Cognitive Systems: Neural Networks and Applied AI (IN2222)

# LECTURES

# Introduction

## 🌸Related topics

|  |  |
| --- | --- |
| **AI:** How to develop a rational agent? | * Narrow problems with an explicit description (e.g., path planning, inference) * Often identified with symbolic reasoning * Often used as an umbrella term for many other methods |
| **ML**: How to fit a model to data? | * Focus on problems with limited information available (e.g., weather forecast, stock market) * Often related to AI in the context of deep learning * Strongly related to mathematical optimization |
| **Cognitive Systems (CoSys)**: How to develop a system which can autonomously act and interact in an environmental niche? | * Defining and characterizing cognition * How it is implemented in the brain of living creatures * Role of the body and the environment * Which cognitive functions required for certain tasks * Which behaviors learned, and which hardwired * How single functions can be integrated to form a coherent system |

# Part 1: Foundations and theory of CoSys

*Introduction to Cognition and Intelligence, Embodiment, Simulation, Neurorobotics*

## 🌸A brief history of cognitive systems

* Motivation of building CoSys?
* First CoSys ever build?
* Robotics and CoSys?
* What will the future bring?

🐇 Cognition: “ability to comprehend”, from Latin cognitionem (nominative cognition) “a getting to know, acquaintance, knowledge”, “to get to know, recognize”. There is no unified definition of what cognition is and what it means.

🐇 Cognition II: the process involved in knowing, or the act of knowing, which in its completeness includes perception and judgement. Cognition includes all processes of consciousness by which knowledge is accumulated, such as perceiving, recognizing, conceiving, and reasoning. Put differently, cognition is an experience of knowing that can be distinguished from an experience of feeling or willing. It is one of the only words that refers to the brain as well as the mind.

* Cognition enables a system to operate in a meaningful way beyond its original pre-programmed behaviour and specification
* Example applications that require cognition: autonomous driving, NLP, human robot interaction, physical interaction, collaborative work with multiple agents

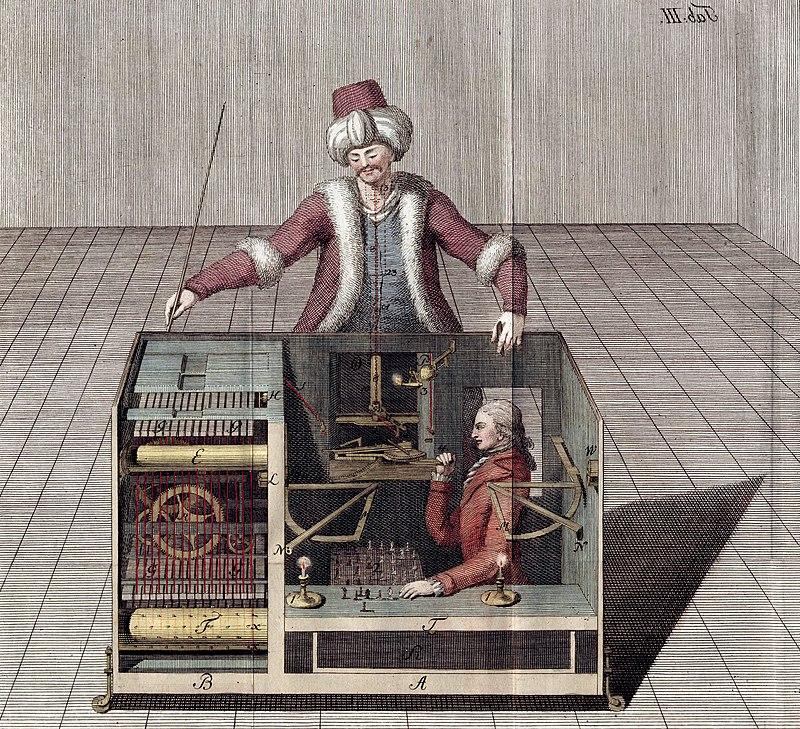
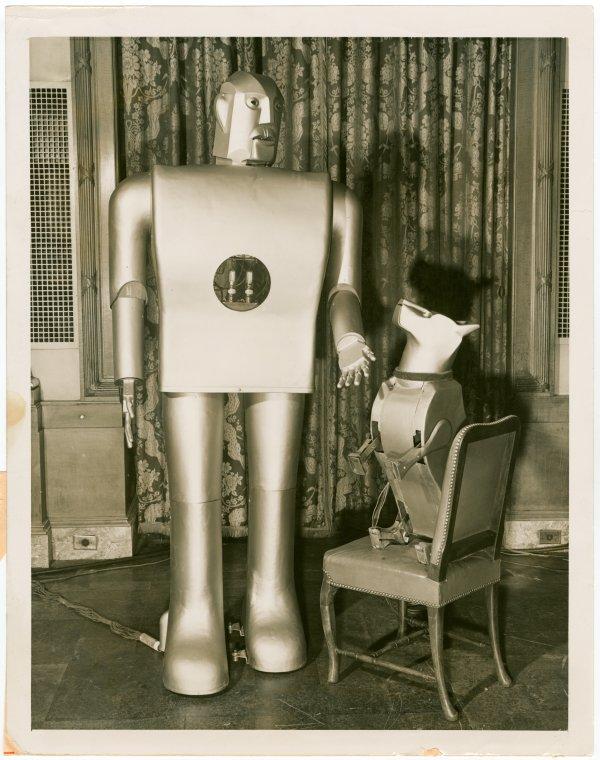
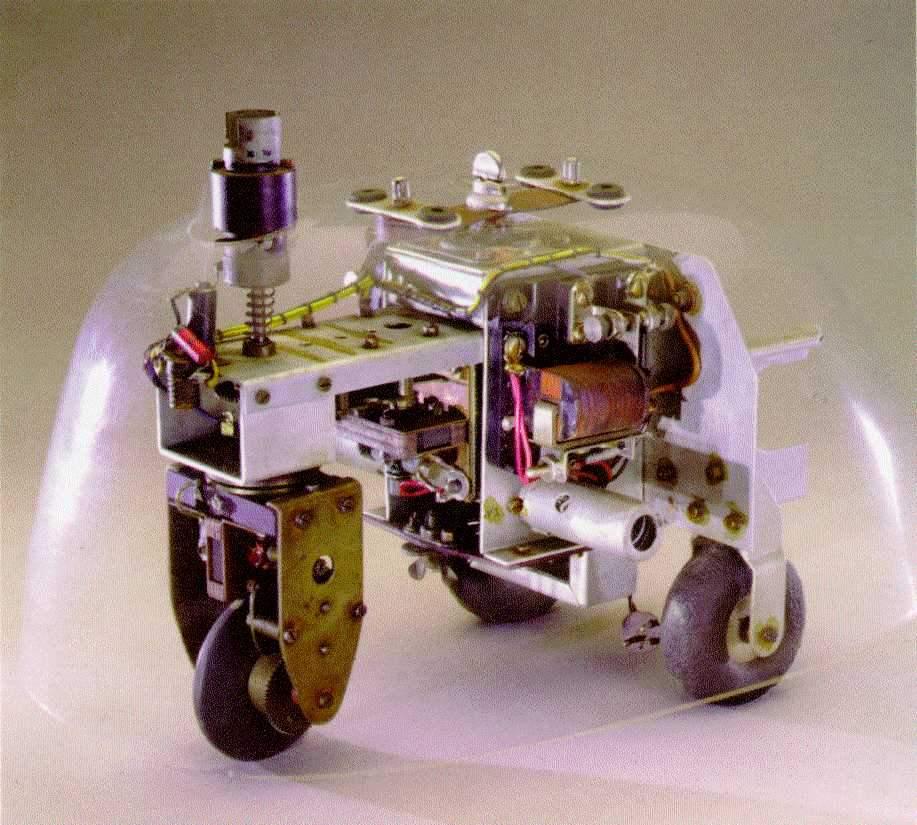
🐇 System: “the whole creation, the universe”, from Latin systema “an arrangement, system”, from Greek systems “organized whole, a whole compounded of parts”, “to place together, organize”.

