# The ShanghAI Lectures

人工智能

上海授课

# The Shanghai Lectures 2022

## Natural and Artificial Intelligence in Embodied Physical Agents

**December 1st, 2022**

**From Zagreb, Croatia**

# Today’s program (CET)

08:30  sites begin connecting
08:55  all sites are ready
09:00 (Fabio) Welcome
09:05 <u>Machine Learning and Deep Learning: an Embodied AI Perspective</u>
10:00 Break
10:10 Guest Lecture by Yulia Sandamirskaya, Intel Germany, Munich, Germany: <u>Neuromorphic Computing</u>
11:00 Wrap-up

# Today’s Guest Lecture

**10:10 Yulia Sandamirskaya,**

**Intel Germany, Munich, Germany**

**«Neuromorphic Computing»**

**Stay tuned!**

# Lecture 5

## Machine Learning and Deep Learning: an Embodied AI Perspective

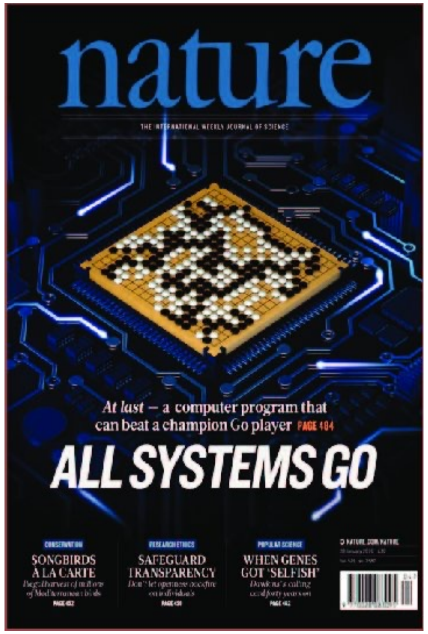
Fabio Bonsignorio
Professor, ERA CHAIR in AI for Robotics

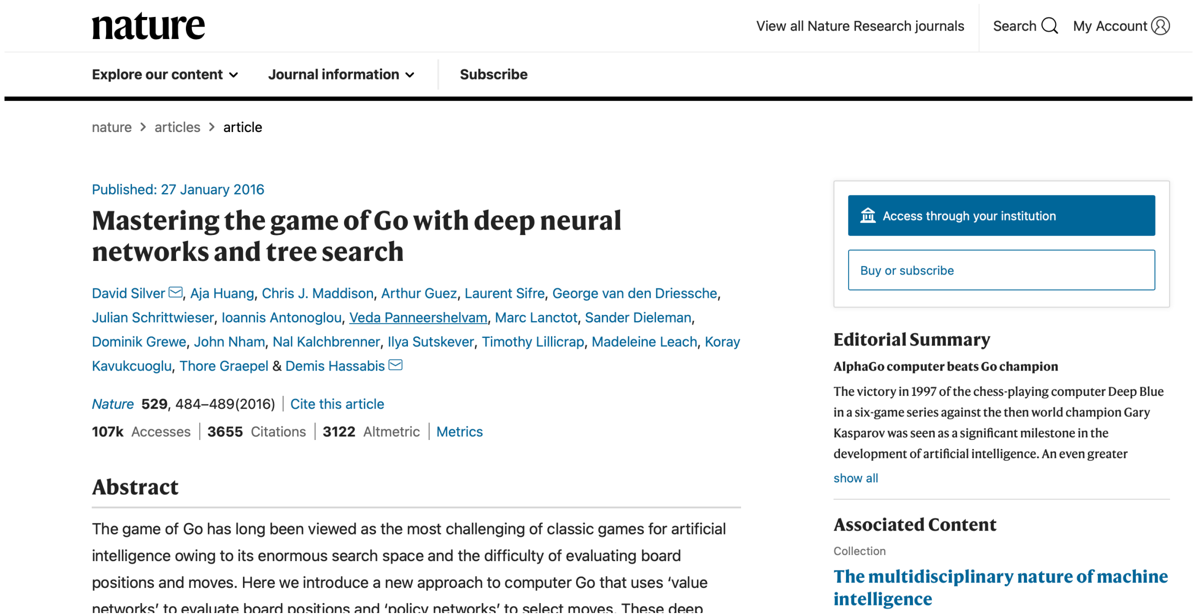
University of Zagreb
Faculty of Electrical Engineering and Computing
Laboratory for Autonomous Systems and Mobile Robotics

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No. 952275

www.heronrobots.com

# Why it matters





nature

View all Nature Research journals | Search | My Account

Explore our content | Journal information | Subscribe

nature > articles > article

Published: 27 January 2016

## Mastering the game of Go with deep neural networks and tree search

David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel & Demis Hassabis

*Nature* **529**, 484–489(2016) | Cite this article

**107k** Accesses | **3655** Citations | **3122** Altmetric | Metrics

### Abstract

The game of Go has long been viewed as the most challenging of classic games for artificial intelligence owing to its enormous search space and the difficulty of evaluating board positions and moves. Here we introduce a new approach to computer Go that uses 'value networks' to evaluate board positions and 'policy networks' to select moves. These deep

Access through your institution

Buy or subscribe

### Editorial Summary

**AlphaGo computer beats Go champion**

The victory in 1997 of the chess-playing computer Deep Blue in a six-game series against the then world champion Gary Kasparov was seen as a significant milestone in the development of artificial intelligence. An even greater

show all

### Associated Content

Collection

**The multidisciplinary nature of machine intelligence**

# Why it matters

nature

View all Nature Research journals Search Logi

Explore our content Journal information Subscribe

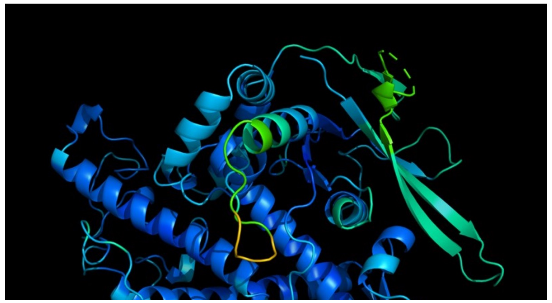
nature news article

NEWS · 30 NOVEMBER 2020

## ‘It will change everything’: DeepMind’s AI makes gigantic leap in solving protein structures

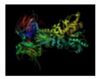
Google’s deep-learning program for determining the 3D shapes of proteins stands to transform biology, say scientists.

Ewen Callaway



RELATED ARTICLES

**AI protein-folding algorithms solve structures faster than ever**



**The revolution will not be crystallized: a new method sweeps through structural biology**

**The computational protein designers**

# Chess: New York, 1997


1 win 3 draws 2 wins

# Go: Hong Kong, 2017

*Google's AlphaGo Defeats Chinese Go Master in Win for A.I.*


Ke Jie, the world's top Go player, reacting during his match on Tuesday against AlphaGo, artificial intelligence software developed by a Google affiliate. China Stringer Network, via Reuters

**By Paul Mozur**

May 23, 2017

阅读简体中文版

HONG KONG — It isn't looking good for humanity.

# A comparison

Chess and GO are ‘perfect information games’

They always have an optimal value function which determines, under perfect game assumptions by all players, the outcome of the game from any initial state s.

The recursion tree in such games will include roughly **$b^d$ moves**

- **<u>Chess</u>: b ≈ 35**, **d ≈ 80**
- **<u>Go</u>: b ≈ 250**, **d ≈ 150**

Interestingly the developers of AlphaGo have implemented an exhaustive testing and evaluation schema to compare and refine different gaming policies by mixing Montecarlo Simulations, Machine learning and guided sampling techniques.

# Remarks

- In **embodied AI (and robotics!)** deterministic approaches are practically impossible to implement -> **No ‘perfect information games’**

- We are very much likely still far from what we have to cope with for a robot operating in the real world, but it can be seen as a better approximation than Chess….and other proposed before.

Silver, D. et al. , Mastering the game of Go with deep neural networks and tree search, Nature 529, 484–489, 2016

# Crash Introduction to ML

Lecture slides from *Deep Learning*

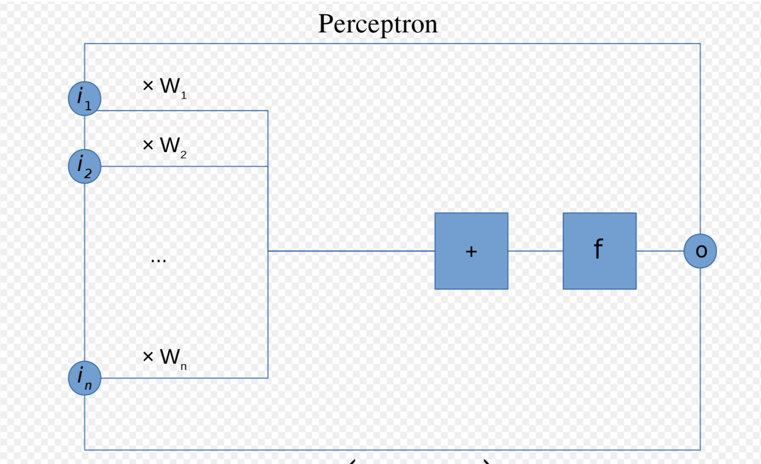
www.deeplearningbook.org

Ian Goodfellow

2016-09-26

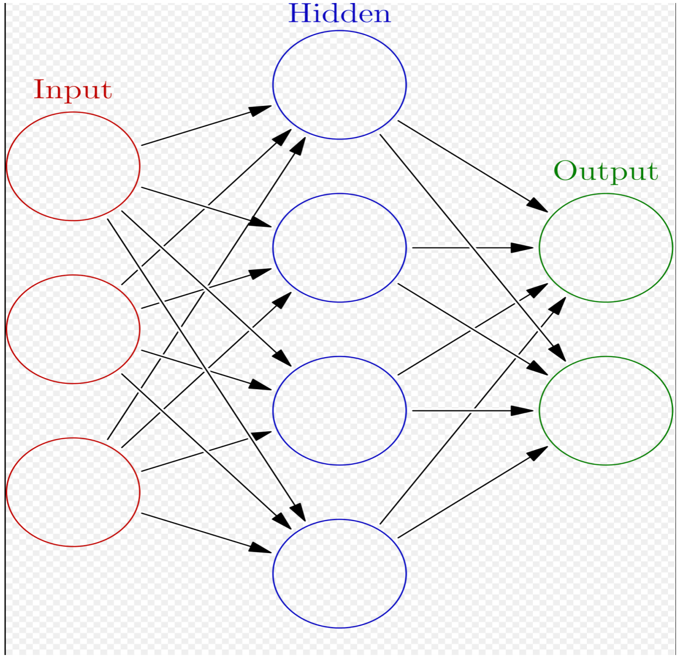
# The beginning

Rosenblatt's Perceptron (1958)




