Machine Learning I

Professor: Petros Koumoutsakos

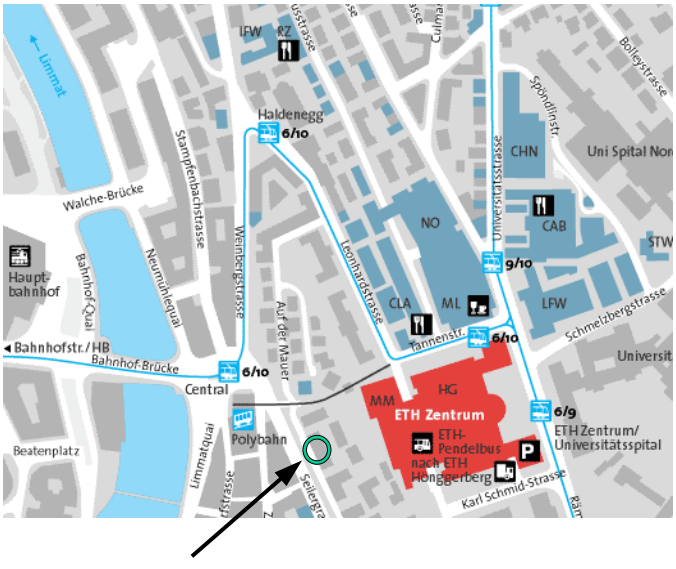
Dozenten: Nicol Schraudolph
Sibylle Müller

WEB PAGE: www.icos.ethz.ch -> Teaching -> Machine Learning

LECTURES: Thursday, 13:15 - 15:00 Uhr – **IFW A36**

EXERCISES: Monday, 15:15 - 16:00 Uhr – **IFW C42**

What (and where) is HRS?



Machine Learning I

Textbooks:

- **Neural Networks for pattern recognition** - C. Bishop
  Oxford U. Press (used in Machine Learning II)

- **Learning from Data** - Cherkassky and Mulier
  J. Wiley and Sons (expensive)

Grades (if needed) :

- **final oral examination** (15 mins)

  if oral exam grade is on the border your **homework** performance will be taken into account

Homework and Exercises

- **Homework**: handed out in class
  or posted on the web site by Friday noon
- bring it to the exercise on Monday
- at least **60%** of homework required for the **testat**

- content of Monday exercises:
  - review of previous lecture
  - review of homework
  - first month: **tutorials** on probability theory, differentiation, and linear algebra

	Exercise		Lecture	
	Mon.	Content	**Thu.**	Content
1		(no exercise)	24.10.	- organisational matters - Machine Learning: **Introduction and History** (K,S)
	Homework: Perceptron learning rule			
2	28.10.	- homework review - **Tutorial** (M): probability theory	31.10.	**Bayes Rule** (S): - central role in statistics - derivation and formula - use for machine inference
	Homework: ovarian cancer screening OR whale sound filtering			
3	04.11.	- homework review - **Tutorial** (M): differentiation	07.11.	**Foundations** (S): - modeling of data - maximum likelihood (ML) - least-squares (LS) regression as an ML procedure

Reading: chapters 1-5 of
Conjugate Gradient Without the Pain

4	11.11.	**Tutorial** (M): differentiation, linear algebra	14.11.	**Optimisation** (S): - simple gradient descent (GD) - basis functions, generalized LS
5	18.11.	**Tutorial** (M): linear algebra	21.11.	**Optimisation** (S): - weakness of simple GD - Newton's method - singular value decomposition
6	25.11.	lecture review	28.11.	**Regularisation** (M): - overfitting - cross-validation - penalisation

7	02.12.	lecture review	05.12.	**Model Selection** (M): - Ockham's razor: learning as compression - structural risk minimization - minimum description length
8	09.12.	lecture review	12.12.	**Neural Networks** (S): - biological background - trainable basis functions - backpropagation algorithm
9	16.12.	**review** (M): backpropagation	19.12.	**Training Methods** (S): - learning rate adaptation - quasi-Newton methods - conjugate gradient