* A group of interacting bodies under the influence of related forces
* The body considered as a functional unit
* A group of devices or artificial objects or an organization forming a network especially for distributing something or serving a common purpose

🐇 Artificial intelligence (AI): draws from mathematics, logic, control theory, decision theory and etc. It has a goal of duplicating human faculties - language, learning, memory, rational thinking, creativity. AI is a tool, not a threat.

## 🌸Trying to build human-like machine

1. Year 62: Heron of Alexandria (Greek) builds *programmable automata* that are used in *puppet theaters*
2. 1136-1206: Ismail al-Jazari (Muslim) develops different *humanoid automata* that *dispense water, flush a basin, or play music*
3. 1495-1497: Leonardo da Vinci (Italian) draws construction plans for a *mechanical knight*
4. 1770: Wolfgang von Kempelen (Hungarian) constructs the „*Mechanical Turk*” a human-operated check-playing fake automaton
5. 1837: Charles Babbage (English) designed Analytical Engine – *first Turing-complete computer*
6. 1921: The *term robot is invented* in the play R.U.R by the check writer Karel Capek (Check)
7. 1935-1938: The computer Z1, Z2 and Z3 constructed by Konrad Zuse (German) which have *floating point arithmetic*
8. 1939: *Elektro* robot introduced in the New York World’s fair. He can speak, walk, move its head, recognize colours through a photocell. 26 pre-programmed actions can be triggered by speech.
9. 1940: *Sparko* – robot dog – developed by Westinghouse, commanded by Elektro.