$$o = f\left(\sum_{k=1}^{n} i_k \cdot W_k\right)$$

ANN (Artificial Neural Network) with hidden layers




Picture source: Wikipedia

# Representations Matters (F-O-R!)

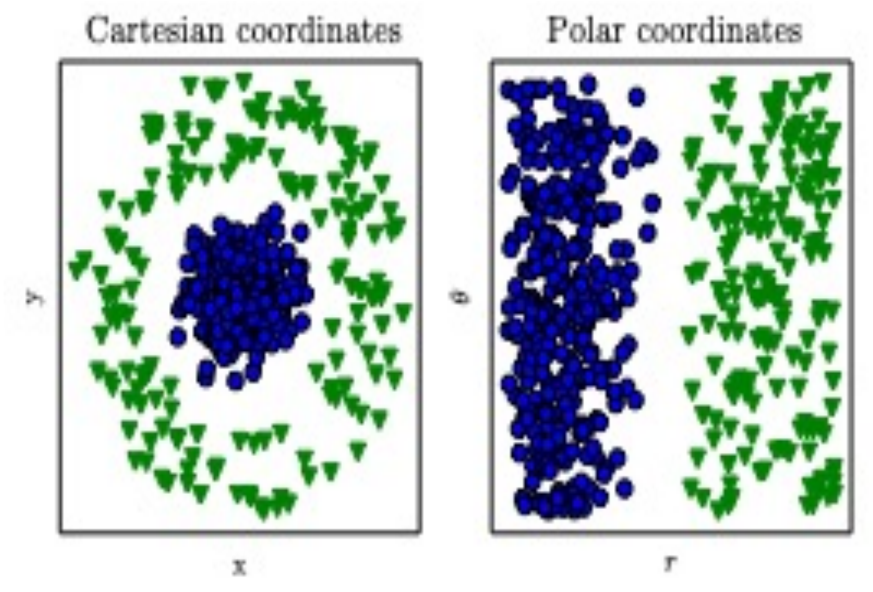



Figure 1.1

# Depth: Repeated Composition

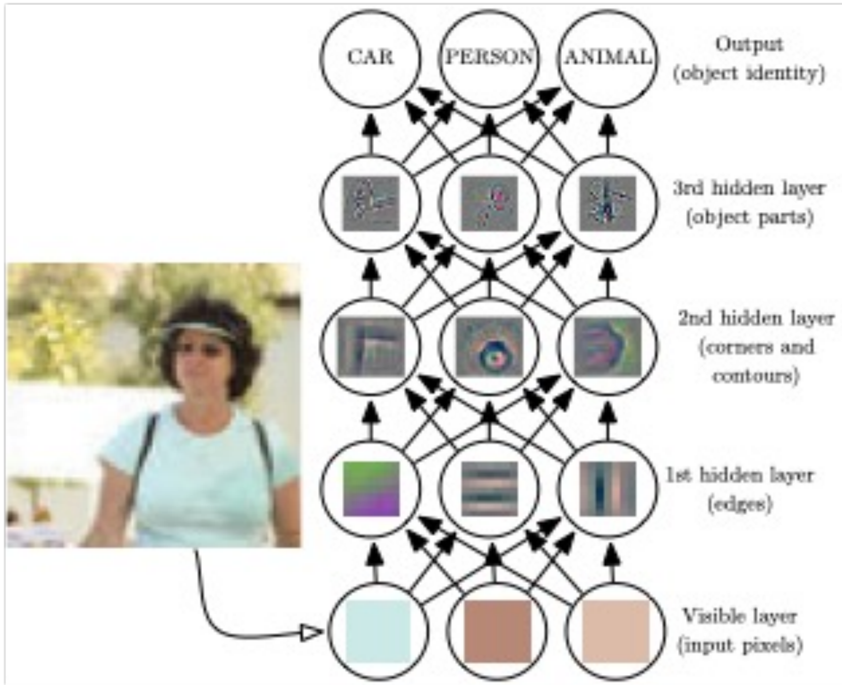



Figure 1.2

# Machine Learning and AI

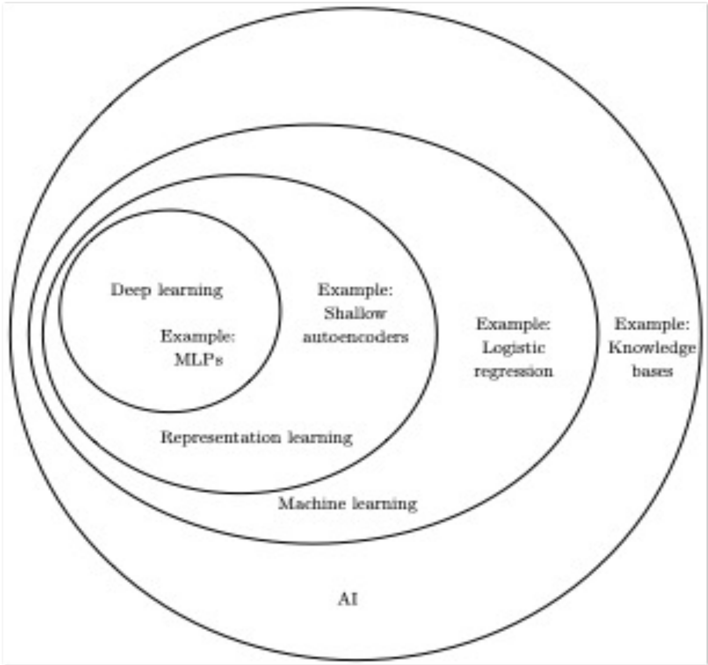



Figure 1.4

# Historical Waves

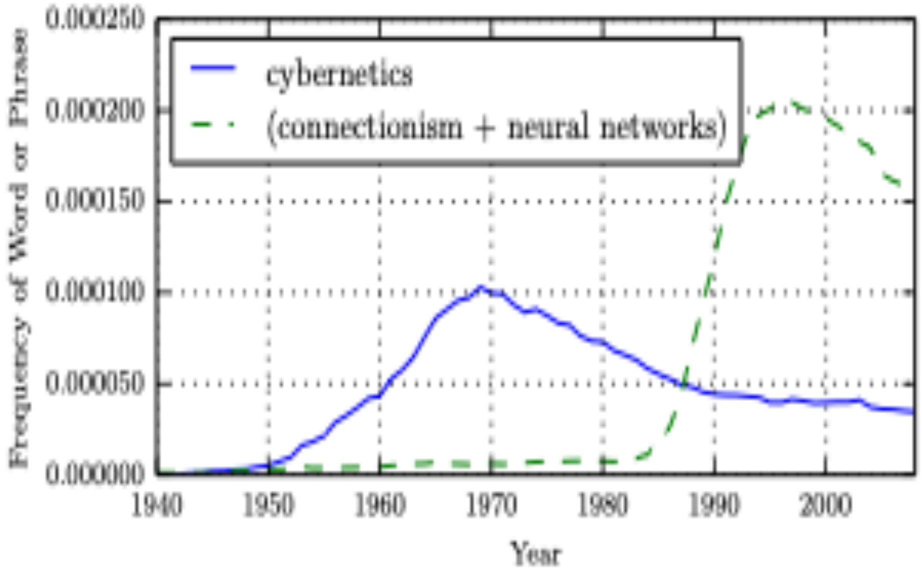



Figure 1.7

# Historical Trends: Growing Datasets

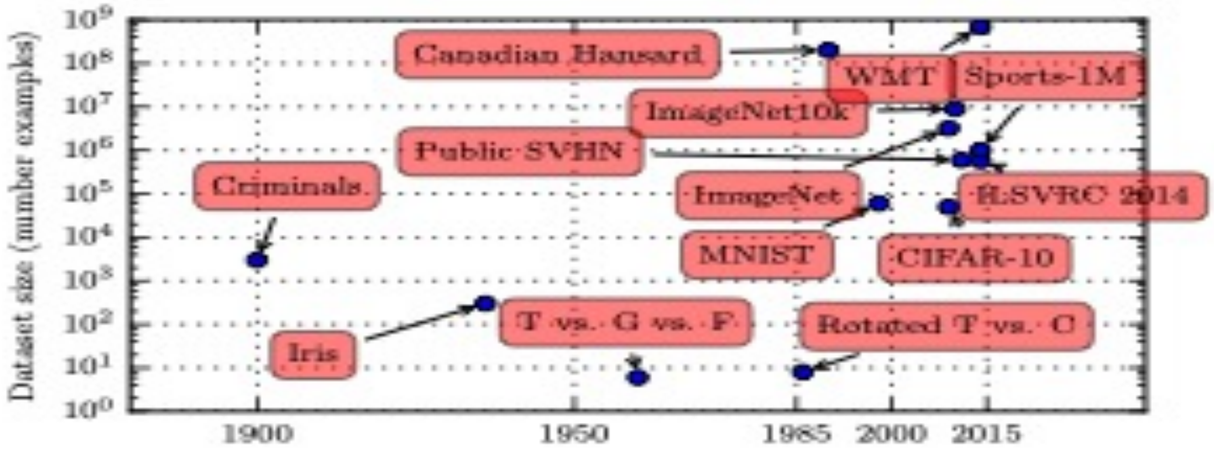