10		(no exercise)	09.01.	**Density Estimation** (S): - Parzen window method - Expectation-Maximisation
11	13.01.	**review** (M): expectation-maximisation	16.01.	**Classification** (K): - Fisher's linear discriminants - k-nearest neighbor - vector quantisation
12	20.01.	lecture review	23.01.	**Dimensionality Reduction** (K): - curse of dimensionality - principal components analysis - nonlinear autoencoding
13	27.01.	lecture review	30.01.	**Self-Organising Maps** (K)
14	03.02.	lecture review	06.02.	**Summary lecture**

Machine Learning : A definition

Given a collection of data ***originating from some functional dependency***

infer this dependency

Machine Learning : A definition

A first cut : Various distribution functions describe well many events of reality.
Examples : Gauss distribution (e.g. usually grades in an exam) - Poisson distribution
(e.g. distribution of bacteria in water).

A second cut : Create descriptions based on an observational framework

How would you describe ?

The flight of a paper airplane ?

The sound of a violin ?

The glow of a campfire ?

Surfing a 4m wave ?

Highway traffic in rush hour ?

The pumping of your heart ?

How would you describe :

Breaking Glass ?

QuickTime™ and a Compact Video decompressor are needed to see this picture.

How would you describe :

Raindrops ?

QuickTime™ and a Microsoft Video 1 decompressor are needed to see this picture.

How would you describe :

how ants find food ?

Social insects, following simple, individual rules, accomplish complex colony activities through: flexibility, robustness and self-organization

Pheromone Trail Following

Ants and termites follow pheromone trails

Asymmetric Bridge Experiment

Goss et al., 1989

Dorigo & Bertolissi, 1998

How would you describe :

Moving on a Real Bridge ?

QuickTime™ and a Sorenson Video decompressor are needed to see this picture.

How would you describe :

A Violin and its playing ?

http://library.thinkquest.org/27178/en/section/1/index.html

Describing a Violin Playing

The Observational (MACHINE LEARNING) Way :

- Determine some correlation between actions and sound that comes from the instrument -
- find an effective description

The Numerical Way :

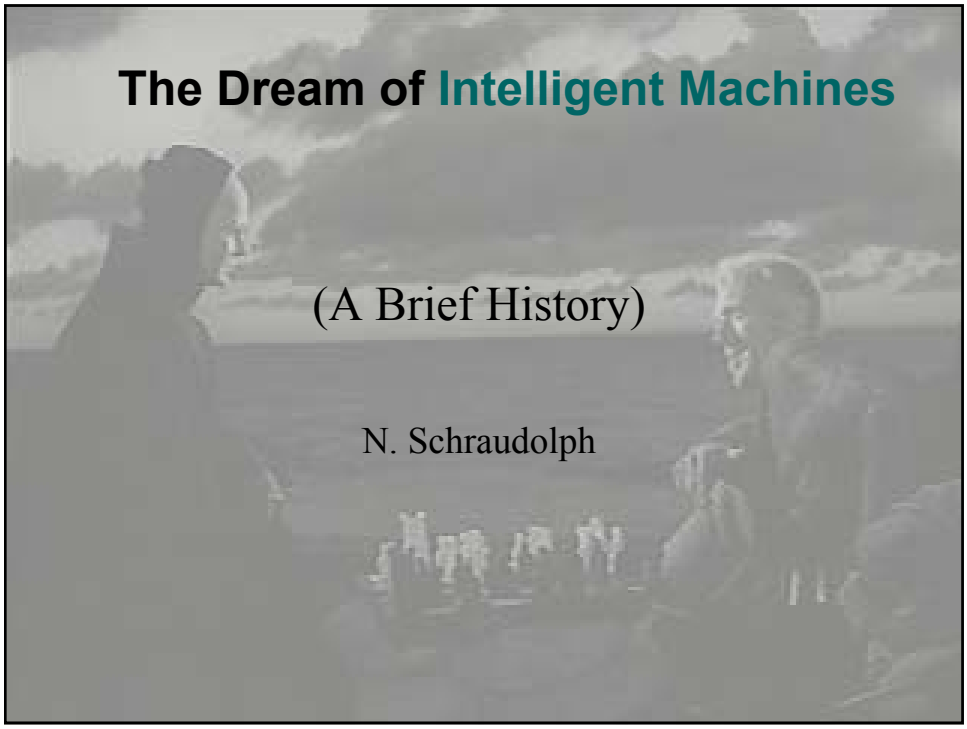
- Determine the Governing Equations from first principles - (This should be good enough to let you distinguish between a Stradivarius and a Guarneri)
- Determine a method for discretizing them
- Solve them numerically - (Use large supercomputers (or MATLAB) to get the desired accuracy)