Mechanical Knight, „Mechanical Turk”, Elektro & Sparko, Elmer and Elsie

1. 1942: *Three laws of robotics* are introduced in Isaak Asimo’s (Russian) story “Run-around”, which later becomes a part of collection I, Robot
2. 1946: *ENIAC is Turing-complete programmable* electronic computer with 18,000 vacuum tubes
3. 1848: Foundation on Cybernetics through Norbert Wiener (American) – theoretical foundation for scientific *study of control and communication* in the animal and the machine.
4. *First autonomous robot* constructed by neurophysiologist William Grey Walter (American) to study the emergence of *complex behaviors* in a simple control system
5. 1950: “An imitation of life” by William Grey Walter presents *Elmer and Elsie*, two machines of his artificial robot species. They implement phototaxis and automatically return to their charging station before batteries are empty. They have “sense organs” for light and touch, a nervous system implemented by two radio tubes, two motors for crawling and steering, the direction of the photo system is controlled by the steering motor. Brain of the robots is based on an analogue electrical circle (CORA, conditioned analog reflex).
6. 1951: Bioinspired robots with “brains” *Squeee* developed by Edmund C. Berkeley – collects nuts and carries them to the nest. Sense organs: two phototubes, acting organs: motors, small brain: relays.
7. 1955: Arthur Samuel (American) implements *ML-based program for playing checkers*. The System improves by playing against itself and then wins against the human in 1962.
8. 1956 – [Cognitivism](#_Paradigm_of_cognitive) started as sister discipline of AI during summer school held at Dartmouth College.
9. 1971: *Intel 4004* – the first commercially available microprocessor
10. 1978 - 1973: Blocks world - AI system *SHRDLU* - interaction with a world composed of simple geometric objects *through natural language*. *Memory* allows the system to preserve a limited understanding of context, *physical simulation* enables the system to identify impossible requests (e.g. blocks can not be positioned on a cone).
11. 1961: First *industrial robot* at General Motor.
12. 1966-1972: *Mobile robot Shakey developed at SRI*, for the first-time combining *computer vision*, *language processing* and the *planning systems STRIPS* to navigate in a simplified world with blocks and ramps.
13. 1975: *Mycin* - expert system to *identify infectious bacteria* and *recommending antibiotics + patient-dependent dosage*. Based on around *600 rules*, provided a sequence of textual queries to the physician, incorp a notion of uncertainty through so-called certain factors.
14. 1980: *Rule-based programming language OPS5 (approx. 2500 rules)*, implementing R1/XCON systems capable of automatically configuring VAX computer systems by DEC, based on customer requirements. OPS5 processes 80 000 orders by 1986 with up to 95% accuracy, saved around $25 million per year.
    1. *VAX is a series of computers featuring a* [*32-bit*](https://en.wikipedia.org/wiki/32-bit_computing) *instruction set architecture (ISA) and virtual memory* that was developed and sold by Digital Equipment Corporation (DEC).
15. 1983: *Delta* was an expert system for *trouble-shooting and repairing diesel electric locomotives*. About *530 rules*, represented knowledge of senior knowledge field engineer, provided text info, CAD diagrams and repair sequences.
16. 1985: *DEX.C3* was an expert system for *error diagnosis in the C3 gearbox from Ford*. Implemented in Germany on XEROX 1108 computer. *140 rules,* used *certainty factors*, *explain to the user* why it made a certain decision.
17. 2011: *Watson* wins in U.S quiz show Jeopardy against two human competitors.
18. 2016: *AlphaGo from DeepMind* beats go champion LeeSedol.

**However, CoSys build so far are:**

→ highly specialized for specific niches

→ living creatures are more versatile and flexible

→ design of CoSys = application as well as extension of what is known about cognition

## 🌸Early applications of AI

*Logical Theorist* was designed to solve general problems defined in logic. Targeted tasks with very narrow scope - known as *microwords*, examples: mathematical proofs, strategic games (chess, checkers), simple geometric problems (“blocks world”).

# Part 2: Neural computation

*Computational Neuroscience, Biological Neural Networks, Brain Simulators, AI, ML, (Deep) Artificial Neural Networks*

## 🌸Questions

* What is cognition and what is it good for?
* How is cognition related to intelligence and learning?
* What are the main components of a cognitive system?
* What are cognitive systems capable of?
* Is it possible to measure cognition?

### Key literature:

Graphical user interface

Description automatically generated with medium confidence

Graphical user interface

Description automatically generated

## 🌸The robotic paradox

* Robots are good in the factory but fail to make simple tasks in real-life environment.
* Because there is a big difference between well-defined tasks (welding task at the factory) and ill-defined tasks (shopping task)

🐇 Cognitive System is an autonomous system that can perceive its environment, learn from experience, anticipate the outcome of events, act to pursue goals and adapt to changing circumstances.

## 🌸Cognition, Intelligence, Learning

* Cognition is a global process at a system level
  + Special cognitive skills are intelligence, learning, memory, interaction etc.; these skills are synergies of a CoSys.
* Theory of CoSys creates a conceptual framework which accommodates many other disciplines: AI, ML, computer vision, control theory.
* Individual aspects of cognitive system: cognitive function, cognitive skills

Diagram

Description automatically generated

## 🌸 The cycle of cognitive processing: Anticipation, Assimilation, Adaption, Cognition, Action, Perception

Diagram

Description automatically generated

Diagram

Description automatically generated

Cognitive capabilities:

* Self-reliance: CoSys must be able to act in and interact with its environment self-reliantly, purposefully and independently. The CoSys can act self-reliantly and independently and therefore do not need complete formal specification to achieve a goal.
  + Goal-directedness: take on goals, formulate predictive strategies to get there and put those strategies into effect
  + Autonomy: operate with varying degree of autonomy
  + Interaction: interact – cooperate, collaborate, communicate – with other agents
* Perception & Action: CoSys must be able to perceive its environment, make sense of its perceptions and predict future events
  + Interpretation
  + Sensing
  + Anticipation
  + Action
* Adaption: CoSys must be able to adapt to changes of itself, others and within the environment
  + Reaction
  + Learning from experience
  + Anomaly detection

## 🌸 How to benchmark cognition?

### Turing test

Proposed by Alan Turing in 1950 is aiming to test if the computer can think. It is mased on imitation game, in which interrogator interacts with a player A and a player B by writing notes. One of the payers is replaced by computer and interrogator should guess if the player is a computer or not.

Interestingly, Eugen Goostman, who is in reality a computer program, fooled 10 out 30 judges into thinking that he is a real person.

### Chinese room argument

Chinese room argument suggests that even if someone behaves in intelligent manner, it does not mean that he is intelligent.

### The Uncanny Valley

The physical appearance of a cognitive system has a huge impact on the interaction with humans. The uncanny Valley phenomenon states that a steady increase in “human-likeliness” of a robot does not yield a steady increasy of familiarity to humans.