Figure 1.8

# The MNIST Dataset

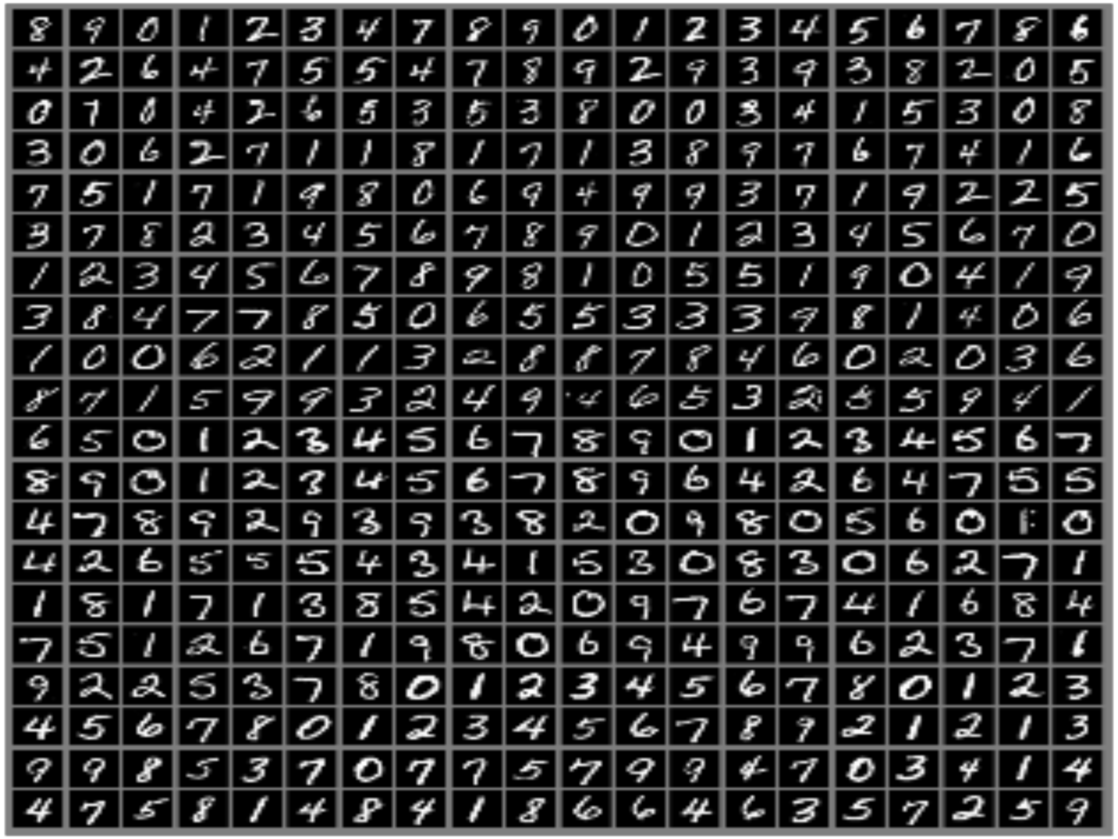


Figure 1.9

# Connections per Neuron

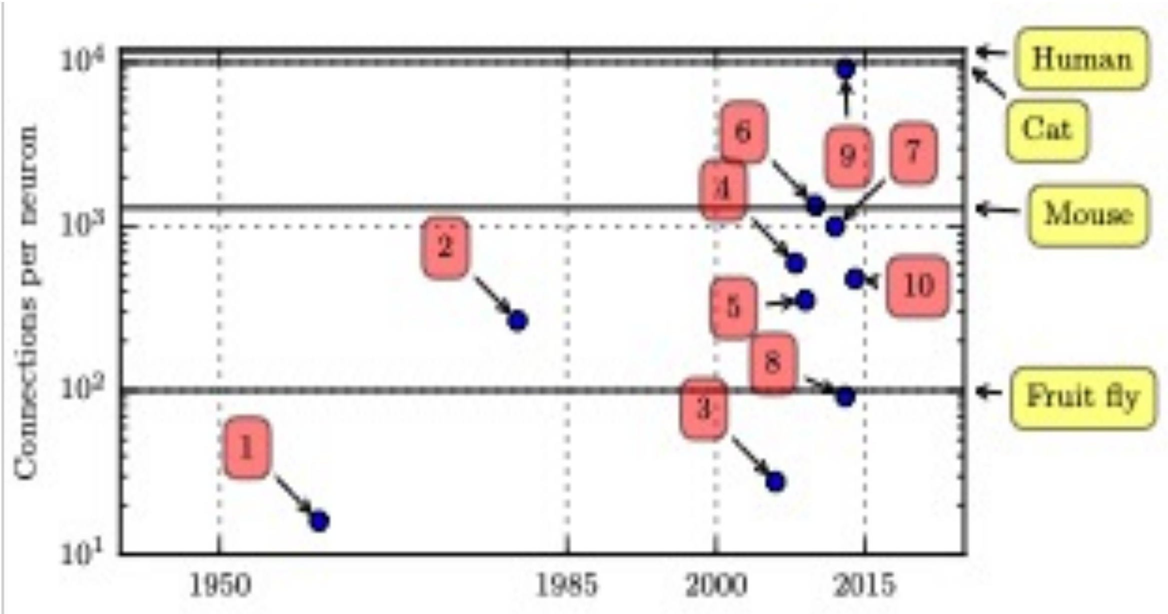



Figure 1.10

# Number of Neurons

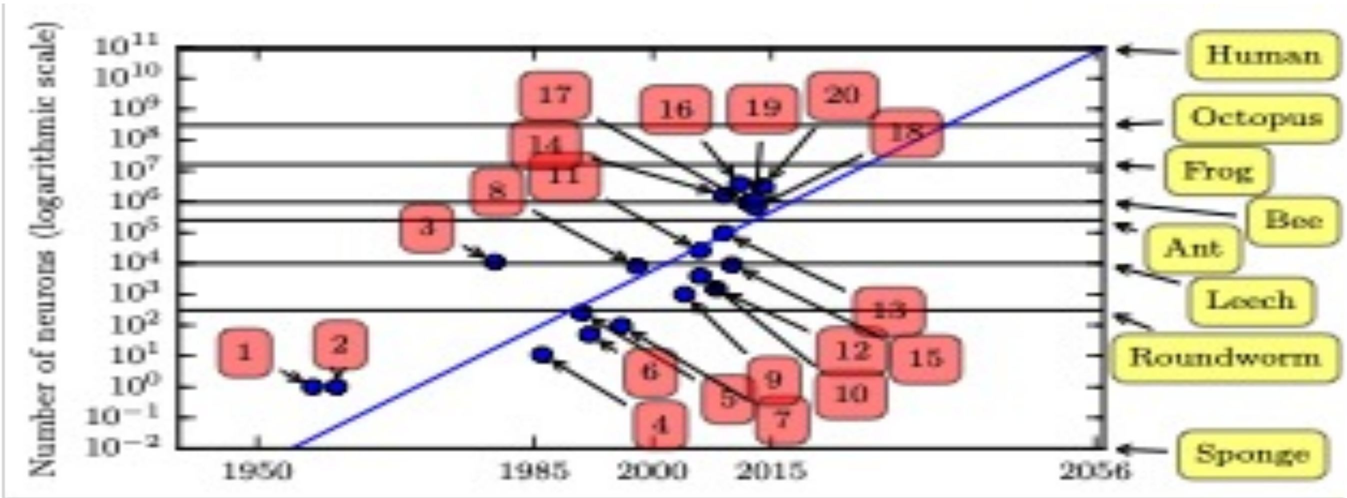



Figure 1.11

# Solving Object Recognition

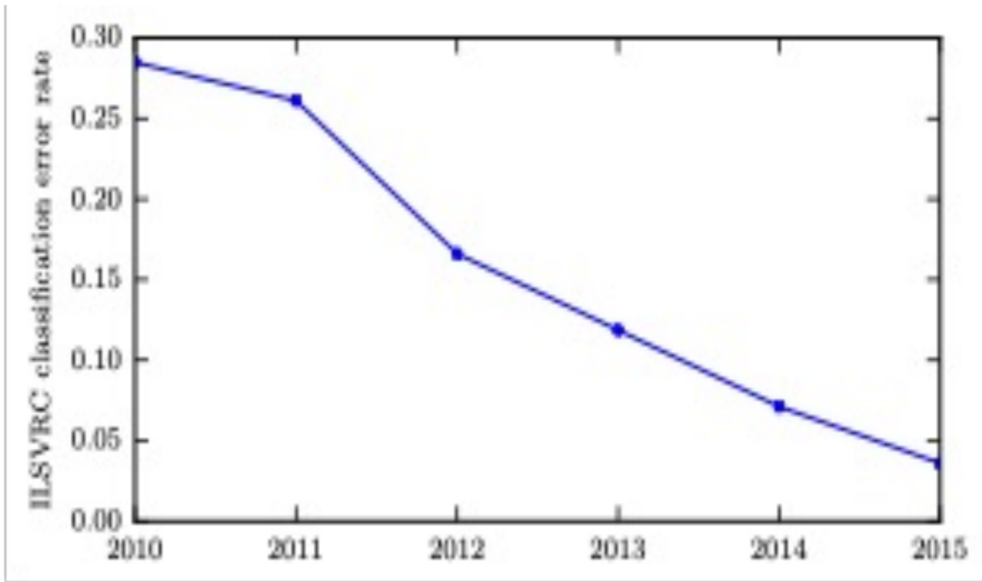



Figure 1.12

# Gradient Descent

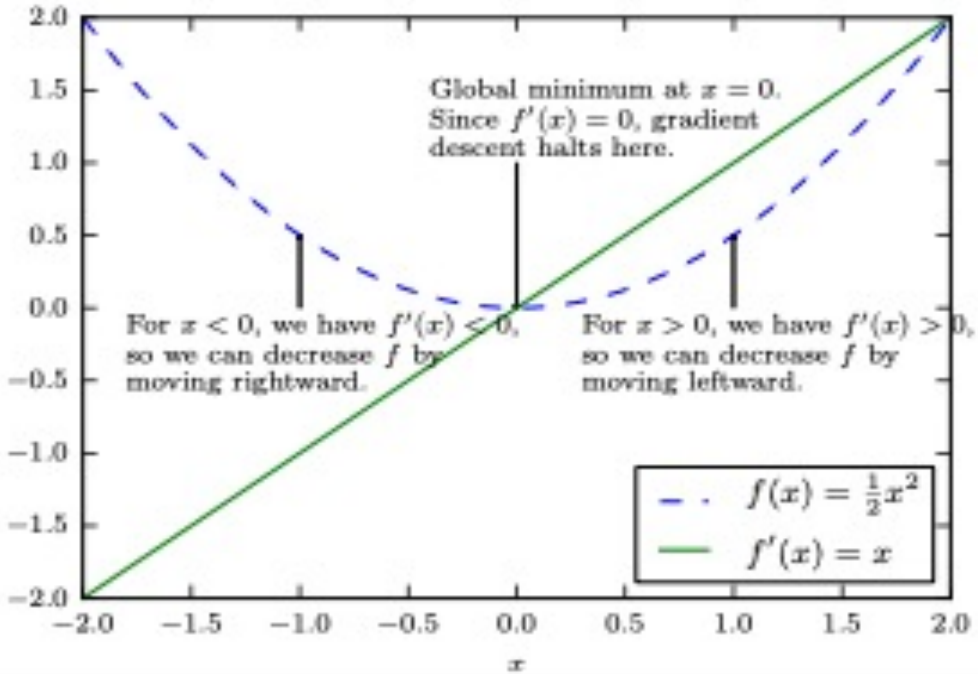



Figure 4.1

# Approximate Optimization

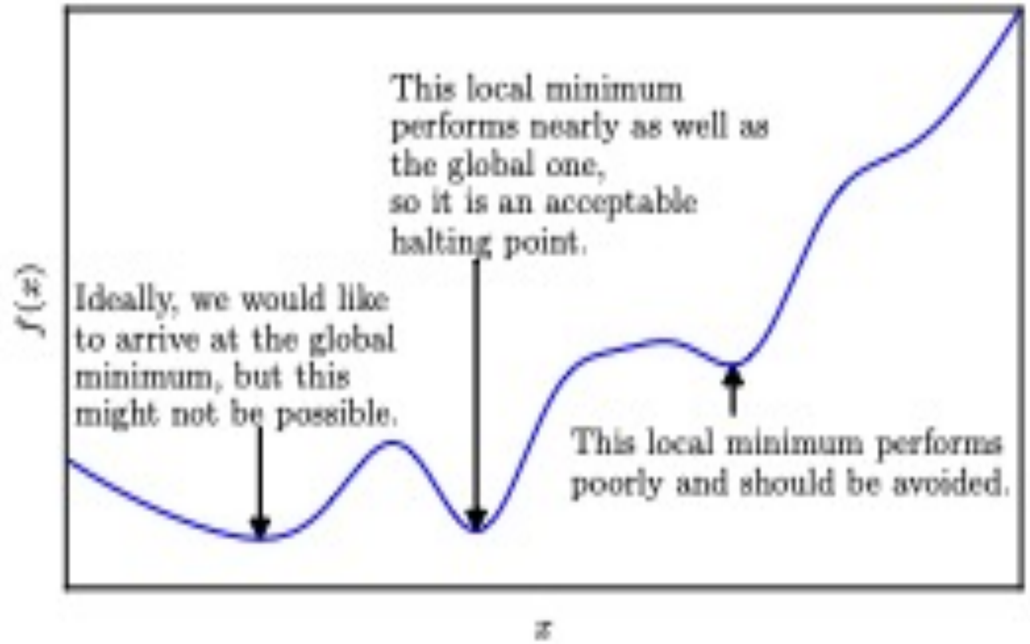
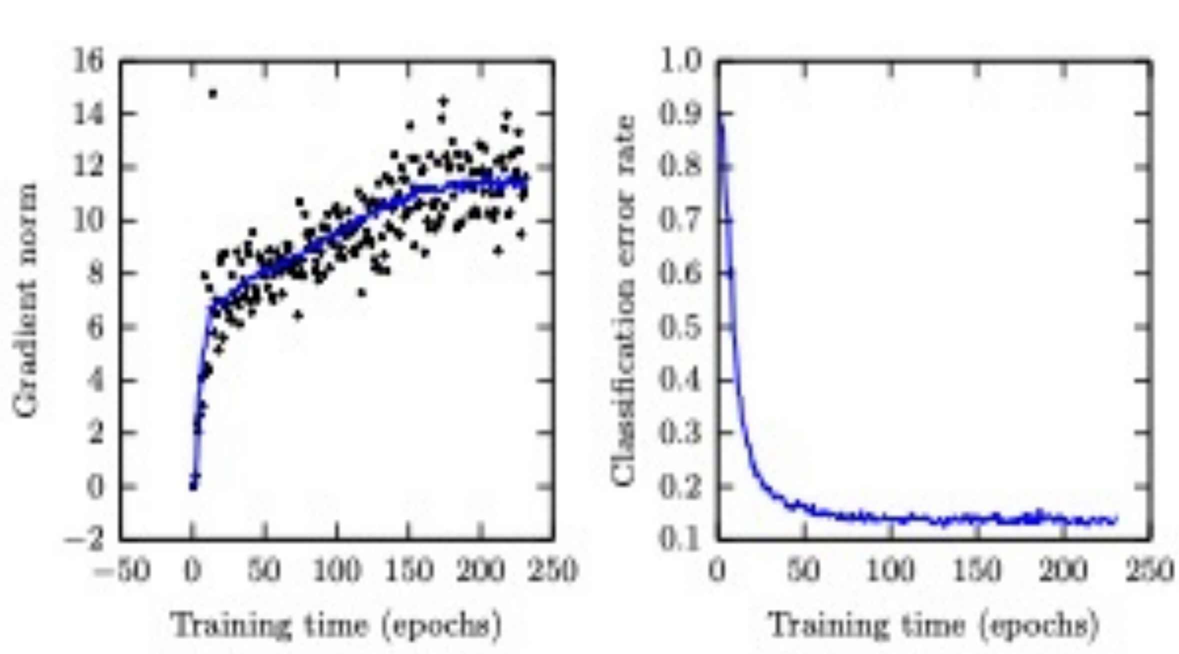



