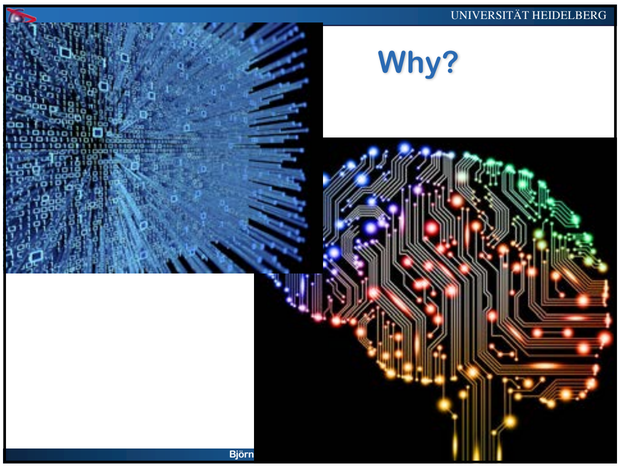
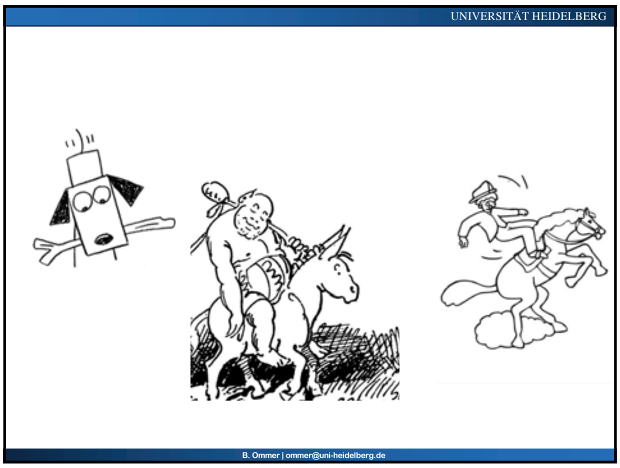
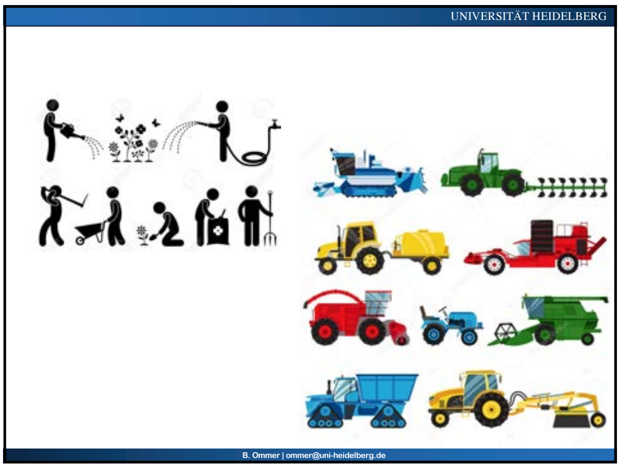

Vision
Computer
group heidelberg
IWR
UNIVERSITÄT
HEIDELBERG
Artificial Intelligence
Prof. Björn Ommer
HCI & IWR
Oct 17, 2016 — 1 Introduction

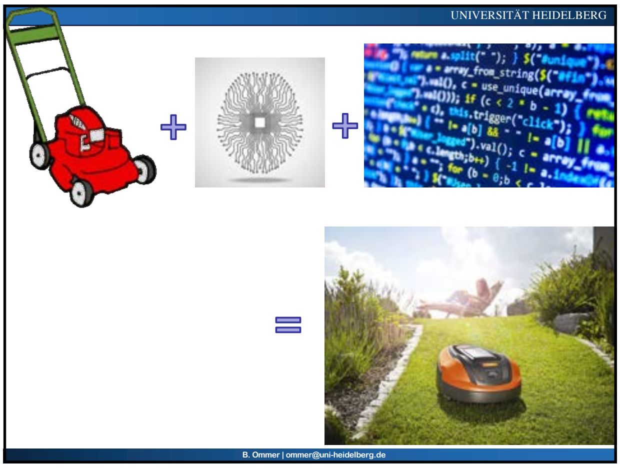

UNIVERSITÄT HEIDELBERG
The Computer Vision Group Heidelberg
Björn Ommer | ommer@uni-heidelberg.de