## 🌸 Guest Lecture: New insights for biological models of cognitive functions

### Cognitive Neuroscience

* How is the mind created by the brain?
* Uses research methods like tomographic brain imaging, electrophysiological methods, single-cell recordings, computer simulations…
* Understanding important insights in neural networks ([class of connectionists system](#_Connectionist_emergent_Systems)) subserving higher cognitive functions, including visual and auditory perception, psychiatric disorders, language, emotion, motor control, memory
* Functional magnetic resonance imaging (fMRI) has gained a leading role in CN research

## 🌸 Modelling Cognitive Processes – tools and methods

### Paradigm of cognitive science (Vernon, 2014)

Diagram

Description automatically generated

Graphical user interface, text, application

Description automatically generated

## 🌸 Cognitivist systems

🐇 Cognitivist systems modeling based on the hypothesis that cognition is a form of computation. Cognitive functions are modelled as working on a computer program.

* Theoretical basis of cognitivism: 🐇 *computational theory of mind (CTM):* mind is a digital computer that stores symbolic representations and manipulates them according to syntactic rules. Thoughts are mental representations – more specifically, symbolic representations in the language of thoughts. Mental processes are causal sequences driven by the syntactic (structure/ form of expression), but not by semantic (meaning), properties of symbols. […]. Cognition can therefore be implemented in a purely physical system.
* Central hypothesis of cognitivism: cognition is achieved by operations on symbols that have a direct correspondence to the real world’s actions and objects.
* External information then gathered and filtered by perceptual processes.
* Over time the focus of AI shifted to purely computational techniques.
* AI with focus on cognition called Artificial General Intelligence.

### Philosophy of Mind

* Challenges cognitivism by asking questions like:
  + What is the basic nature of mental phenomena like thought, feeling, perception, consciousness and sensory experience?
  + Can these phenomena only be described by chemical/ physical processes or can they exist independently of them (inside the PC).
  + How to explain subjective conscious experiences (qualia)?
    - Qualia is often mean phenomenal properties of experience. Qualia is a perceived sensation of pain of a headache, the taste of wine, etc.

### Mind-body problem

* How does the mental world relate to the physical world?
* Two schools of thought: Substance Dualism & Monism
  + 🐇 Substance Dualism (Rene Descartes): mind and body are 2 kinds of different substances; soul is immortal and there is a free will.
  + 🐇 Monism (Physicalism): there is only one substance which means that mental states are physical states. Individuals that share a mental property also share a physical property. Non-reductive physicalism therefore postulates that mental properties are not identical physical ones even through there is only one substance.

### Physical Symbol Systems (PSS)

* From Newell and Simon – theoretical background for the emergence of general intelligence from symbolic systems.
* PSS has physical patterns = symbols, which form expressions that are created, modified, copies, destroyed by process. The system can interpret an expression if it designates a process that can be carried out by the system.
* Problem solving is done by searching.
* Both, symbols, and processes are defined recursively: symbols can designate symbols and processes can produce other processes.

#### Physical symbol system – PSS (Vernon, 2014)

Diagram

Description automatically generated

### Two hypotheses on PSS

1. PSS Hypothesis, Newell & Simon
   1. Physical system has all necessary means and sufficient for general intelligence action
      * Every CogSys is a physical symbol system
2. Heuristic Search Hypothesis, Newell & Simon
   1. PSS is intelligent in problem-solving by search, by generating and progressively modifying symbol structures until it produces a solution structure.
      * Cognition is reflected in an intelligent search strategy that finds a solution without exhaustive brute-force search

### Rationality

* Agent acts in a sensible and purposeful way to achieve its goals. Rational agent always will use as much of its knowledge as possible to guide its behavior.
* Two views on rationality: Optimal rationality & Bounded rationality
  + 🐇 Optimal rationality: ideal agent given its belief-desire system, optimizes its choices -> not feasible in reality.
  + 🐇 Bounded rationality: shown by decision-makers of limited abilities. Rationality depends on success criterion, agent’s prior knowledge of the environment and the actions the agent can perform and past percepts.

## 🌸 Emergent systems

🐇 Emergent systems modeling based on the idea that cognition is a continuous self-organized process that is driven by interaction between agent and its environment (this is how it develops its own models)

* Rise of systems that cannot be predicted or explained from antecedent conditions
* Knowledge and representation are specific to the emerging systems and depend on its history of interaction
* Knowledge is acquired in the form of additional anticipatory skills
* Emergent systems require embodiment and development

### Connectionist Emergent Systems

* Based on networks of simply interconnected computational units
* Cognitive processes here distributed, parallel and based on stat properties instead of formal rules
* Can be [dynamic system](#_Dynamic_Emergent_Systems) (through recurrent connectivity). Connectionist modeling addresses the microscopic level.

#### Main components:

* Units: simple elements
* Equations defining an activation value for each unit at each point in time
* Weighted connections between units: permit activation of one unit influence activity of other units
* Optionally: learning rules

Knowledge is stored in the connection weights. Neural networks are a class of connectionist system.

### Enactive Emergent Systems

* Develops its own understanding of the world around it through interactions
* Governed by two main mechanisms: Phylogeny & Ontogeny
  + 🐇Phylogeny means that interaction is structurally determined by the innate embodies physical capabilities
  + 🐇 Ontogeny means that the system is structurally coupled with its environment and makes sense of it based on its actions, perception of its environment. It generates its own epistemology through sense-making. This epistemology increases its cognitive capacity by capturing regularities of interaction. Gradual increase of cognitive capacity results in development.