Figure 4.3

# We usually don't even reach a local minimum



# Iterative Optimization

- Gradient descent
- Curvature
- Constrained optimization

# Critical Points

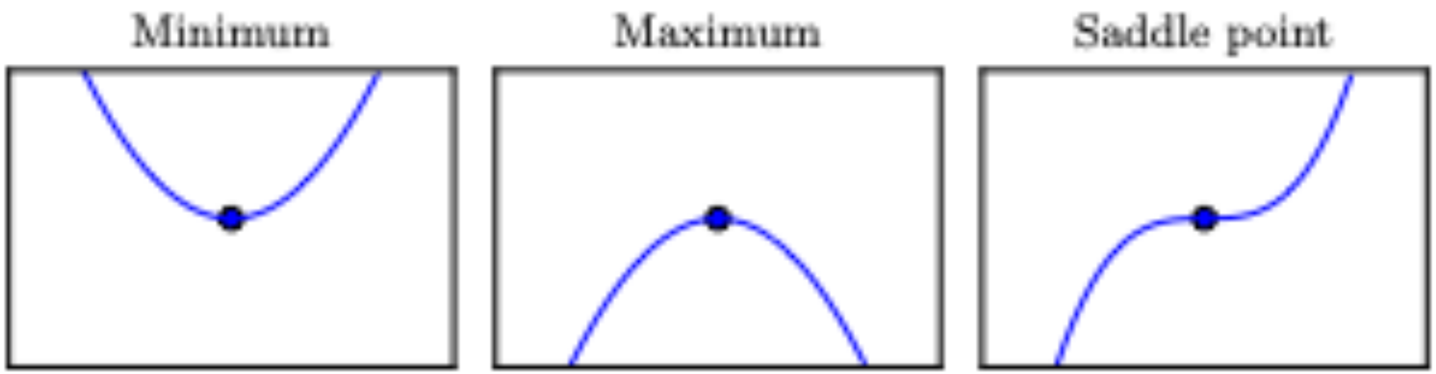



Figure 4.2

# Saddle Points

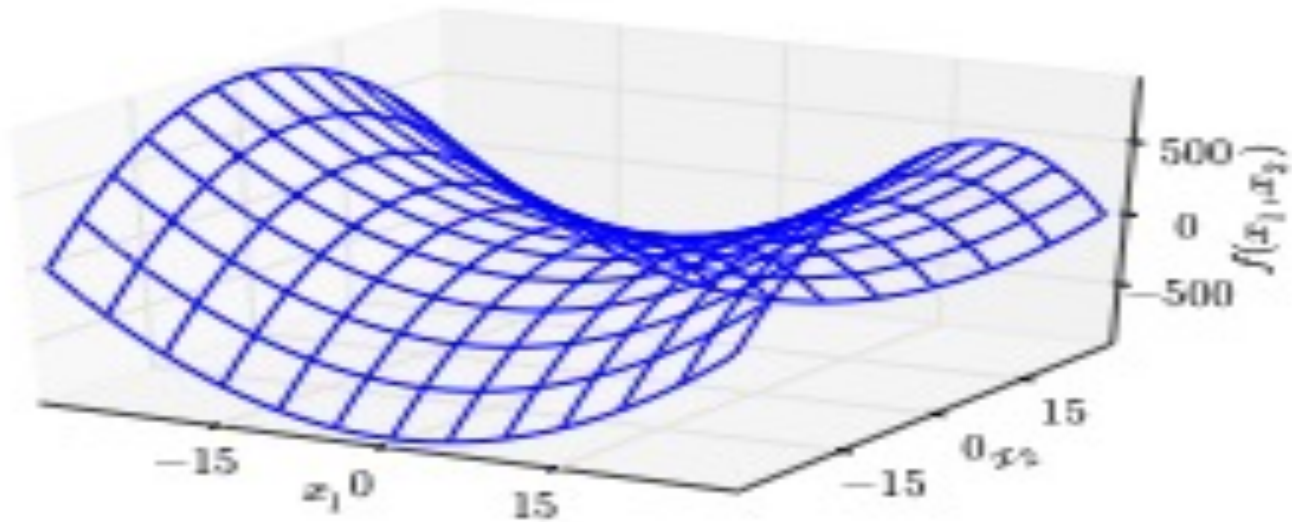



Figure 4.5

Saddle points attract
Newton's method

(Gradient descent escapes,
see Appendix C of "Qualitatively
Characterizing Neural Network
Optimization Problems")

# Curvature

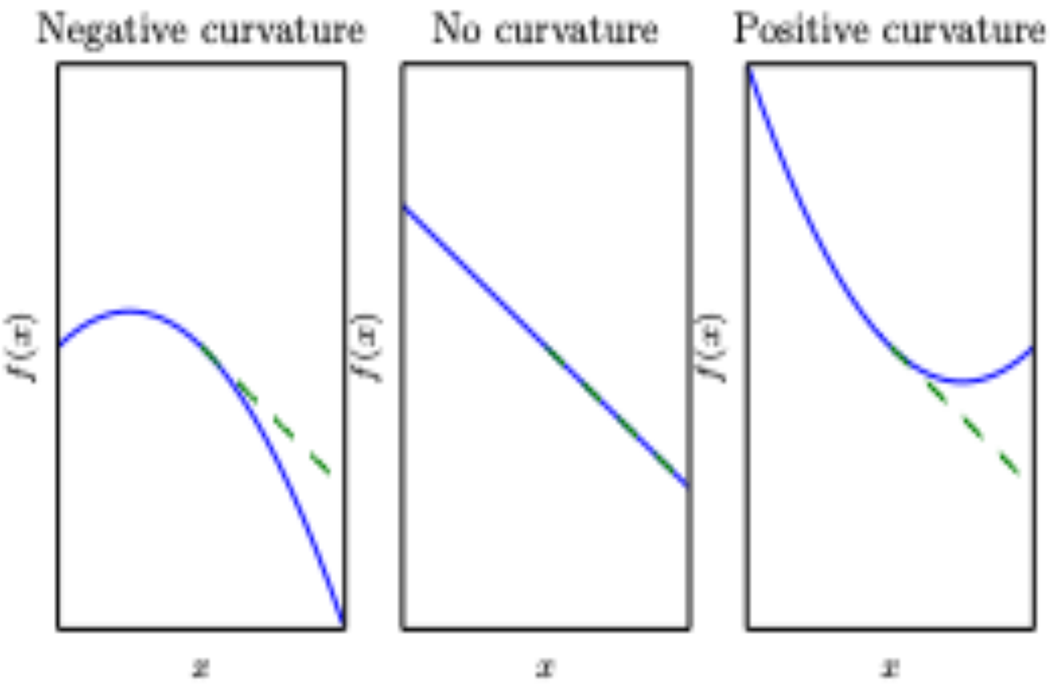
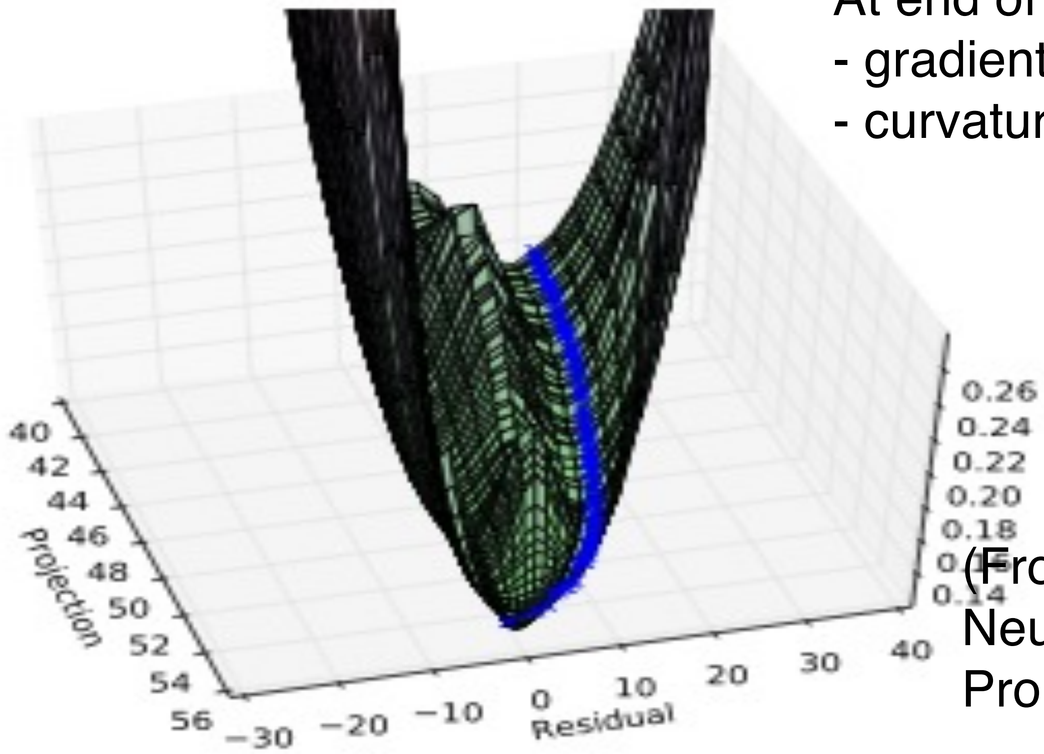



Figure 4.4

# Neural net visualization




At end of learning:
- gradient is still large
- curvature is huge

(From "Qualitatively Characterizing Neural Network Optimization Problems")

# Iterative Optimization

- Gradient descent
- Curvature
- Constrained optimization

# Roadmap

- Iterative Optimization
- Rounding error, underflow, overflow

# Numerical Precision: A deep learning super skill

- Often deep learning algorithms “sort of work”
  - Loss goes down, accuracy gets within a few percentage points of state-of-the-art
  - No “bugs” per se
- Often deep learning algorithms “explode” (NaNs, large values)
- Culprit is often loss of numerical precision

## Rounding and truncation errors

- In a digital computer, we use `float32` or similar schemes to represent real numbers
- A real number *x* is rounded to `x + delta` for some small delta
- Overflow: large *x* replaced by `inf`
- Underflow: small *x* replaced by `0`