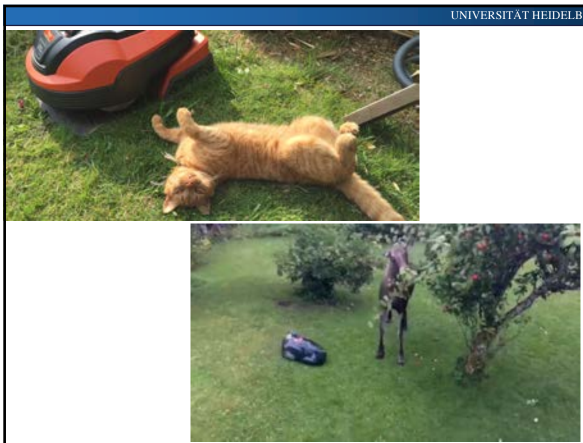

UNIVERSITÄT HEIDELBERG
Why?
Björn


UNIVERSITÄT HEIDELBERG
B. Ommer | ommer@uni-heidelberg.de

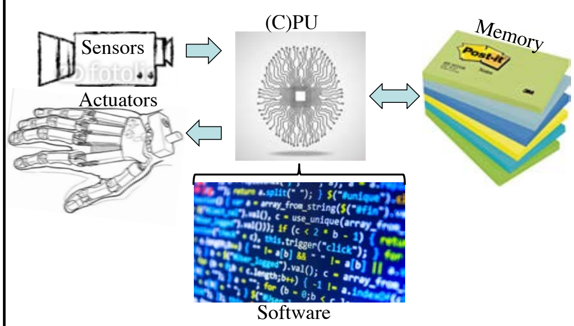

UNIVERSITÄT HEIDELBERG
B. Ommer | ommer@uni-heidelberg.de


UNIVERSITÄT HEIDELBERG
B. Ommer | ommer@uni-heidelberg.de

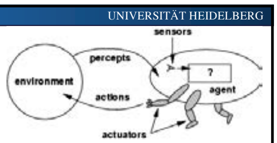


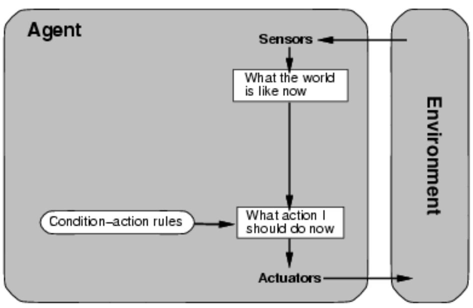


A Generic Agent

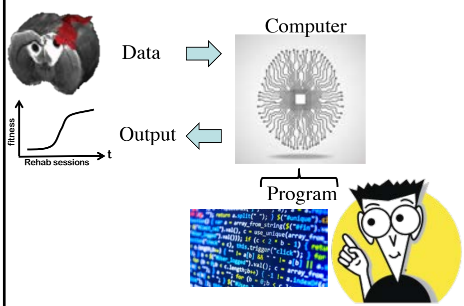


Simple reflex agents

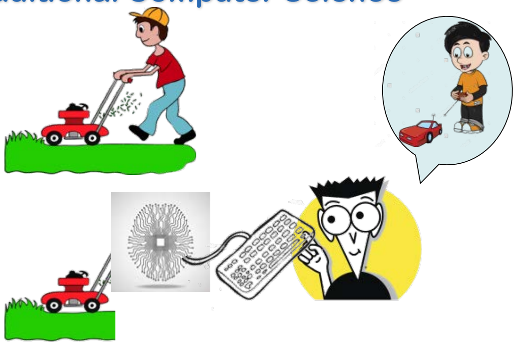




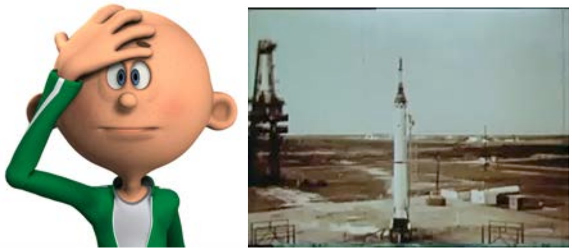
Traditional Computer Science



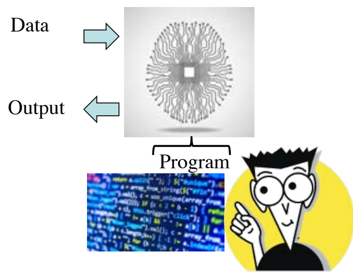
Traditional Computer Science



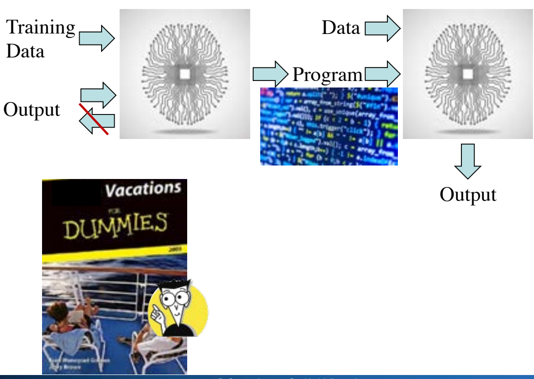
- The Neverending Story of <If .. Then .. Else.. >
- what if ....?



Traditional Computer Science



Artificial Intelligence



Autonomous Driving

Outline

- What is AI?
- A brief history
- The state of the art

Automatic Intelligent Information Processing

- Abilities of computers:
  - memorize large volumes of data
  - fast execution of complex algorithms
  - exact computation
  - never tired, bored, ...
- Limitations of Computers:
  - dealing with fuzzy information
  - deriving abstract concepts
  - creativity, empathy, ...