The Analytical Way (pencil/paper or MAPLE) **:**

- make assumptions
- Simplify the Governing Equations -
- Solve them analytically - (Solution may not sound very good)

Levels of Description for a Problem

Specific	⟷	General
Model	⟷	First principles
Numerical	⟷	Analytical
Molecular	⟷	Macroscopic
Discrete	⟷	Continuous
Qualitative	⟷	Quantitative



The Big Question

What **is** life, intelligence, conciousness? Is it caused by

- divine fire (the **Prometheus** legend)?
  It does look like some kind of magic, but:
  any sufficiently advanced technology is indistinguishable from magic (Arthur C. Clarke). So perhaps it is due to
- some physics we don't understand well yet?
  - electricity? (Mary Shelley's **Frankenstein**, 1817)
  - quantum mechanics? (Penrose, 1989)
  - *emergent complexity of adaptive systems?*

A Pragmatic Approach

We may never know what life/intelligence/consciousness **is**, but we can try to build machines that look (and act) as if they possessed it.

If physics is the answer, we may yet become Prometheus, creating ``real'' artificial life/intelligence/consciousness.

If divine fire is required, we may at least create simulacra that are ``good enough'' for many practical purposes.

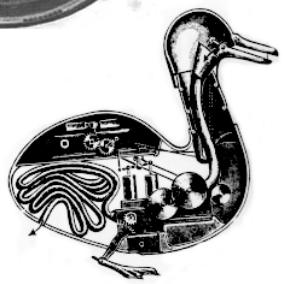
Through the ages, this program was pursued using the most advanced technology of the time. Before the advent of electronic computing (and with it, software) that meant building clever hardware: mechanical **automata**.

Early Automata

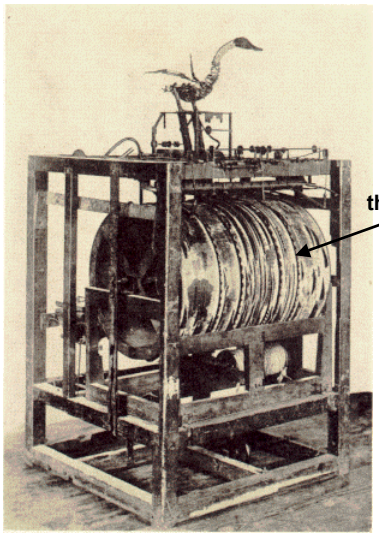
- **400 BC** Philosopher and mathematician Archytas of Tarentum built a wooden dove that could flap its wings and fly.
- **Early 16th Century** Hans Bullmann creates the first androids - simulated people that can play musical instruments for the delight of paying customers.
- **1533** In his laboratory at Nuremburg, scholar Johann Müller, aka Regiomontanus, is reputed to have created an iron fly and an artificial eagle, both of which could take to the air.
- **1543** In England, John Dee creates a wooden beetle that can fly for an undergraduate production of Aristophanes' Pax.
- **1725** At the Heilbrunn chateau in Germany, a mechanical theatre is created featuring 119 animated figures that perform a play about village life to the accompaniment of a water-powered organ.

Jacques de Vaucanson

- Born **1709** in Grenoble
- While training as a Jesuit, builds flying angels which cause him to be thrown out of the order.
- **1737** Vaucanson creates a mechanical flute player that can play 11 different tunes. He also creates an automatic duck that can drink, eat, paddle in water, digest and excrete like a real duck. Each wing reputedly contains more than 400 moving parts.
- ''*a rival to Prometheus*'' - Voltaire