### Dynamic Emergent Systems

* Based on time-dependent differential equations which make it possible to compactly represent complex system behavior and application of analysis methods from dynamic systems theory (attractors, stability, bifurcations…).
* They can represent entities by means of system states or trajectories. These representations are context dependent.
* Time here is an active factor of computation
* The synchrony between agent and environment imposes constraints on the processing speed of the system.

#### Mian properties:

* Dissipation: the number of reachable states reduced over time
* Non-equilibrium system: stable function requires and external energy supply
* Non-linearity: complex behavior can emerge from a small set of state parameters
* Collective variables: the system is represented by a small set of state variables

## 🌸 Neurons and Brain

### The Human Brain

A picture containing diagram

Description automatically generated

#### The Vertebrate Nervous System (Нервная система позвоночника)

* Comprised of:

|  |  |
| --- | --- |
| * + the central (CNS): brain + spinal cord – collects and distributes data through the system   + and peripheral neural system (PNS) – transmits signals between sensory organs and the CNS | Diagram  Description automatically generated |

##### The Anatomy of Human **CNS**: Cerebrum (Мозжечок)

* accounts for sensory integration, voluntary motion and higher -level cognition function.
* two hemispheres are connected by corpus callosum.

Diagram

Description automatically generated A picture containing text, book

Description automatically generated

**Cerebral cortex** = grey matter is the folded outer layer of the cerebrum that is mainly composed of cell bodies. The surface of it is subdivided into three anatomical features:

* Gyri – crests (гребни)

Diagram

Description automatically generated with low confidence

(not all the gyros are listed here, there are more)

* Sulcus – трещина / fissure between two sulci
* Lobes

A picture containing diagram

Description automatically generated

White matter – inner part of the cerebrum – is a core of nerve fibres that connect the cortical region.

##### The Anatomy of Human **CNS**: Brainstem (Ствол головного мозга)

* Brainstem connects cerebrum with the spinal cord. Divided into four sections: diencephalon
  + - 1. Epithalamus: hormone production – melatonin
      2. Thalamus: relay and distribution of sensory and motor signals to the different regions of the cerebral cortex
      3. Hypothalamus: control atomic functions – temperature regulation, appetite, behavior, hormone production
      4. Subthalamus

Midbrain: control of eye movement, auditory processing, visual processing.

Pons: transmits sensory and motor signals between the brain and facial region.

Medulla oblongata: connects brain with the spinal cord.

Diagram

Description automatically generated

#### The human peripheral nervous system

* Autonomic nervous system (вегетативная нервная система): responsible for homeostasis (self-regulation) and operates mainly unconsciously. Includes two antagonist subsystems:
  + Sympathetic nervous system: connecting internal organs to the brain, prep for the stress (heartbeat increase, changing blood flow to the muscles)
  + Parasympathetic nervous system sets the body to resting state and increase digestive function

#### Brain Networks

* Brain networks have static and dynamic connections
* Can be analyzed with graph theory, using nodes, degrees, clustering coefficients, hubs etc.
* Related to information integration

### Cognition without brains

* Jellyfish: have diffused neural system
* Sponges have no neurons at all
* In humans, endocrine system – is a regulatory system which works in parallel to the nervous system by secreting hormones.

### Myths about human brain

#### Brain sizes

* The size of the brain does not mirror the level of intelligence.

#### Encephalization quotient

Table

Description automatically generated

### Cognition: The cerebral cortex

Grey matter = cerebral **cortex** plays the highest cognitive function. Cortex is organizes in layers (neocortex) and columns (cortical microcircuit).

Neocortex is a part of cerebral cortex and considered to be center of higher cognitive function.

Functions of different areas of cerebral cortex:

* Frontal lobe: short-term memory, action planning, movement control
* Parietal lobe: somatic sensation, body image
* Occipital lobe: vision
* Temporal lobe: hearing, learning, memory, emotion

Neocortex layers:

* Layer 1: Molecular Layer
* Layer 2: external Granular Cell Layer
* Layer 3: External Pyramidal Cell Layer
* Layer 4: Internal Granule Cell Layer
* Layer 5: Internal Pyramidal Cell Layer
* Layer 6: Multiform Layer – layer with heterogeneous neurons which blinds into the white matter

Cortical microcircuit

* Consist of the most basic building blocks – nanocolumns, which consist of 80-100 neurons. Cortical columns or modules.

#### The cortical homunculus

### Cognition in other brain regions

#### The hippocampus

### Distinctive features of the human brain

#### Sequence learning

### Modeling and simulating the brain

#### The discovery of neurons

#### Some history

### Neuron models

#### Biological Neurons and spiking neural models

#### Analogue neurons

#### Common activation functions

#### Spiking neurons

#### Levels of modeling

#### Whole brain modeling

##### The multilevel organization of the brain

##### Central questions in brain research

# Part 3: Components of cognitive systems

*Vision, Audition, Touch, Multisensory Integration and Data Fusion*

# Part 4: Cognitive architectures

*Classification Criteria, Production Systems, Discussion of Selected Architectures, Introduction to ACT-R*

# EXERCISES

Exercise 01: [Python Cheat sheet](#_Python_Cheat_sheet)

Exercise 01: [Machine Learning](#_Basic_Concepts_in) & Exercise 02: [Machine Learning Part 2](#_Basic_Concepts_in)

Exercise 03: Neurorobotics Platform Introduction

Exercise 04 - Biological Neural Networks

Exercise 05: Learning Part 1 (slides)

Exercise 05: Learning Part 2 (slides)

Exercise 05: Learning (problem sheet)