# Machine Learning Basics

# Linear Regression

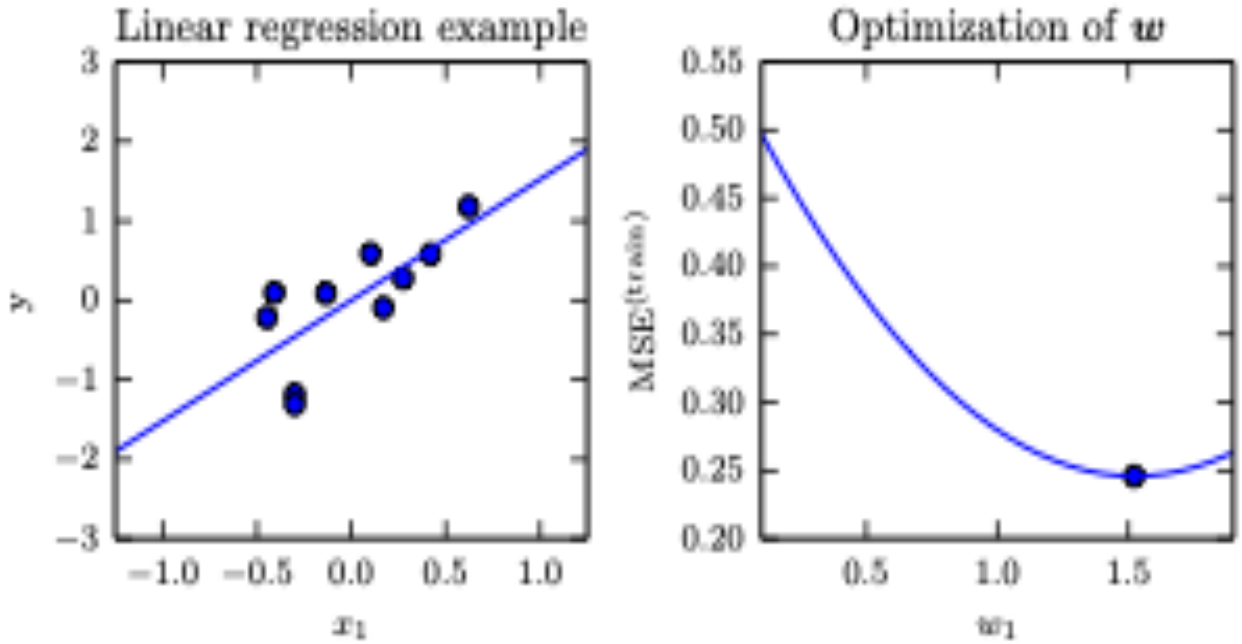



Figure 5.1

# Underfitting and Overfitting in Polynomial Estimation

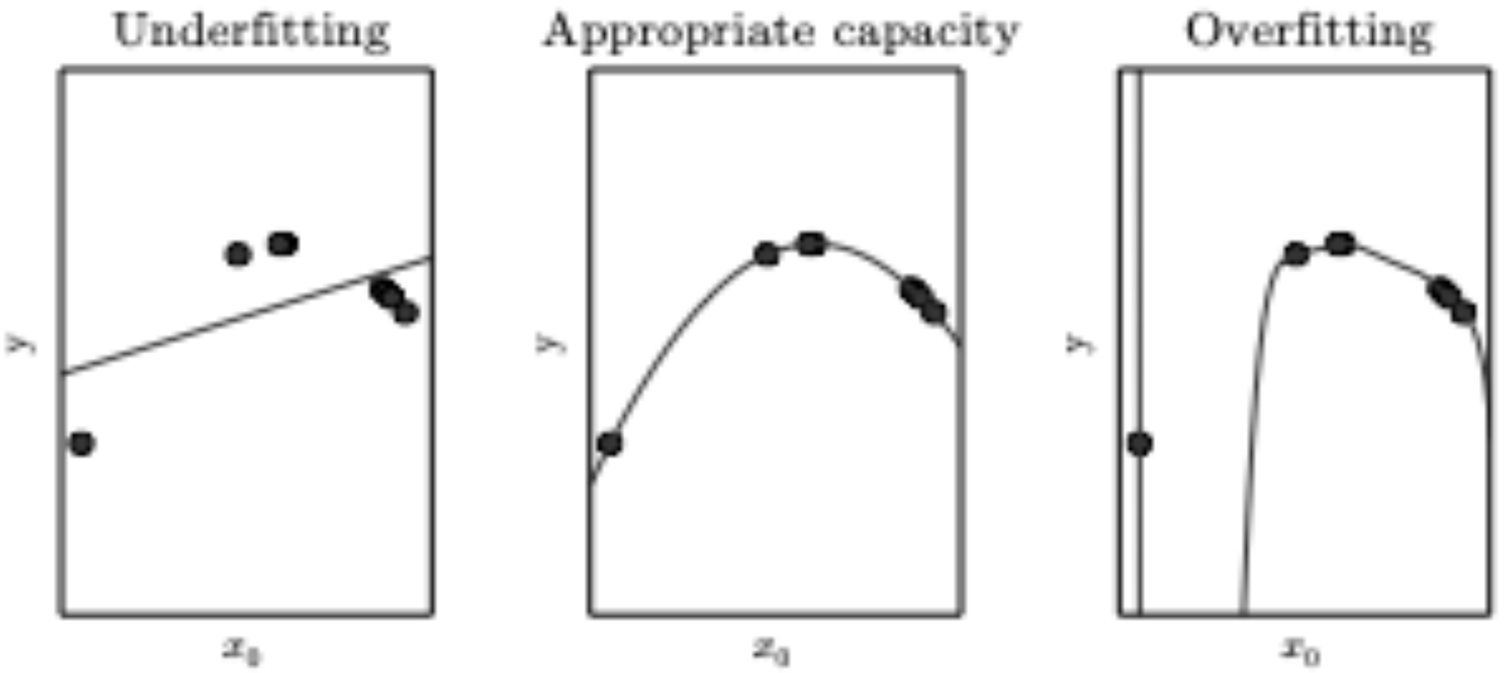



Figure 5.2

# Generalization and Capacity

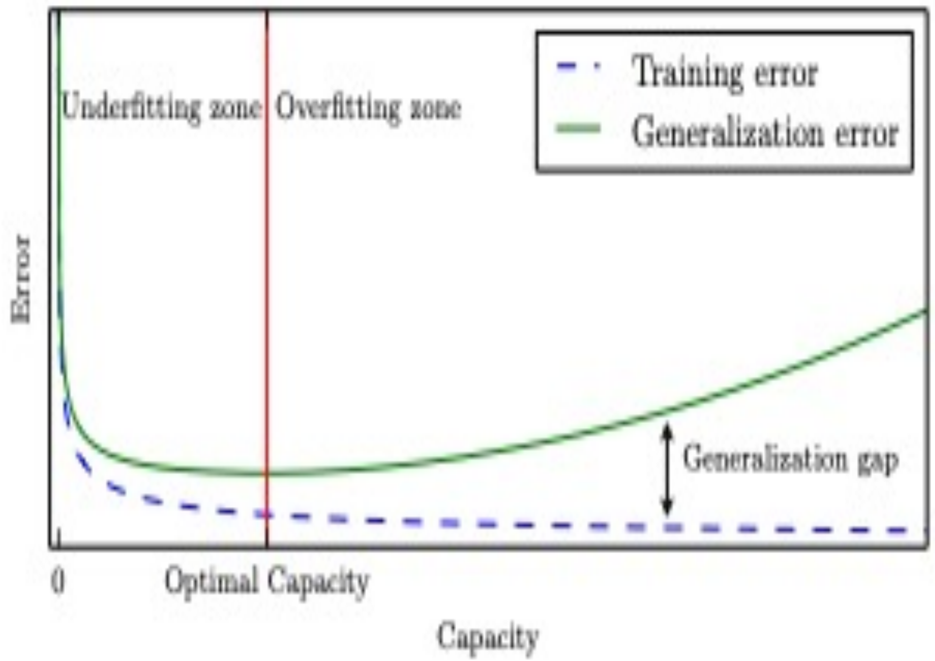
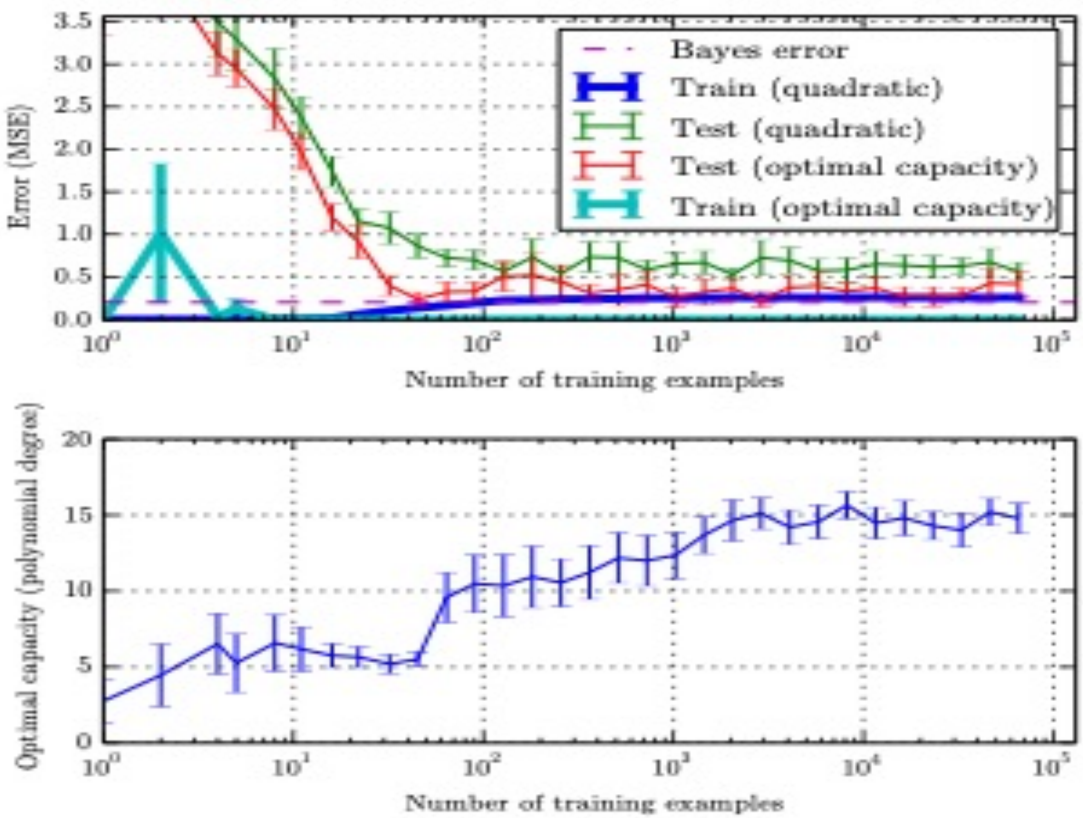