➪ Syntax vs. Semantics

Artificial Intelligence

Goals (very abstract):

- Automatic, *intelligent* information processing
- Efficient analysis of large volumes of data

Desire: Understand & reconstruct “intelligence“ and make it applicable to real world problems

The need for automatic data processing & AI:

- „We are drowning in information and starved for knowledge“, *John Naisbitt: Megatrends*

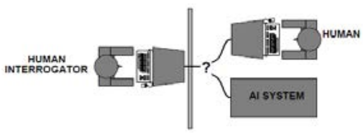
What is AI?

Systems that

think rationally	think like humans
act rationally	act like humans

1 Acting humanly: The Turing test

- Turing (1950) "Computing machinery and intelligence":
- “Can machines think?“ ➪ “Can machines behave intelligently?“
- Operational test for intelligent behavior: the Imitation Game



Acting humanly: The Turing test

- Predicted that by 2000, a machine might have a 30% chance of fooling a lay person for 5 minutes
- Anticipated all major arguments against AI in following 50 years
- Suggested major components of AI: knowledge, reasoning, language understanding, learning

Problem: Turing test is not reproducible, constructive, or amenable to mathematical analysis

ELIZA

- By Joseph Weizenbaum in 1966: simple pattern matching and reacting to keywords

- Examples:
  - Reformulating comments in questions
    - User: I'm having trouble with my car.
    - ELIZA: Why do you say you're having trouble with your car?
  - Reacting to keywords:
    - User: I'm having trouble with my father.
    - ELIZA: Tell me more about your family.
  - Misleading keywords:
    - User: War is the father of all things.
    - ELIZA: Tell me more about your family.

Subfields of AI

- Capabilites needed to pass Turing test:
  - Knowledge presentation
  - Automated reasoning
  - Machine learning
  - Natural language processing
- Total Turing test:
  - Computer vision
  - Robotics
- These capabilites compose most of AI

2 Thinking humanly: Cognitive Science

- 1960s "cognitive revolution": information-processing psychology replaced prevailing orthodoxy of behaviorism
- Requires scientific theories of internal activities of the brain
  - What level of abstraction? "Knowledge" or "circuits"?
  - How to validate? Requires
    - 1) Predicting and testing behavior of human subjects (top-down)
    - or 2) Direct identification from neurological data (bottom-up)
- Both approaches (roughly, Cognitive Science and Cognitive Neuroscience) are now distinct from AI

Thinking humanly: Cognitive Science

- Both share with AI the following characteristic:
  - the available theories do not explain (or produce) anything resembling human-level general intelligence
- Hence, all three fields share one principal direction, but have not achieved their ultimate goal!

Why don't we just copy the human brain?

•People try to but we don't yet have a sufficient understanding of how our visual system works.

•$O(10^{11})$ neurons used just for vision

•By contrast, moderate CPUs have $O(10^{9})$ transistors (most are cache memory)

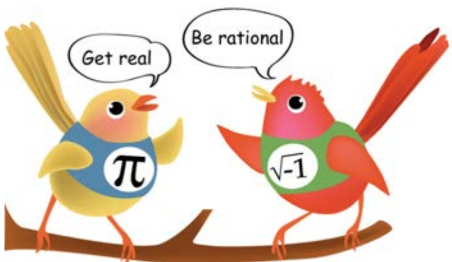
•Very different architectures:
- Brain is slow but parallel
- Computer is fast but mainly serial

•Bird vs Airplane
- Same underlying principles
- Very different hardware

3 Thinking rationally: Laws of Thought

Systems that

think rationally	think like humans
act rationally	act like humans



3 Thinking rationally: Laws of Thought

- Normative (or prescriptive) rather than descriptive
- Aristotle: what are correct arguments/thought processes?
- Several Greek schools developed various forms of logic:
  - notation and rules of derivation for thoughts;
- May or may not have proceeded to the idea of mechanization
- Direct line through mathematics and philosophy to modern AI

Thinking rationally: Laws of Thought

- Problems:
- 1) Not all intelligent behavior is mediated by logical deliberation
- 2) What is the purpose of thinking? What thoughts should I have out of all the thoughts (logical or otherwise) that I could have?

Example: Constructivism

Constructivism: Perception is the result of unconscious inferences about the scene most likely to have caused the retinal image or event.

Hermann von Helmholtz originated the idea of <u>unconscious inference</u> and the likelihood principle.

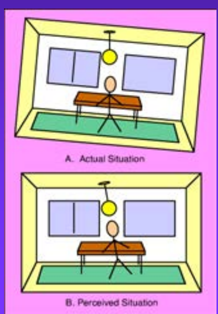
Constructivism

Unconscious Inference: the process of recovering environmental information by logically combining retinal information with heuristic assumptions.

Tilted room illusion:
If you assume that the walls and floor of the room are vertical and horizontal, then <u>you</u> must be tilted —and you <u>feel</u> that way!