Remains of Vaucanson's Duck



Jaquet-Droz

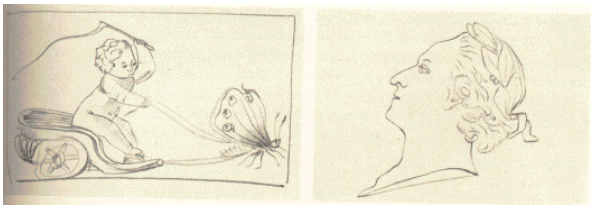
- Pierre Jaquet-Droz (1721-1790), was born in Neuchâtel (Switzerland) and became an engineering clockmaker. He was interested very early by applied mechanics. His instruction and his intelligence allowed him to be one of the largest mechanists of his time.
- His son Henri-Louis (1752-1791) became his collaborator at a young age and brought his taste for the arts, music in particular. Among his apprentices was Jean Frederic Leschot (1746-1824) who was his right arm. He took over the management of the house after the death of his two leaders.

Jaquet-Droz Automata (1774)

The Draughtsman

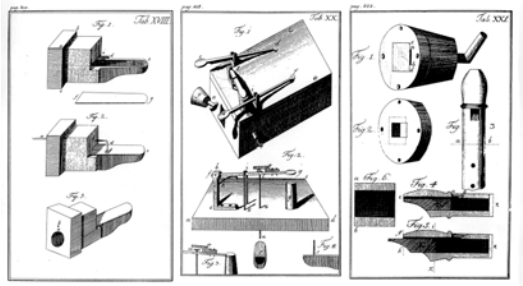
- 3 famous automata: the writer, the musician, and the draughtsman
- Presented to Swiss high society, to Louis XV in Paris, in Brussels, London, Russia, Madrid, etc...
- Now in the Museum of Art and History, Neuchatel, Switzerland.

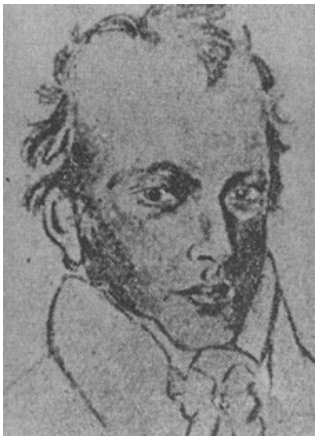
sample drawings by the draughtsman:



Baron Wolfgang von Kempelen

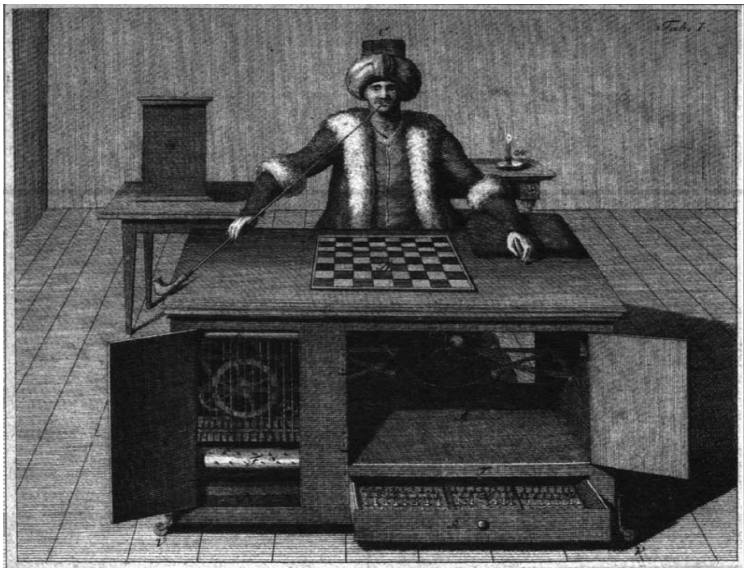
- born 1734 in Bratislava
- scientist and inventor at the court of Maria Theresa in Vienna
- constructed the world's first speech synthesizer (today in the Deutsches Museum, Munich)





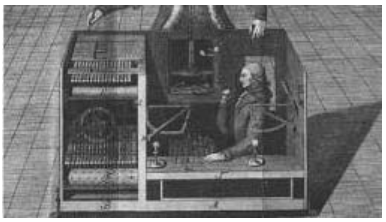
But infamous for…

von Kempelen's Chess Turk (1769)