Figure 5.3

# Training Set Size

Figure 5.4



# Weight Decay

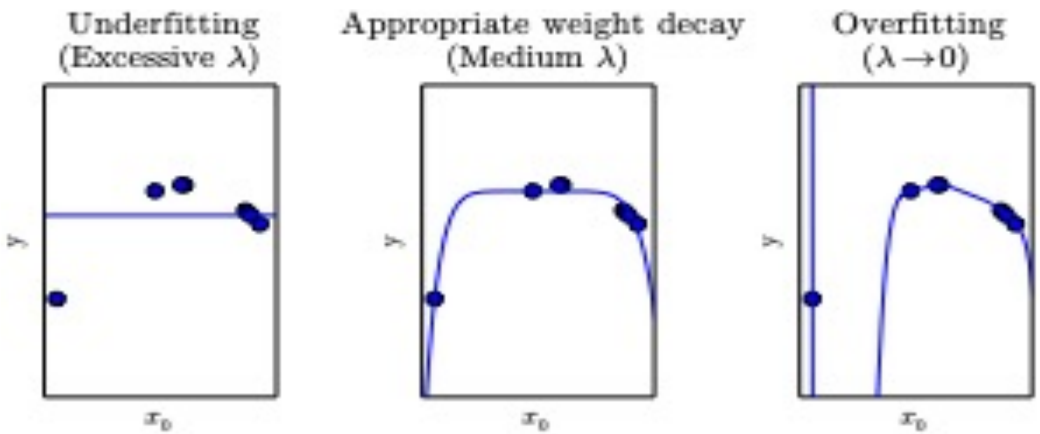



Figure 5.5

# Bias and Variance

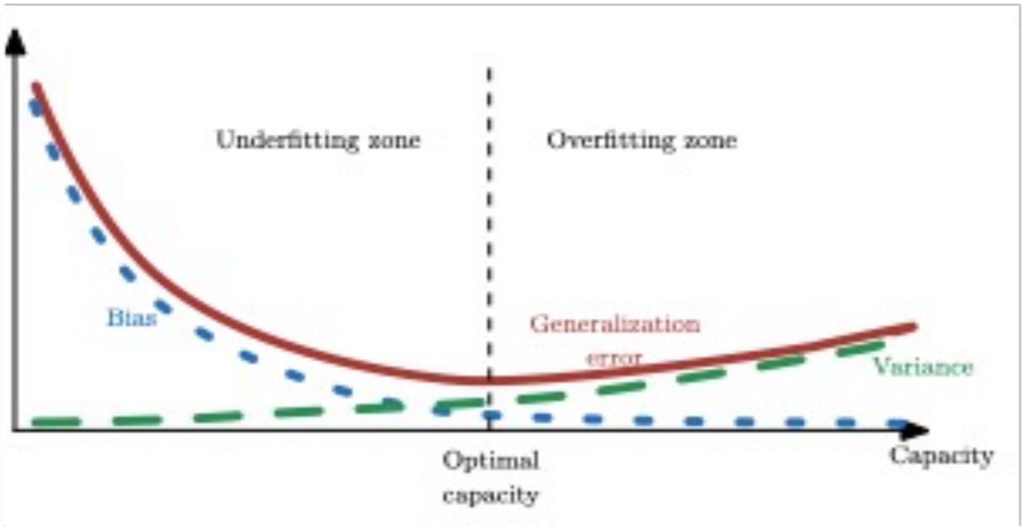
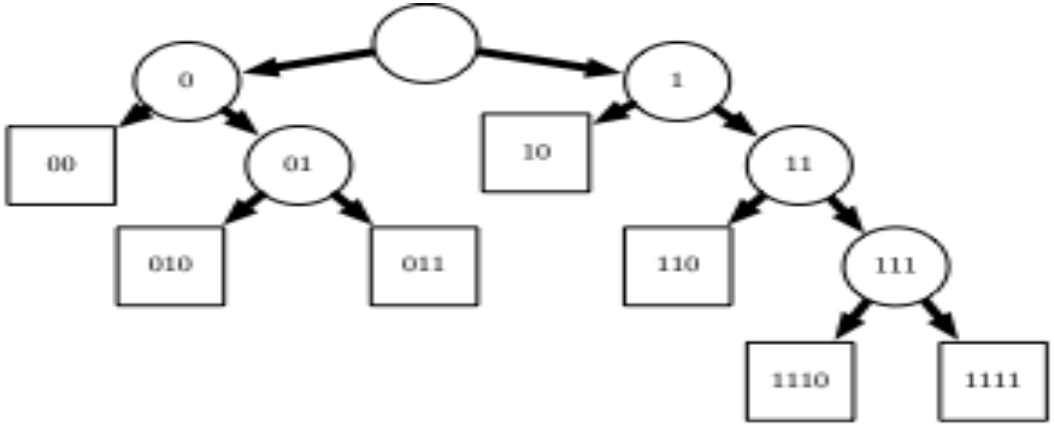



Figure 5.6

# Decision Trees



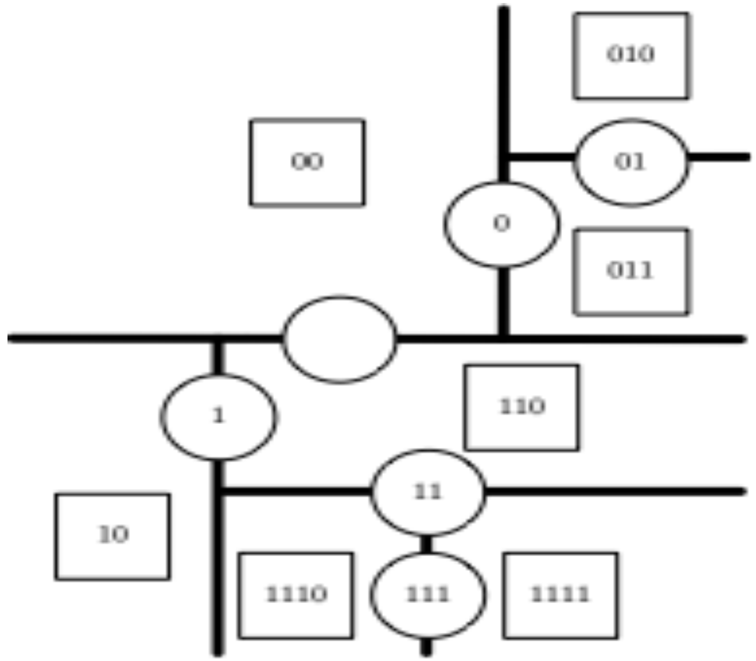




Figure 5.7

# Principal Components Analysis

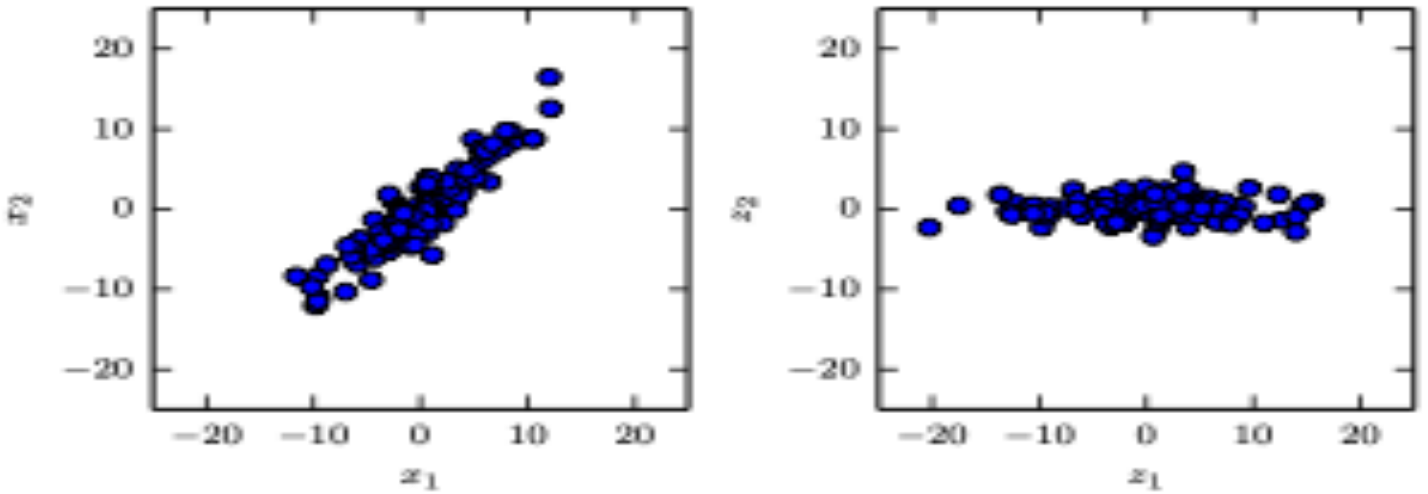



Figure 5.8

# Curse of Dimensionality

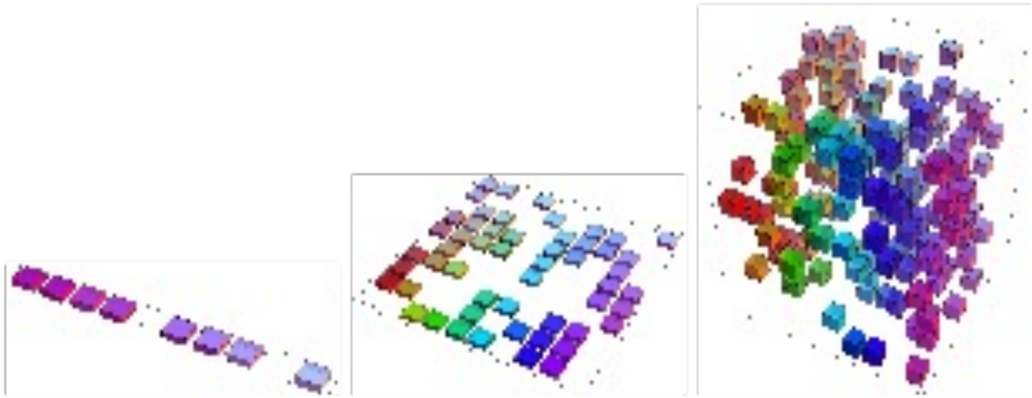


Figure 5.9

remember: Chess vs Go

# Nearest Neighbor

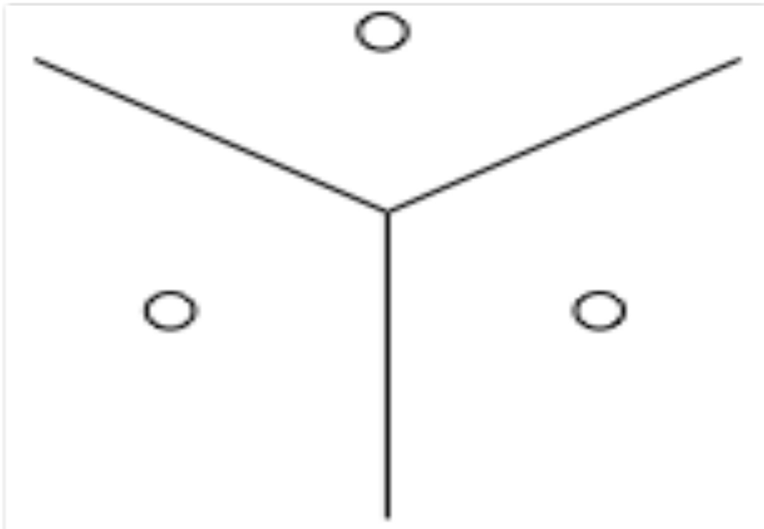


Figure 5.10

# Manifold Learning

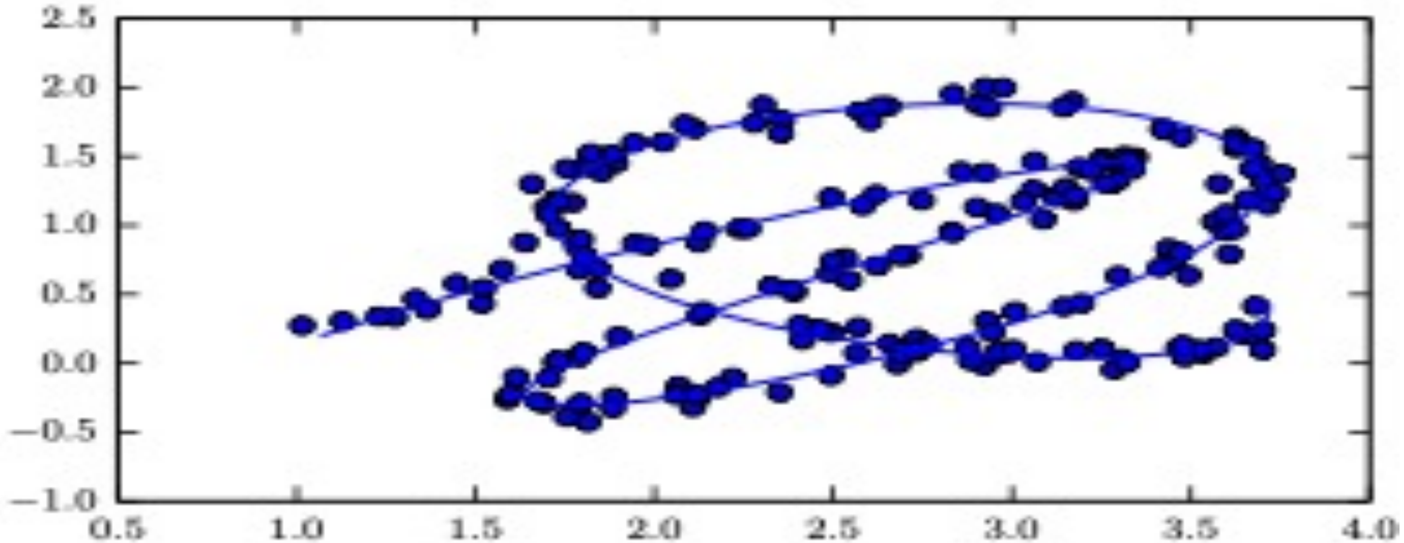



Figure 5.11

# Convolutional Networks

# Convolutional Networks

- Scale up neural networks to process very large images / video sequences
  - Sparse connections
  - Parameter sharing
- Automatically generalize across spatial translations of inputs
- Applicable to any input that is laid out on a grid (1-D, 2-D, 3-D, …)