Constructivism

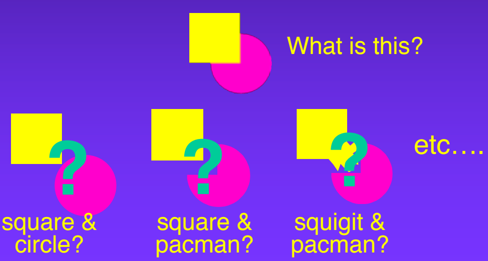
Tilted room illusion



Constructivism

Prägnanz vs. Likelihood

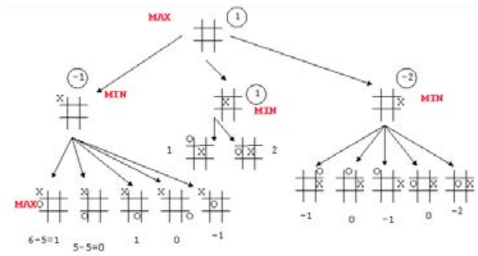
What governs what we see: goodness or probability?



4 Acting rationally

- Rational behavior: doing the *right* thing
- The right thing: that which is expected to maximize goal achievement, given the available information
- Doesn't necessarily involve thinking - e.g., blinking reflex - but thinking should be in the service of rational action
- Aristotle (Nicomachean Ethics):
  - Every art and every inquiry, and similarly every action and pursuit, is thought to aim at some good

- Rational planning by searching for the best strategy *(later in this course)*



Rational agents

- An agent is an entity that perceives and acts
- This course is about designing rational agents
- Abstractly, an agent is a function from percept histories to actions: $f : \mathcal{P}^* \rightarrow \mathcal{A}$
- For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance
- Caveat: computational limitations make perfect rationality unachievable

⇨design best program for given machine resources

AI Prehistory

Philosophy	logic, methods of reasoning
	mind as physical system
	foundations of learning, language, rationality
Mathematics	formal representation and proof
	algorithms, computation, (un)decidability, (in)tractability
	probability
Psychology	adaptation
	phenomena of perception and motor control
	experimental techniques (psychophysics, etc.)
Economics	formal theory of rational decisions
Linguistics	knowledge representation
	grammar
Neuroscience	plastic physical substrate for mental activity
Control theory	homeostatic systems, stability
	simple optimal agent designs

A brief History of AI

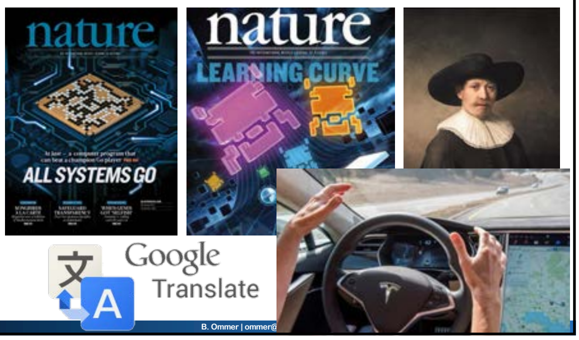
1943	McCulloch & Pitts: Boolean circuit model of brain
1950	Turing's "Computing Machinery and Intelligence"
1952–69	Look, Ma, no hands!
1950s	Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
1956	Dartmouth meeting: "Artificial Intelligence" adopted
1965	Robinson's complete algorithm for logical reasoning
1966–74	AI discovers computational complexity
	Neural network research almost disappears
1969–79	Early development of knowledge-based systems
1980–88	Expert systems industry booms
1988–93	Expert systems industry busts: "AI Winter"
1985–95	Neural networks return to popularity
1988–	Resurgence of probability; general increase in technical depth
	"Nouvelle AI": ALife
1995–	Agents, agents, everywhere . . .
2003–	Human-level AI back on the agenda
2010s	Deep Learning

State-of-the-Art

- Deep Blue defeated the reigning world chess champion Garry Kasparov in 1997
- Proved a mathematical conjecture (Robbins conjecture) unsolved for decades
- No hands across America (driving autonomously 98% of the time from Pittsburgh to San Diego)
- During the 1991 Gulf War, US forces deployed an AI logistics planning and scheduling program that involved up to 50,000 vehicles, cargo, and people
- NASA's on-board autonomous planning program controlled the scheduling of operations for a spacecraft
- Proverb solves crossword puzzles better than most humans

State-of-the-Art

- … and this was all prior to the latest revival.
- Now:



Automatically Generated Stories

One day Joe Bear was hungry. He asked his friend Irving Bird where some honey was. Irving told him there was a beehive in the oak tree. Joe threatened to hit Irving if he didn't tell him where some honey was. The End.

Joe Bear was hungry. He asked Irving Bird where some honey was. Irving refused to tell him, so Joe offered to bring him a worm if he'd tell him where some honey was. Irving agreed. But Joe didn't know where any worms were, so he asked Irving, who refused to say. So Joe offered to bring him a worm if he'd tell him where a worm was. Irving agreed. But Joe didn't know where any worms were, so he asked Irving, who refused to say. So Joe offered to bring him a worm if he'd tell him where a worm was …

Once upon a time there was a dishonest fox and a vain crow. One day the crow was sitting in his tree, holding a piece of cheese in his mouth. He noticed that he was holding the piece of cheese. He became hungry, and swallowed the cheese. The fox walked over to the crow. The End.