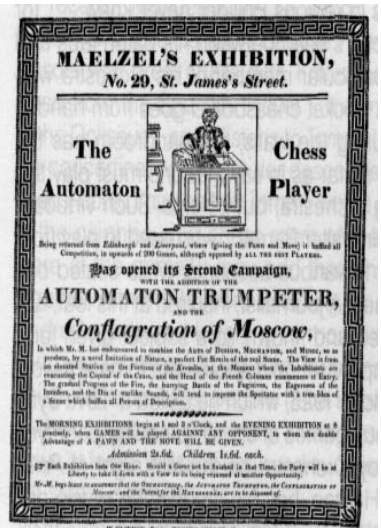


The Chess Turk

- von Kempelen called it "a bold illusion"; it was of course a hoax!



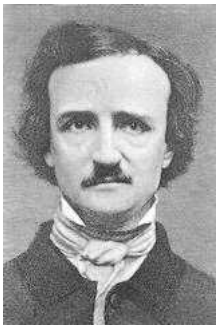
- nonetheless extremely influential in establishing the idea of intelligent machines
- toured the world for almost 70 years; destroyed by fire in 1854



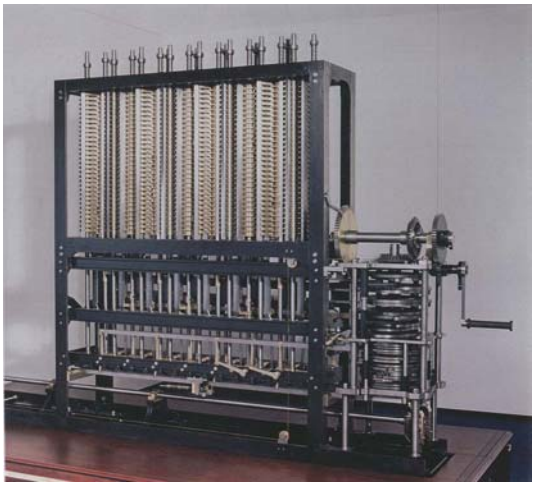
Famous Chess Turk Opponents

- **Frederick the Great** lost against it, as did **Benjamin Franklin**
- **Napoleon Bonaparte** tried to cheat against it, but lost as well

- his stepson was so desperate to find out how the Turk worked, he bought it for 30'000 francs
- **Edgar Allan Poe** wrote an investigative essay on it
- last but not least: **Charles Babbage** played against it, then went on to design his famous compute engines



Charles Babbage (1791 – 1871)



the Difference Engine (modern reconstruction)

The Advent of Computing

The advent of electronic computers brough two important changes to the quest for machine intelligence:

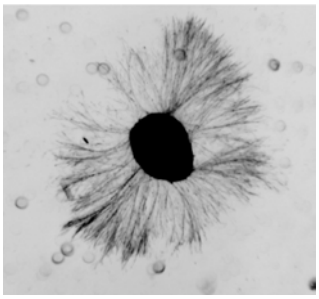
- it could be decoupled from special-purpose, limited performance, mechanical or electrical **hardware**
- it brought a growing realization that our **brains** are in fact computing devices

The focus thus shifted to trying to design intelligent **software**.

The science of **cybernetics** began to systematically examine two-way coupling between an (analog) computer and its environment through **adaptive feedback loops**.

The Perceptron

- **1943**: McCulloch & Pitts propose the first computational abstraction for what biological **neurons** might be doing
- **1958**: Rosenblatt proposes the first training procedure for the McCulloch-Pitts **perceptron**.



- **1969**: Minsky & Papert show that simple perceptrons can only learn simple **(linearly separable)** problems. Symbolic AI wins the battle for funding over machine learning. Rosenblatt, discredited and depressed, dies in a "boating accident".

Neural Networks

- **1980s**: a training procedure for **multi-layer perceptrons**, that overcome the Minsky/Papert limitation, is found independently by Parker, Widrow, Rumelhart/Hinton/Williams
- meanwhile, symbolic AI is stagnating. **Neural networks**, and with them machine learning in general, take over as the leading paradigm for the Promethean quest.
- **today**: machine learning is an integral part of computer science, with applications in many aspects of everyday life.
- artificial life/intelligence/consciousness, however, remains elusive. Perhaps it's quantum mechanics after all?