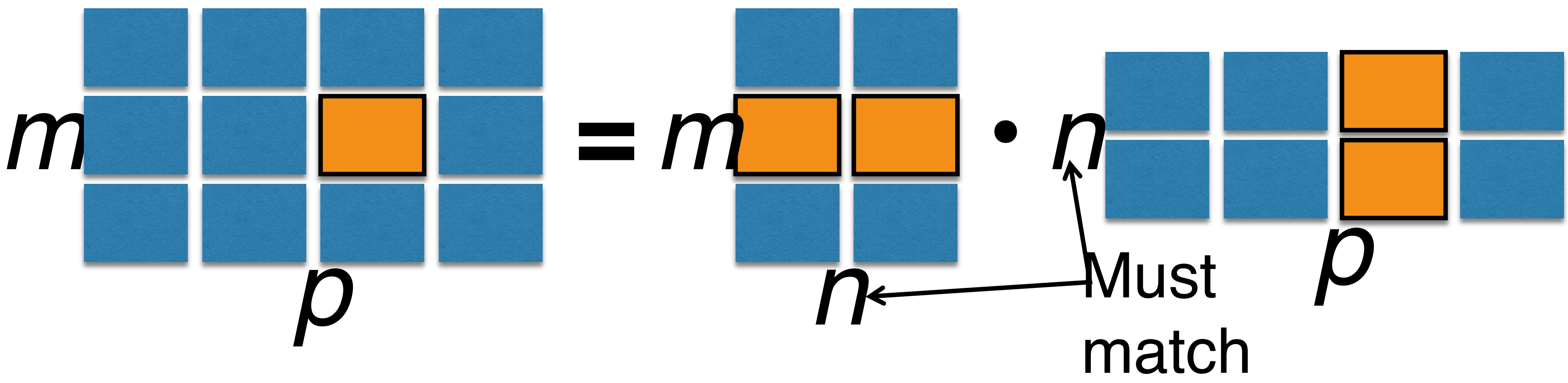
# Key Idea

- <u>Replace matrix multiplication in neural nets with convolution</u>
- Everything else stays the same
    - Maximum likelihood
    - Back-propagation
    - etc.

# Matrix (Dot) Product

$$C = AB. \tag{2.4}$$

$$C_{i,j} = \sum_k A_{i,k} B_{k,j}. \tag{2.5}$$



# 2D Convolution

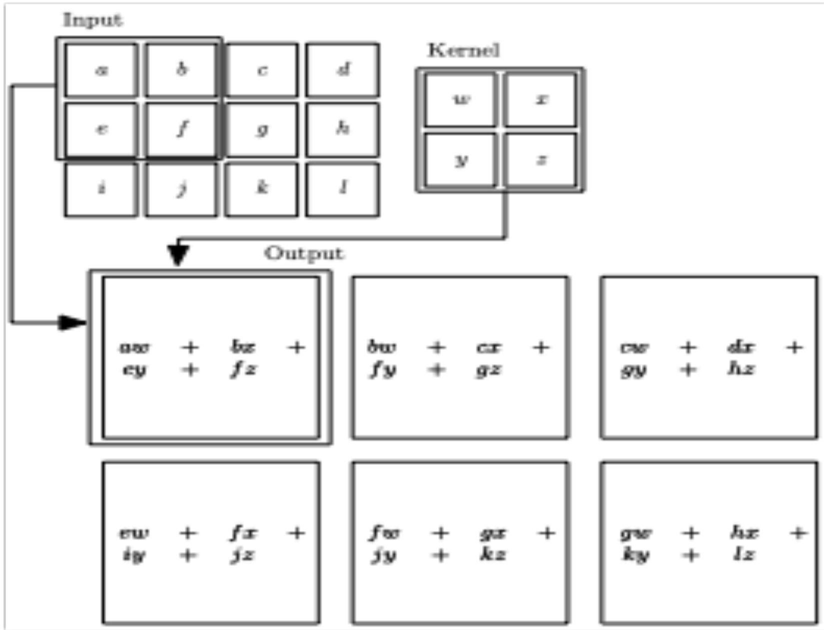



Figure 9.1

# Edge Detection by Convolution

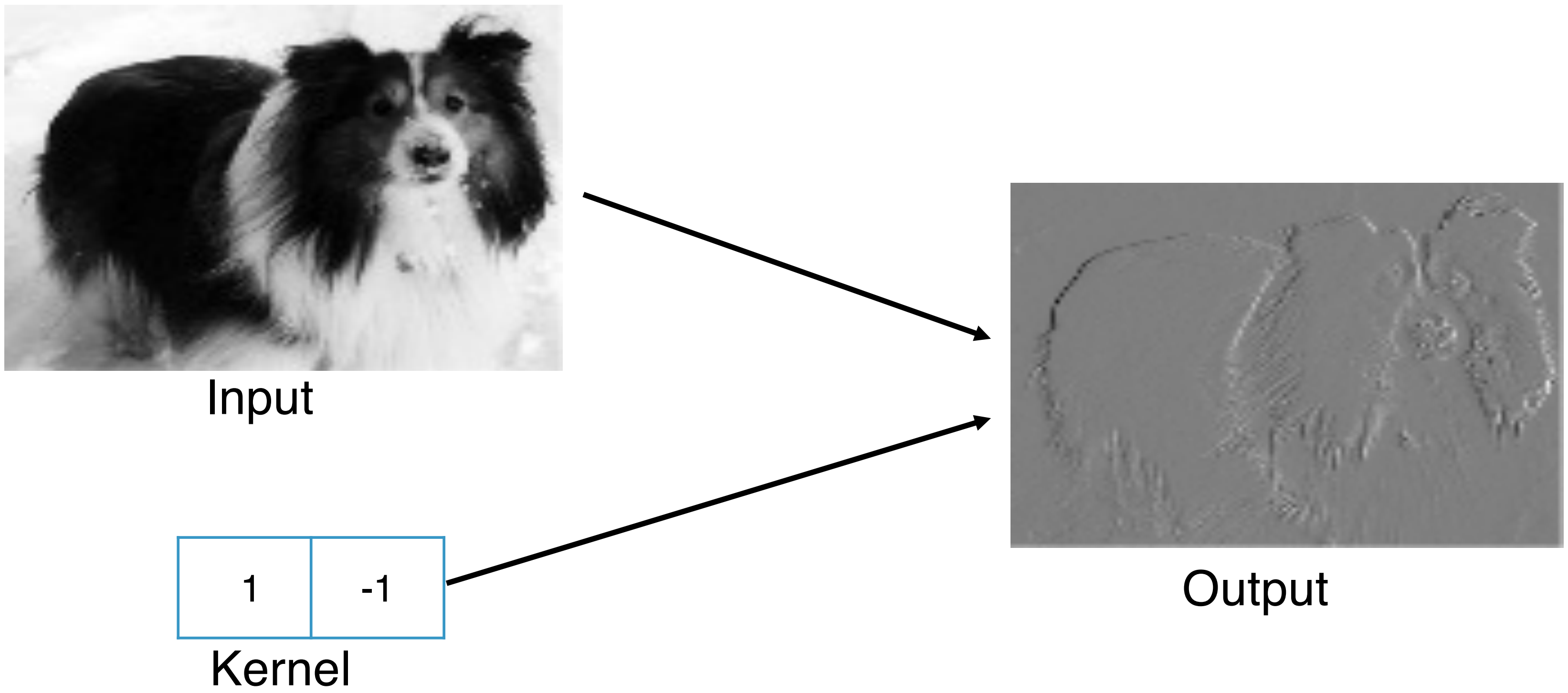
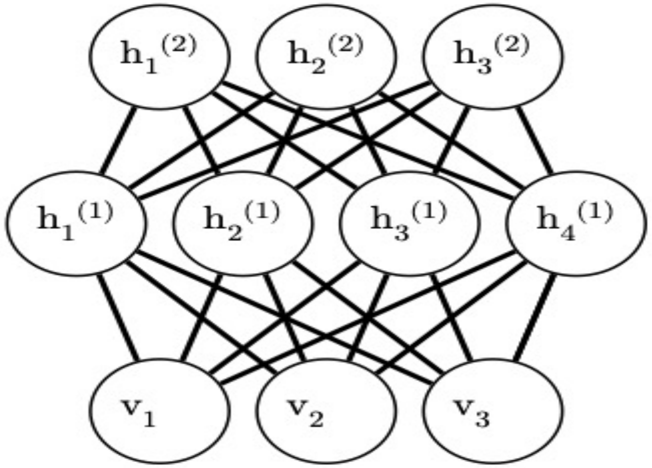



Figure 9.6

# Practical Methodology

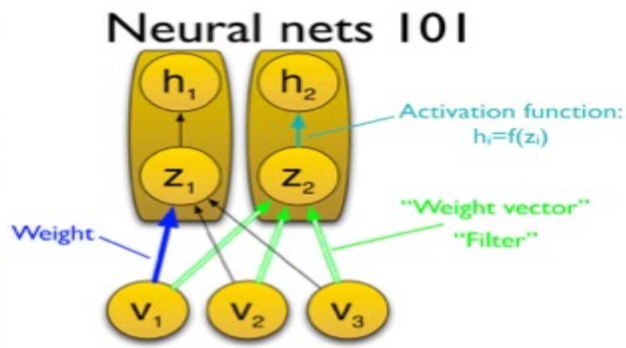
# What drives success in ML?

Arcane knowledge
of dozens of
obscure algorithms?

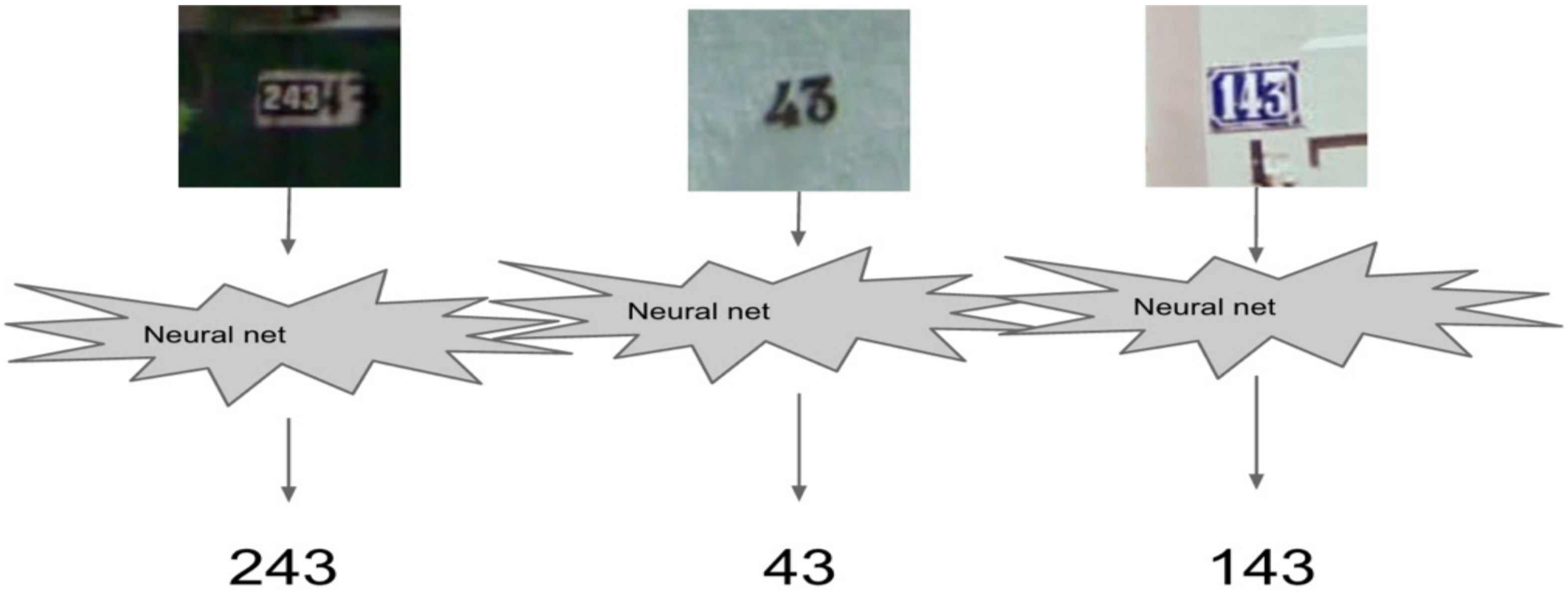



Mountains
of data?