Successful Applications

- Speech recognition (e.g. speed-dialing)
- Hand-written character recognition (e.g. letter delivery)
- Information retrieval (e.g. image & video indexing)
- Vision (e.g. face detection & surveillance)
- Operation systems (e.g. caching)
- Fraud detection (e.g. credit cards)
- Text filtering (e.g. email spam filters)
- Game playing (e.g. strategy prediction)
- Robotics (e.g. self localization & mapping)

Subproblems of AI

- Deduction, reasoning, problem solving
  - Using logic or probabilistic reasoning
- Knowledge representation
  - Knowledge bases, modeling
- Planning
  - Defining goals, subgoals, and finding ways to achieve them
- Learning
  - Adapting to a complex world by learning from the past
- Vision
  - Visual object recognition, action recognition, scene understanding
- Natural language processing
  - Understanding text and speech and articulation
- Motion and manipulation
  - Interacting with the world: robotics, autonomous vehicles,…
- Perception
  - Sensing the world (visual, auditory, tactile, …) and finding meaningful representations for percepts
- Social intelligence
  - Social interaction with users
- Creativity, General intelligence
  - Strong AI, AI-completeness

Administrative Interlude

Selected Topics of this Class

- Rational agents
- Solving problems by search: informed/uninformed
- Constrained satisfaction problems
- Game playing & adversarial search
- Inference using logic & knowledge bases
- Reasoning under uncertainty
- Learning
- Neural Networks
- Deep Learning
- Applications
- …

Administrative Issues

- Website for the class:
  - *https://hci.iwr.uni-heidelberg.de/Teaching_AI_Lecture & Moodle*
- TAs:
  - Uta Büchler, Nikolai Ufer, Timo Milbich, Biagio Brattoli, Artsiom Sanakoyeu, Ekaterina Sutter, Tobias Dencker
  - <firstname.lastname@iwr.uni-heidelberg.de>
- Lectures / Exercise groups:
  - Lecture: Tue 14-16, Mathematikon (INF 205), HS 1
Ex Grp: Fri 14-16, Mathematikon, SR C
  - Exercises starting in week no. <*tba*>
- “Scheinkriterien”:
  - Participation in exercises
  - Successfully handing in >70% of all exercises
  - Proposing & presenting one practical student project
- Enter your name & email in form and register in Moodle with correct address so we can contact you!

Textbook

- Useful book:
  - Russel & Norvig: Artificial Intelligence: A Modern Approach
- Additional resources will be specified during class

…And

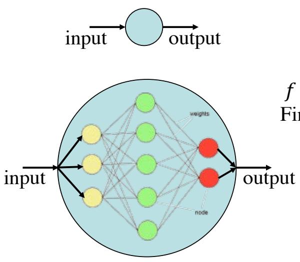
Ask Questions!

Deep Learning in a Nutshell

- … an outlook

Artificial Neural Networks

- Inspired by the brain … but NOT the brain!
- Devide-and-conquer approach to learning complex models

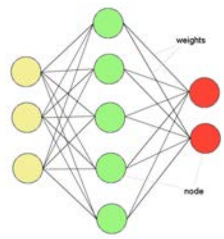


$f = a \circ b \circ c$
Find a, b, c

ANNs – The basics

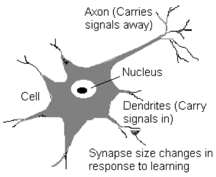
- ANNs incorporate the two fundamental components of biological neural nets:

1. Neurons (nodes)
2. Synapses (weights at connections btw. nodes)

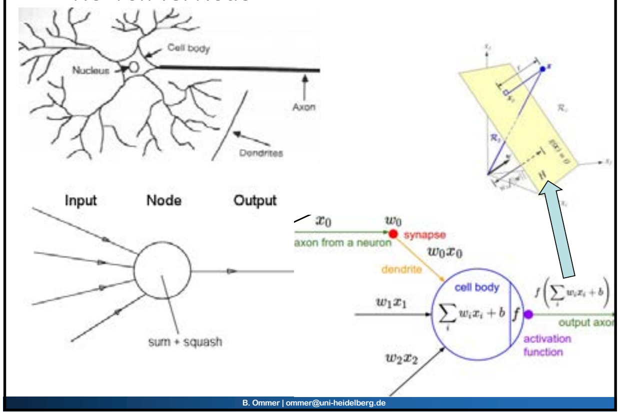


Inspiration from Neurobiology

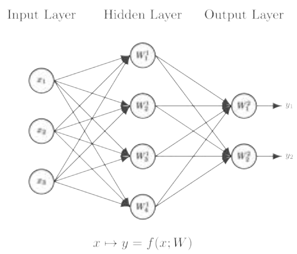
- A neuron: many-inputs / one-output unit
- output can be *excited* or *not excited*
- incoming signals from other neurons determine if the neuron shall excite ("fire")
- Output subject to attenuation in the *synapses,* which are junction parts of the neuron



- Neuron vs. Node



Deep Learning – The Classical NN way



$x \mapsto y = f(x; W)$

Layer k: n_k Nodes
Total Weights: $n_1 \times n_2 + n_2 \times n_3 +$ (some extra)

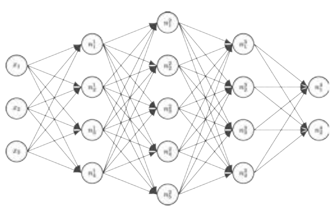
Neural Networks – Expressional Power

Expressional power:

Node:	'and', 'or'
One Hidden Layer:	continuous functions
Two Hidden Layers :	any function

Goal: Learn W^k from training samples $\hat{x}^s, \hat{y}^s, s = 1, \dots, S$??.