Exercise 05: Learning (problem sheet w/ solutions)

Exercise 06: Spiking Neural Networks (slides)

E07: Neurorobotics Platform Exercise CPG

# Python Cheat sheet

1. Jupyter notebook: <https://jupyter.org/try-jupyter/lab/>
2. About python language:
   1. mutable objects - lists, instances of custom object classes - can be changed, changes to a mutable objects are visible from all names
   2. immutable - numbers, strings, tuples - objects can not be changed
   3. Conditional branching
      1. there are if-statements
      2. there is no “switch statement”, multiway branches are emulated with dictionaries
   4. Loops
      1. while-loop
      2. for-loop
   5. Functions
      1. def fun(a, b, c = 0, \*arg\*):
   6. Python supports object-oriented programming with classes and multiple inheritance
   7. Modules
      1. import module
      2. import module as mod
   8. Scopes
      1. global, self (local)
      2. names are resolved according to the LEGB rule: local scope, scopes of enclosing functions, global scope, built-in names
3. NumPy Package
   1. module for fast numeric computation

## 🌸 Tasks

* The fibonacci number // TODO
* Execution speed of python //TODO

# Basic Concepts in ML

🐇 Learning : acquisition of new information or knowledge or the process to acquire knowledge or skill by systematic study or by trial and error.

🐇 Machine Learning (ML) : the field of study that gives computer the ability to learn without being explicitly programmed (as dynamics and objects are too complex, e.g. faces, animals; the system itself is subjected to change, e.g. through aging growth; environment is continuously changing). Practical applications of ML: image classification, speech recognition, autonomous driving, recommender systems, threat protection, control systems.

## 🌸 4 Main components of ML dataset

1. *Dataset S*: set of samples, generated by some system or process; can be single data points or pairs input-output values. *Types of datasets*:
   1. Unlabeled data. Most real-world data is unlabeled (emails, images, videos), label assignment is usually a manual task. Natural datasets usually have specific structural features - big data, data mining.
   2. Labeled data. [MNIST database](http://yann.lecun.com/exdb/mnist/) has 70 000 labeled handwritten pixel digits.
   3. Mixed data: U
   4. “Dynamic data”. Many real-world tasks have complex dynamics and unknown goal representation and therefore cant be modeled as pairs of input and desired output.
2. *Model M*: representation of relations of a certain class of input/output that is hypothesized to be able to model the system or process, which generates S.
   1. The model encodes, stores and retrieves the outcome of the learning process.
   2. Optimal model M\*, corresponding to search in a hypothesis space H, is found through adaptation of M by learning algorithms A.

Global minima is almost impossible to find, therefore we are mainly searching for local optima.

1. *Objective function L*: function which encodes the current performance of M (loss - minimizing losses, or reward).
2. *Algorithm A*: learning algorithm that adjusts M based on S and L.

Machine Learning in Artificial CoSys

Learning and Development

…from **biological** learning **mechanisms**

Memory, Knowledge, and Internal Simulation

…**encoding, storage and retrieving of facts**, experiences and actions

Perception

…basic features to **detect** and **categorize** perceptual **stimuli**

Autonomy

…dynamic **adaptation** to changes in the environment

🌸 The general workflow of ML



🐇 Feature engineering : the selection of the right features, which should contain required information for predictions. Selected features are forming feature vector, e.g. picture has v:=(color, edges, location, …)

🐇 ML task: includes an overall goal, in which we *train model M* in a *hypothesis H* using *learning algorithm A* so that *M minimizes loss of* objective function *L* for *dataset S*. Depending on properties of our dataset S we choose hypothesis H and function L.

## 🌸 Types of ML

Amount of information provided during training

**Unsupervised learning**

* Unlabeled data only
* Discovery of structural features in the data set
* Often applied to labeled datasets as a preprocessing tool for initial data analysis

**Semi-supervised learning**

* Mixed (unlabeled + labeled) data
* Require a priori assumption on input data

**Reinforcement learning**

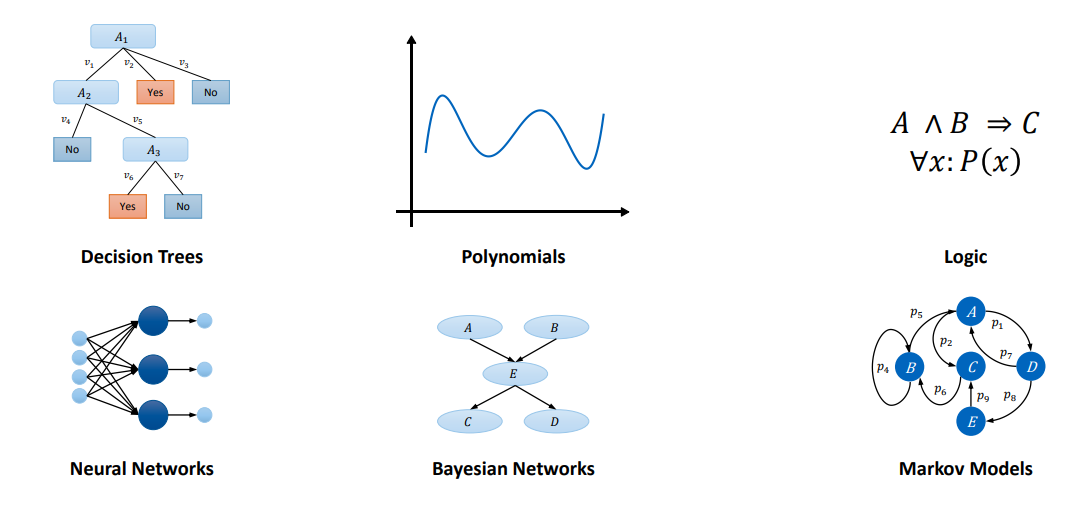
* Interaction with the environment - dynamic datasets
* Reward signal to encode feedback for the policy

**Supervised learning**

* Labeled data only
* Often requires manual assignment of labels
* Desired output specified exactly

## 🌸 Examples of hypothesis spaces

* Different hypothesis spaces support *incorporation of priori knowledge* about the learning task to different degrees (e.g. natural networks or logic)



## 🌸 Ensemble methods and boosting

🐇 Ensemble methods: simple way to extend hypothesis spaces by combining the set of hypothesis h1, h2, …hn to a new hypothesis .

