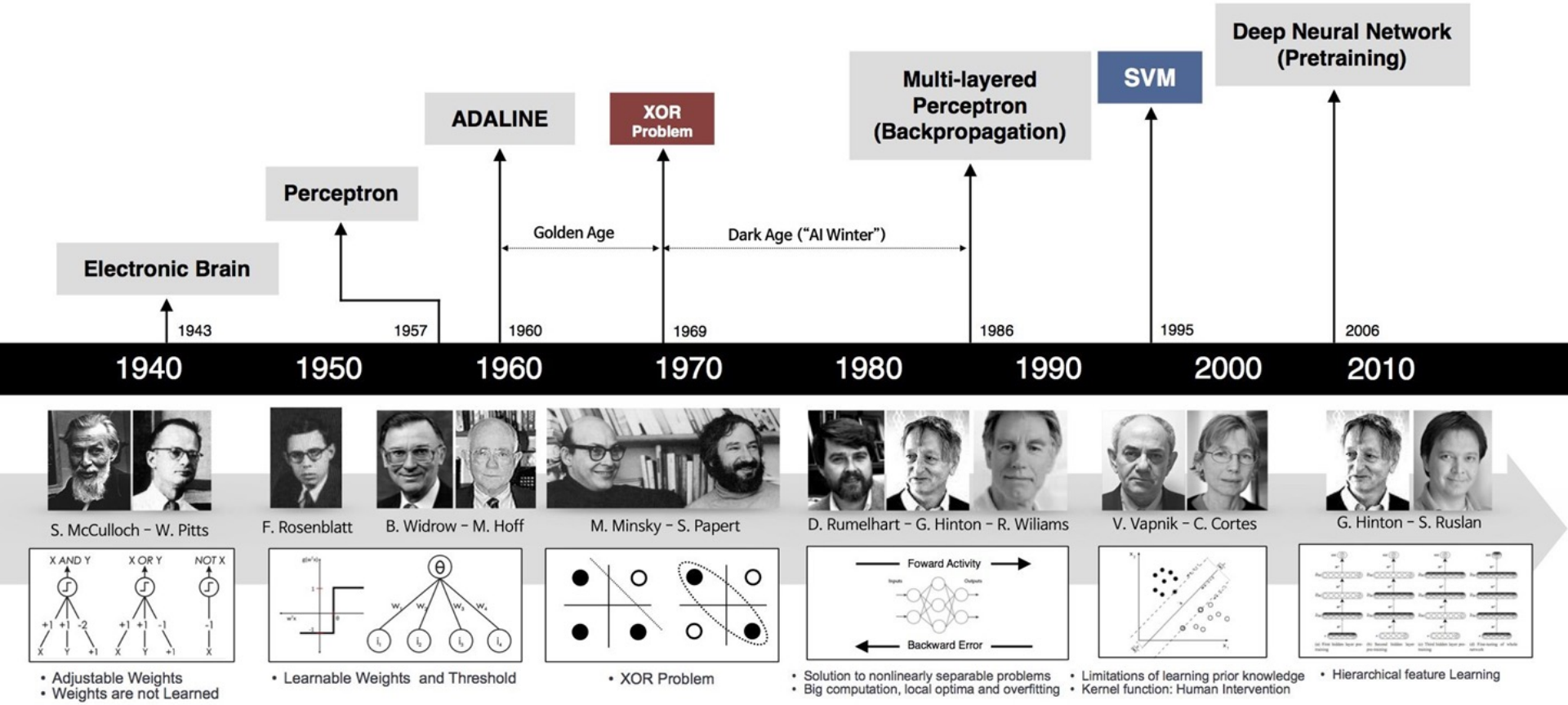
•One of the problems of ANN learning (also similar for human beings) concerns the ability to generalize. When we teach an ANN to behave in a certain way, has it learned to solve that particular problem or can it generalize and answer correctly for similar situations?

•To verify this issue there are specific methods which usually include using different training and test sequences.

•Once the model is built it can be used to represent the behavior of the system, to control it or to do classification (among other tasks).