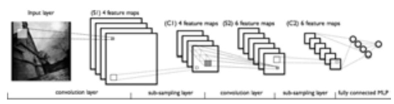
Deep Multilayer Perceptrons



Example:

Input image with N pixels
K layers – approximately same size
Total Weights: approx. KN^2
$\rightsquigarrow$ large set of parameters

The Way Out: Convolutional Neural Networks

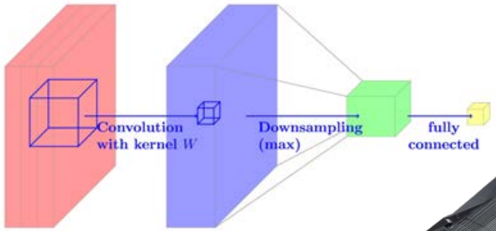


(Image taken from deeplearning.net)

Way out:

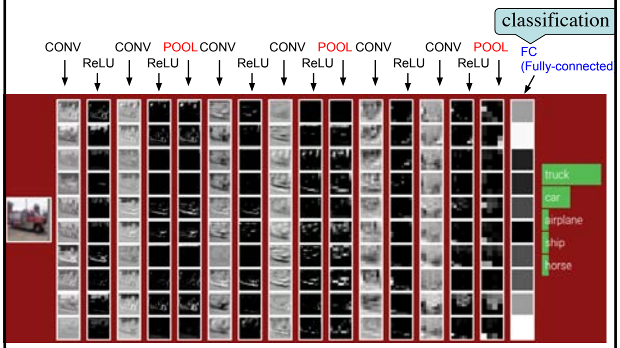
- Use Convolutional Networks: same weights for different image locations
  $\rightsquigarrow$ convolution
- Downsampling (pooling)

Key ingredients of a Convolutional NN



- Same weights for all locations
- All weights learned automatically
- Architecture predefined

UNIVERSITÄT HEIDELBERG

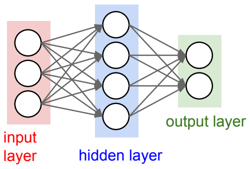


B. Ommer | ommer@uni-heidelberg.de

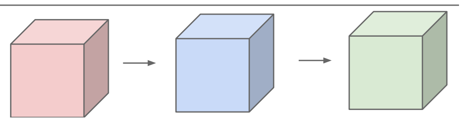
UNIVERSITÄT HEIDELBERG

Convolutional Networks

before:



now:

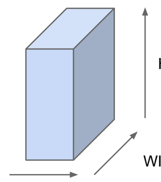


B. Ommer | ommer@uni-heidelberg.de

UNIVERSITÄT HEIDELBERG

Size of Activations

All Neural Net activations arranged in **3 dimensions:**

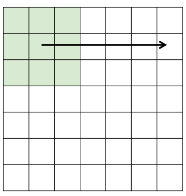


B. Ommer | ommer@uni-heidelberg.de

UNIVERSITÄT HEIDELBERG

Convolution kernel

Replicate this column of hidden neurons across space, with some **stride.**



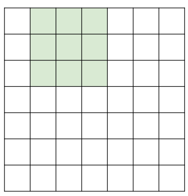
7x7 input
assume 3x3 connectivity, stride 1

B. Ommer | ommer@uni-heidelberg.de

UNIVERSITÄT HEIDELBERG

Convolution kernel

Replicate this column of hidden neurons across space, with some **stride.**

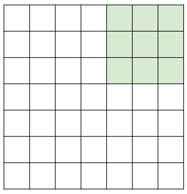


7x7 input
assume 3x3 connectivity, stride 1

B. Ommer | ommer@uni-heidelberg.de

Convolution kernel

Replicate this column of hidden neurons across space, with some **stride.**



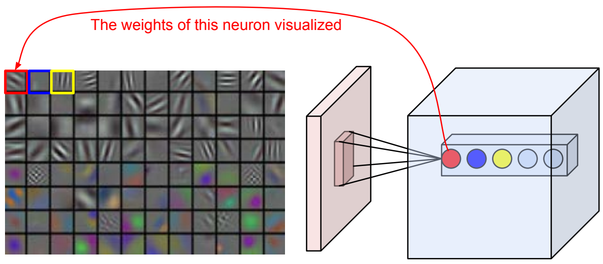
7x7 input
assume 3x3 connectivity, stride 1
=> **5x5 output**