Is it possible to build a strong learner by combining several weaker learners?

🐇 Boosting: the algorithm computes a strong learner by incrementally constructing an ensemble of hypotheses:

* Every training sample is assigned a weight initially all weights are set to the same value
* Weights of incorrectly learned samples are increased
* The training of new hypotheses focuses on samples with high weights

## 🌸 Underfitting, Overfitting and Generalization

H performance depends on how well it predicts the seen data from the training set and unseen data.

* “*Fit*”: the performance of h on the training data (the value for the objective function L for h and S)
* *Generalization*: predictive performance of h on data that were not considered during the training phase.

Issues occurring during the learning:

* *Underfitting*: h fits the training data poorly and does not model the underlying process because H is not expressive enough.
* *Overfitting*: h first the training data very well but does not model the underlying process because it does not generalize well.
  + *Detection of overfitting*: overfitting can be detected by applying the model h to unseen data samples and computing the objective function L. Especially in supervised learning, the data are therefore split into *three different and disjoint subsets*:

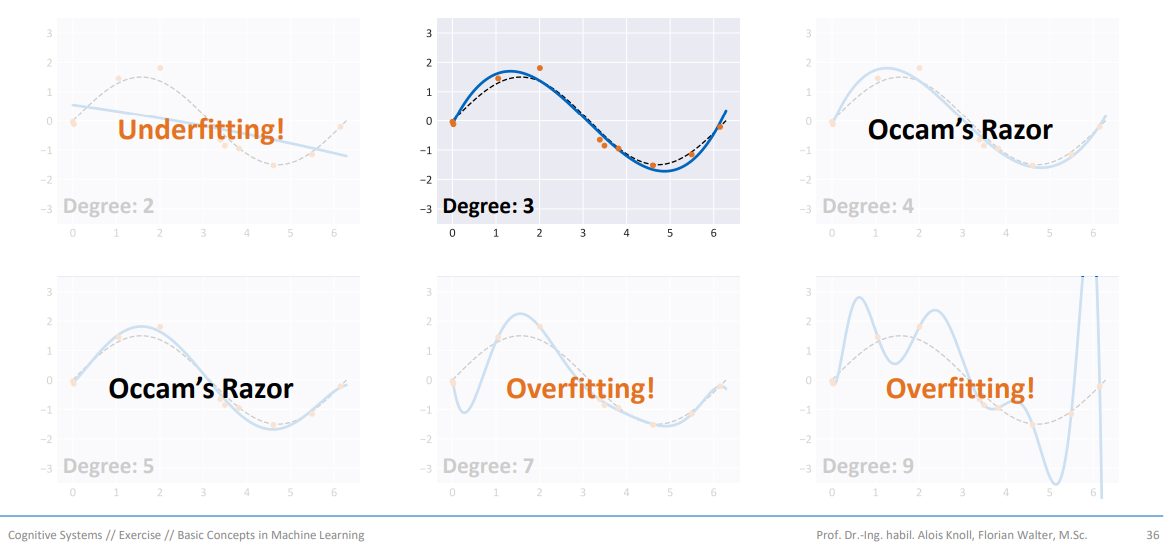
1. Training set
2. Validation set
3. Test set