• The derivative based algorithms were very important but were responsible for making it impossible to use more than 2 hidden layers (the updates of the weights become almost 0). **Deep Learning was only possible after the paradigm was changed.**

ANN history resume



AI and ANN application areas

- AI and ANN are used in many different areas:
- Social Sciences: credit decisions (economy); estimate physical parameters (sport).
- Health: estimate disease risks; medical image analysis (medicine); diagnosis methods (medicine).
- Chemistry: control of chemical reactions.
- Electricity: tools for predicting energy production from wind or solar.

Conclusion

ANNs are used in many different fields of application and in general AI is a growing field with more and more applications every year.

ANNs can be implemented though shallow and deep models and they can be used in supervised, unsupervised and mixed methods.

They became a relevant tool and now, mostly due to the advent of Deep Learning, everyone wants to know about them.

Conclusion

BIESA

biesa.arditi.pt

Relevant publications

1. A Review of Approaches for Sleep Quality Analysis, Mendonça, F., Mostafa, S., Ravelo-García, A., Morgado-Dias, F. and Penzel, T., IEEE Access 7: 24527-24546, 2019. (IF=3.557, Q1)

2. “A Review of Obstructive Sleep Apnea Detection Approaches”, Fábio Mendonça, Sheikh Shanawaz Mostafa, Fernando Morgado-Dias, Antonio G. Ravelo-García, Thomas Penzel, IEEE Journal of Biomedical and Health Informatics, 2019, (IF=3.85, Q1)

3. “Sleep quality estimation by cardiopulmonary coupling analysis”, Fábio Mendonça ; Sheikh Shanawaz Mostafa ; Fernando Morgado-Dias ; Antonio G. Ravelo-García, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2019, (IF=3.972, Q1)

4. "Machine Learning PV System Performance Analyser”, Sandy Rodrigues, Helena Ramos, F. Morgado-Dias, Progress in Photovoltaics, 2018 (IF=6.456, Q1).

5. “Devices for Home Detection of Obstructive Sleep Apnea: A Review”, Fábio Mendonça, Sheikh Shanawaz Mostafa, Fernando Morgado-Dias, Antonio G. Ravelo-García, Thomas Penzel, Sleep Medicine Reviews Volume 41, pp. 149-160, 2018, (IF: 10.602 , Q1).

6. “Economic analysis of photovoltaic systems for the residential market under China's new regulation”, Sandy Rodrigues, Xiaoju Chen, F. Morgado-Dias, Energy Policy, Vol.101, 2017,467-472 (IF=4.039, Q2).

7. “Automatic general-purpose neural hardware generator”, F. D. Baptista and F. Morgado-Dias, Neural Computing and Applications, 2017 (IF=4.213, Q1).

8. “NanEye-An Endoscopy Sensor with 3D Image Synchronization”, Ricardo M. Sousa, Martin Wäny, Pedro Santos, Morgado-Dias, IEEE Sensors, 2016, (IF=2.512, Q1).

9. “Economic feasibility analysis of small scale PV systems in different countries”, Sandy Rodrigues, Roham Torabikalaki, Fábio Faria, Nuno Cafôfo, Xiaoju Chen, Ashkan Ramezani Ivaki, Herlander Mata-Lima and F. Morgado-Dias, Solar Energy, 2016, (IF=4.018, Q2).

10.“Artificial Neural Networks: a Review of Commercial Hardware”, Morgado Dias, A. Antunes, A. Mota, Engineering Applications of Artificial Intelligence, IFAC, Vol. 17/8, pp. 945-952, 2004, (IF=0.641, Q3).

Bibliography

Redes neuronais Artificiais, do algorítmo à implementação, Morgado Dias, Tese de Doutoramento, Universidade de Aveiro, 2005.

Circuits et Systemes de Modelisation Analogique de Reseaux de Neurones Biologiques: Application au Developpement D'Outils por les Neurosciences Computationnelles, Sylvain SAÏGHI, Université Bordeaux I, École Doctorale des Sciences Physiques et de L'Ingenieur, 2004.

Hodgkin AL, Huxley AF (April 1952). "The components of membrane conductance in the giant axon of Loligo". The Journal of Physiology. 116 (4): 473–96. doi:10.1113/jphysiol.1952.sp004718. PMC 1392209 Freely accessible. PMID 14946714.

Adaptation in Natural and Artificial Systems, John Holland, 1975.