UNIVERSITÄT HEIDELBERG

- **Details to follow …**

B. Ommer | ommer@uni-heidelberg.de

UNIVERSITÄT HEIDELBERG

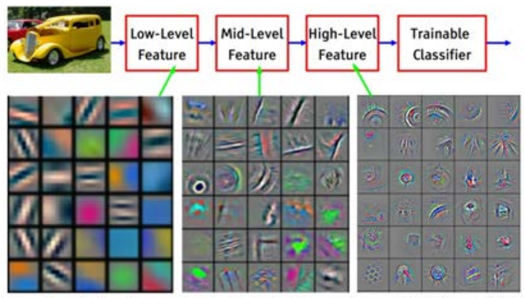


B. Ommer | ommer@uni-heidelberg.de

UNIVERSITÄT HEIDELBERG

A Learned Filterbank

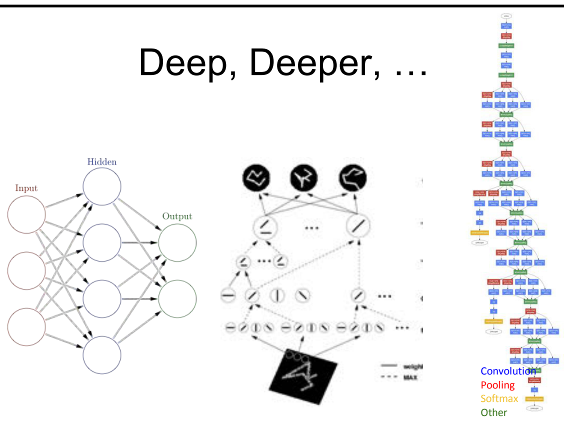
- **CNN is in effect an overcomplete "basis"**



Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]

B. Ommer | ommer@uni-heidelberg.de

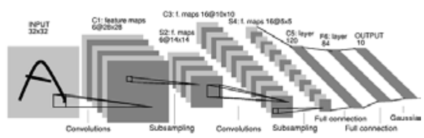
Deep, Deeper, …



Neural Networks, 1980s 2004f 2015f

1998

LeCun et al.



# of transistors

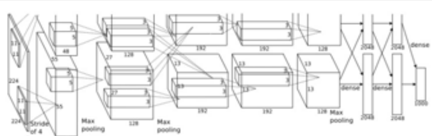
10^6

# of pixels used in training

10^7 NIST

2012

Krizhevsky et al.



# of transistors GPUs

10^9

# of pixels used in training

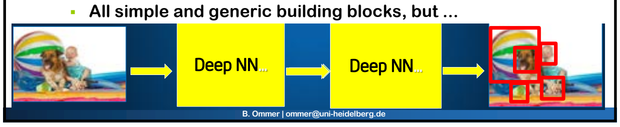
10^{14} IMAGENET

UNIVERSITÄT HEIDELBERG

Deep Learning & CNNs…

- “Classical” Approach:
  - Develop feature representation
  - Develop Learning algorithm

- Promise of deep Learning
  - Build features automatically, but …
  - Combine feature extraction and classification, but …
  - All simple and generic building blocks, but …

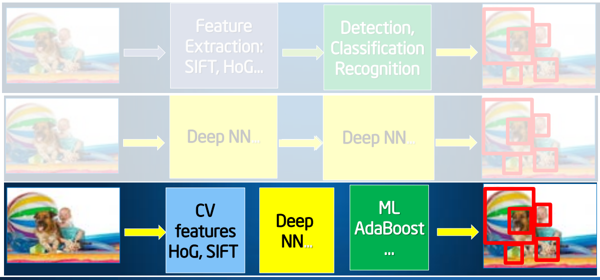


B. Ommer | ommer@uni-heidelberg.de

UNIVERSITÄT HEIDELBERG

Deep Learning & CNNs…

- Best of both worlds
  - Combine predefined features with learned ones
  - Machine learning beyond SGD



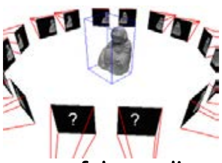
B. Ommer | ommer@uni-heidelberg.de

UNIVERSITÄT HEIDELBERG

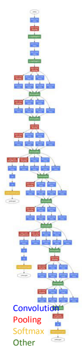
Deep Learning & CNNs…

GoogLeNet

- End-to-end learning with few generic building blocks that are deeply stacked
- Lots of parameters … and lots of guidelines for dealing with them
- Big data



- CNNs are a powerful paradigm, but there are other views on learning that are still helpful

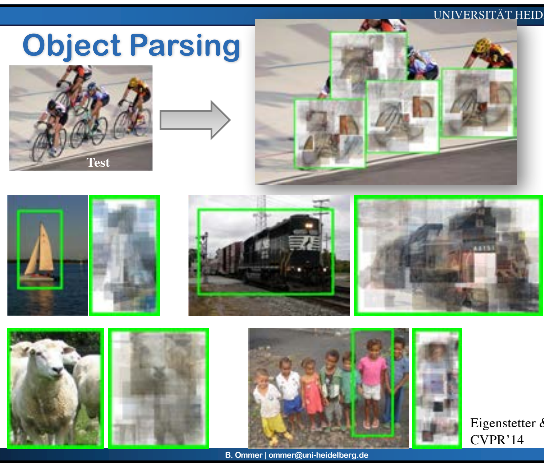


[Szegedy arxiv 2014]