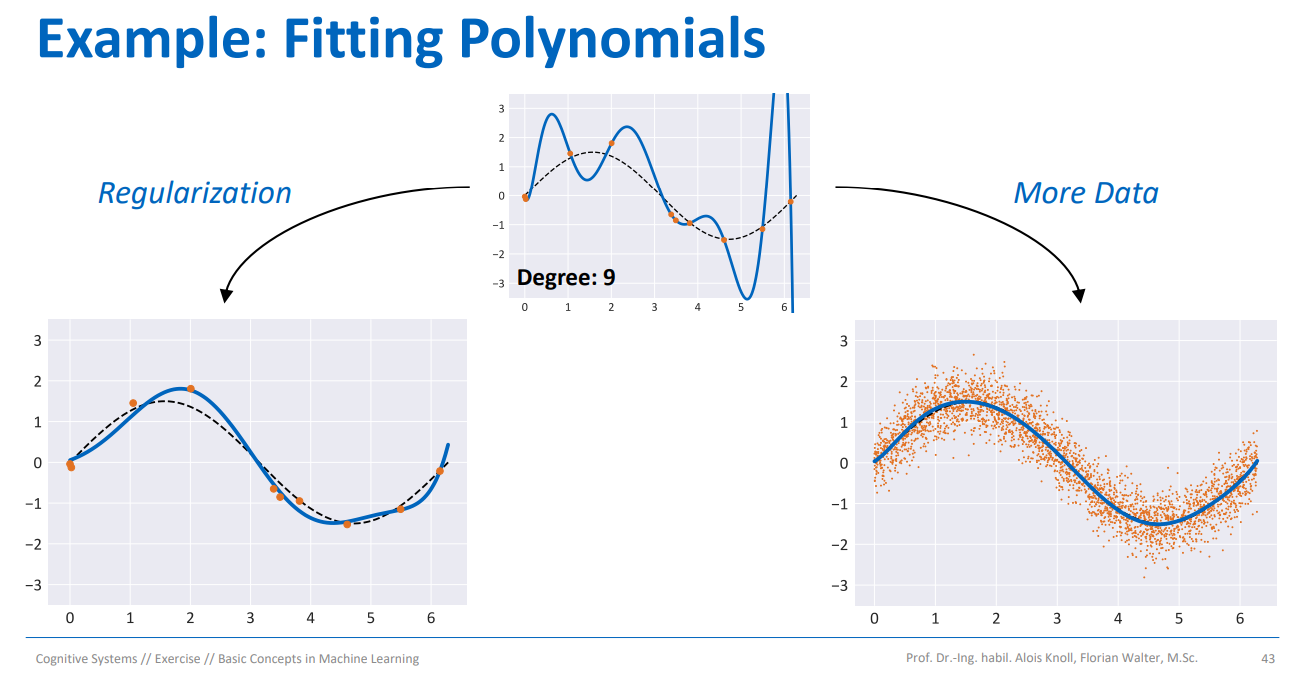
* If h is modified based on the performance of the test set, the test set is invalidated.
* *Cross-validation*: k-fold cross-validation enables an efficient use of data at the cost of computational complexity: data is partitioned into k subsets and learned in k iterations. In every iteration, a different subset is selected as a validation set. The overall performance corresponds to the average performance of the K iteration.

### Avoidance of overfitting

* *Regularization term* in the objective function L guides the learning process towards simpler solutions by punishing complexity.
* *More training data* - increasing the size (i.e. the complexity) of the dataset.
* *Dataset augmentation* - if not enough training data is available, the dataset can be increased by applying transformations to the training samples.

🐇Occam’s Razor: in general, it is desirable to choose a model that is as simple as possible. This law of parsimony is stated in the principle of occam’s razor: “of two competing theories, the simpler explanation of an entity is to be preferred”.





## 🌸 Generative and discriminative models

* *Descriptive models* are based on the posterior probabilities
* *Generative models* are based on the prior probabilities ; predictions can be computed by applying Bayes’ theorem: . Generative models are compact representations of the training data that have considerably less parameters than the dataset S.
* Morphing Faces: <https://vdumoulin.github.io/morphing_faces/>

## 🌸 Generative and discriminative models

### Type of learning

* Supervised learning
* Semi-supervised learning
* Reinforcement learning
* Unsupervised learning

### Model

* Deterministic - Stochastic
* Parametric - Nonparametric
* Generative - Discriminative

### Value domain

* Discrete (classification)
* Continuous (regression)

### Reasoning

* Inductive
* Deductive
* Transudative

## 🌸 Perception learning rule

Goal: separate points through a line

An analogues artificial neuron is defined by its synaptic weights w and an activation function A:

Diagram

Description automatically generated

🐇 The perception is a linear classifier that is based on a single neuron with a digital threshold function.

🐇 By learning rule we mean a procedure for modifying the weights and biases of a network

The perception criterion only pushes incorrectly classified samples.

Perception gradient:

All correctly classified samples should hold this:

The online weighted update rule can be derived by applying stochastic gradient descent and adding learning rate :

The perception learning rule can be derived by applying the weight update:

### Properties of perception learning rules

* If solution exist, e.g., the dataset is linearly separatable, then the perception learning algorithm finds a solution within a finite number of steps (perceptron convergence theorem)
* The solution computed by the algorithm depends on the initialization of the parameters and the order of presentation of the training samples
* The algorithm does not converge for non-linearly separatable data sets, therefore it cannot classify XOR datasets
  + This led to AI winter and abandonment of **connectionism** for almost two decades
    - Connectionism is an approach in the field of cognitive science that hopes to explain [mental](https://en.wikipedia.org/wiki/Mind) phenomena using artificial neural networks (ANN).

Graphical user interface, application

Description automatically generated , linearly non-separatable: Chart

Description automatically generated

# Supervised learning

## 🌸 Basic Regression and Classification

Diagram

Description automatically generated

## Regression: Task definition

* Predict y for given x with minimum loss L.

### Linear basis function models

🐇 Linear basic function models are the class of hypothesis spaces that are linear combinations of functions on the samples x.

Examples of basis functions: linear functions, powers of x, radial basis functions, sigmoid functions

### Computing the error for a weight vector w

* The sum of squares error function

### Interpolation vs Regression

Graphical user interface, chart

Description automatically generated

### Learning w with Least Squares

### Learning w with Linear Least Squares

### Geometrical interpretation

### Learning w with Maximum Likelihood

* Probabilistic Modeling
* Maximum Likelihood
* Maximizing the log-Likelihood

## Classification: Task Definition

* Predict y for given x with minimum error.

### Generalized linear models for classification

* [Analogue neuron models](#_🌸_Perception_learning) are generalized linear models

### Linear separation for binary classification

### Classifying multiple classes

### Representing classes

### One-hot encoding

### Least squares for classification

* Issues of least squares linear classification

### Support vector machines

Support vector machines (SVMs) help to minimize the generalization error by computing a hyperplane that maximizes the margin of the classifier, i.e., the smallest distance between the decision boundary and training samples.

* Computing distances from the decision boundary
* Optimization problem