B. Ommer | ommer@uni-heidelberg.de

UNIVERSITÄT HEIDELBERG

Object Parsing



Eigenstetter & Ommer, CVPR’14

B. Ommer | ommer@uni-heidelberg.de

UNIVERSITÄT HEIDELBERG

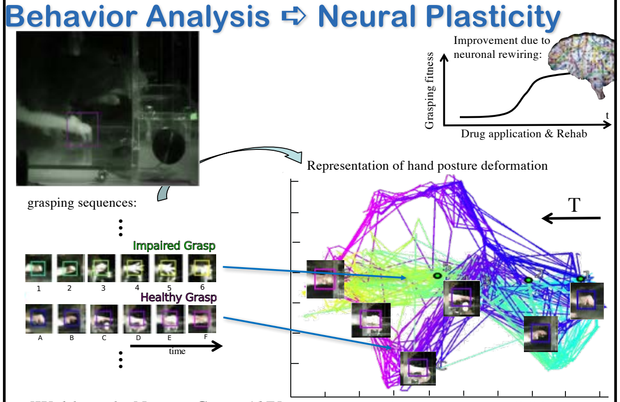
Video Understanding

Milbich & Ommer, ICCV’17

B. Ommer | ommer@uni-heidelberg.de

UNIVERSITÄT HEIDELBERG

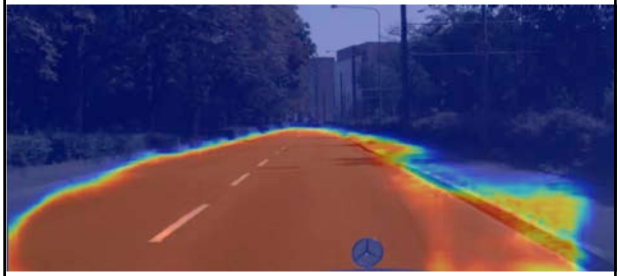
Behavior Analysis ➪ Neural Plasticity



[Wahl et al., Nature Comm‘17] *[Antic et al., MICCAI 2015]*

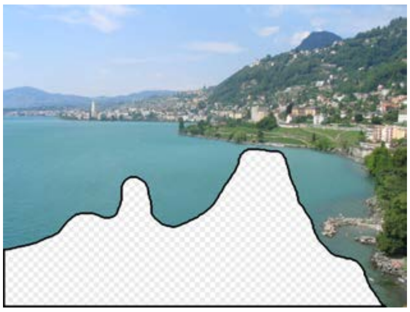
B. Ommer | ommer@uni-heidelberg.de

Unsupervised Learning of Drivable Area

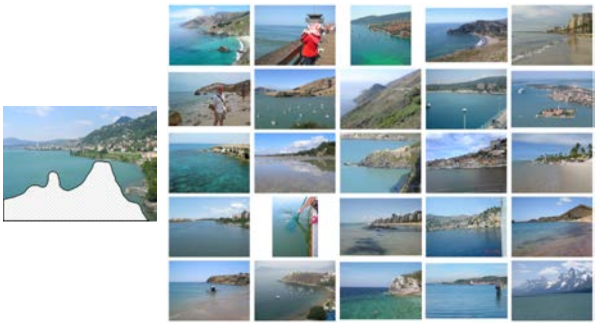


Bautista & Ommer, GCPR'17

Scene Completion

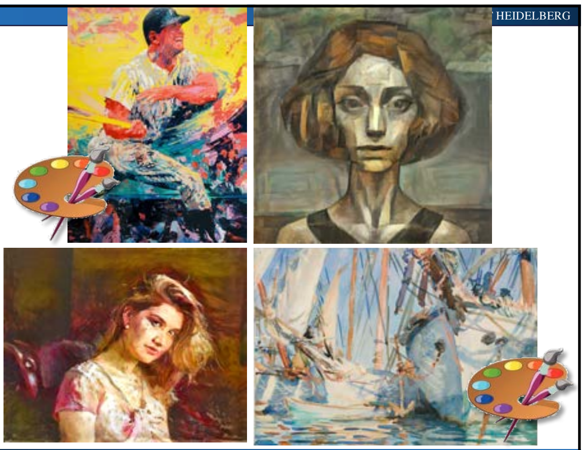


[Hays and Efros. Scene Completion Using Millions of Photographs. SIGGRAPH 2007 and CACM 2008.]



Nearest neighbor scenes from database of 2.3 million photos

Graph cut + Poisson blending



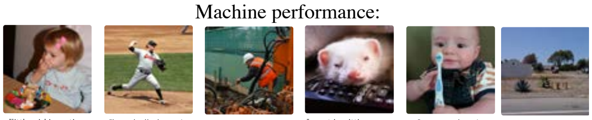
Scene Captioning

Human performance:

PT = 500ms

Some kind of game or fight. Two groups of two men? The foregound pair looked like one was getting a fist in the face. Outdoors seemed like because i have an impression of grass and maybe lines on the grass? That would be why I think perhaps a game, rough game though, more like rugby than football because they pairs weren't in pads and helmets, though I did get the impression of similar clothing. maybe some trees? in the background. (Subject: SM)

Machine performance:



"little girl is eating piece of cake." "baseball player is throwing ball in game." "construction worker in orange safety vest is working on road." "a cat is sitting on a couch with a remote control." "a young boy is holding a baseball bat." "a horse is standing in the middle of a road."

20