Knowing how
to apply 3-4
standard techniques?



# Example: Street View Address Number Transcription




(Goodfellow et al, 2014)

# Three Step Process

- Use needs to define metric-based goals
- Build an end-to-end system
- Data-driven refinement

# Identify Needs

- High accuracy or low accuracy?
- Surgery robot: high accuracy
- Celebrity look-a-like app: low accuracy

# Choose Metrics

- Accuracy? (% of examples correct)
- Coverage? (% of examples processed)
- Precision? (% of detections that are right)
- Recall? (% of objects detected)
- Amount of error? (For regression problems)

# End-to-end System

- Get up and running ASAP
- Build the simplest viable system first
- What baseline to start with though?
  - Copy state-of-the-art from related publication

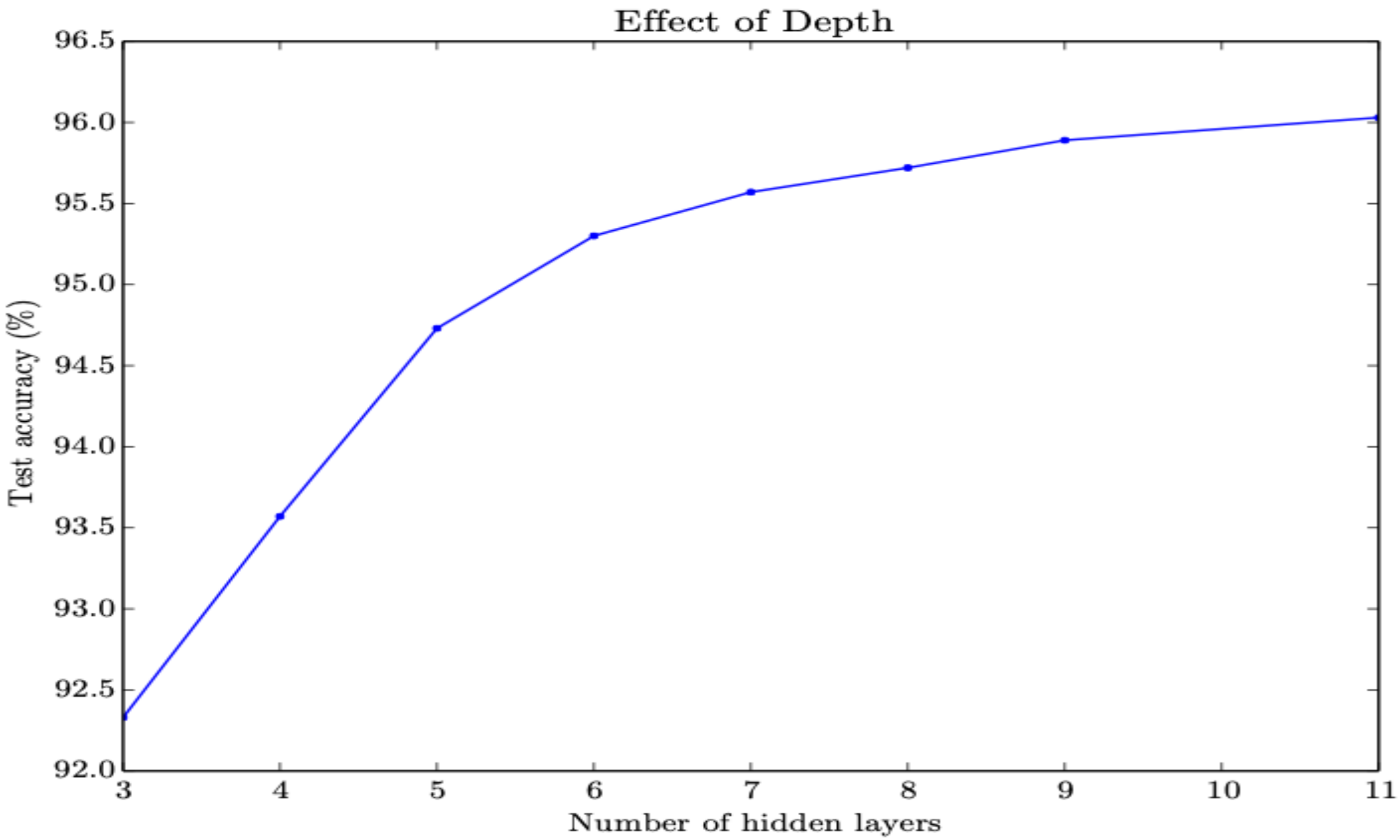
# Deep or Not?

- Lots of noise, little structure -> not deep
- Little noise, complex structure -> deep
- Good shallow baseline:
  - *Use what you know*
  - Logistic regression, SVM, boosted tree are all good

# Choosing Architecture Family

- No structure -> fully connected
- Spatial structure -> convolutional
- Sequential structure -> recurrent

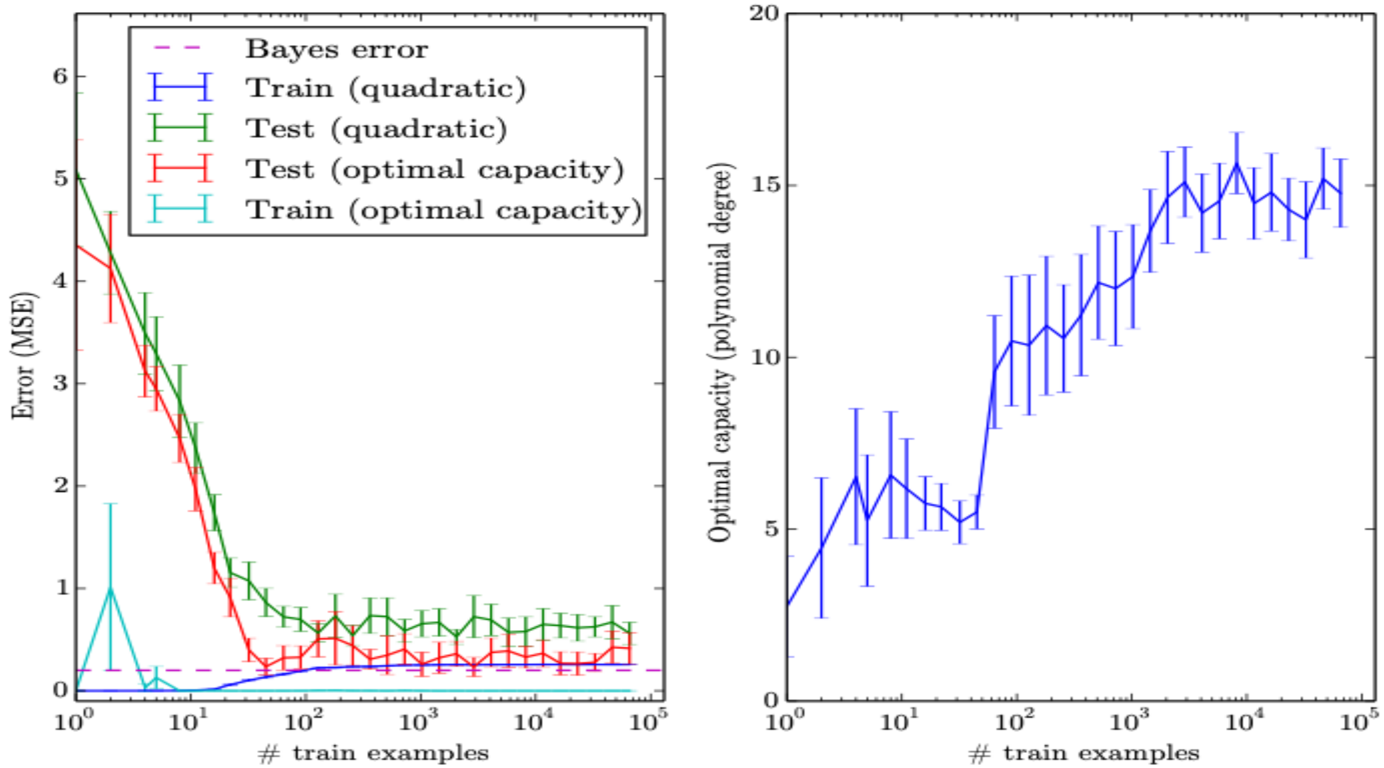
# Increasing Depth



# High Test Error

- Add dataset augmentation
- Add dropout
- Collect more data

# Increasing Training Set Size



# Tuning the Learning Rate

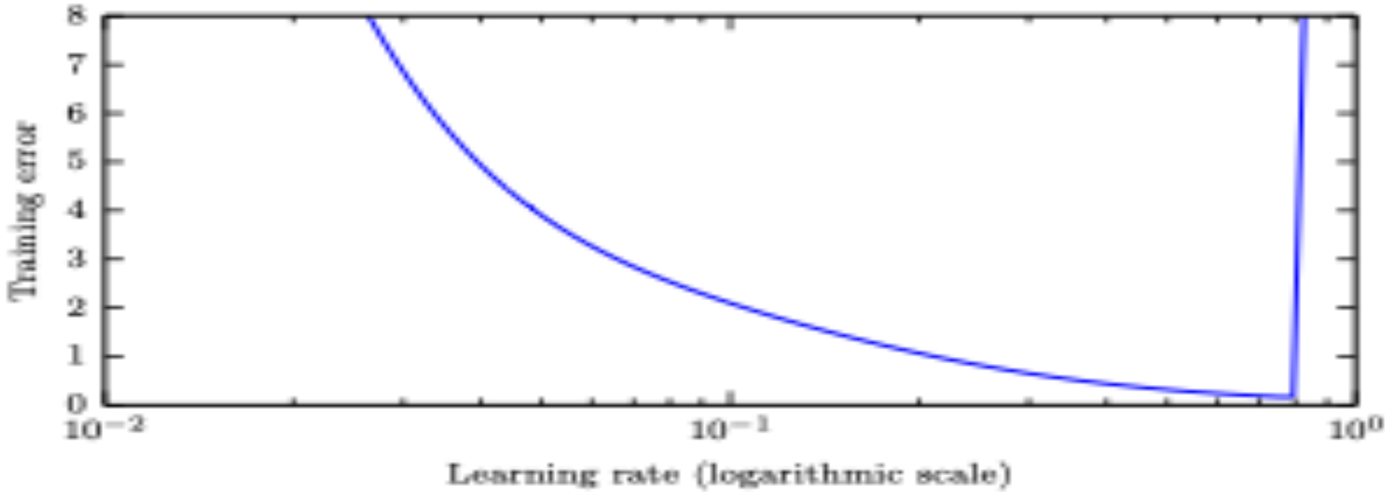



Figure 11.1

# Monte Carlo Methods

# Roadmap

- Basics of Monte Carlo methods
- Importance Sampling
- Markov Chains

# Randomized Algorithms

| | **Las Vegas** | **Monte Carlo** |
|---|---|---|
| **Type of Answer** | Exact | Random amount of error |
| **Runtime** | Random (until answer found) | Chosen by user (longer runtime gives less error) |

# Estimating sums / integrals with samples

$$s = \sum_{\boldsymbol{x}} p(\boldsymbol{x}) f(\boldsymbol{x}) = E_p[f(\mathbf{x})] \tag{17.1}$$

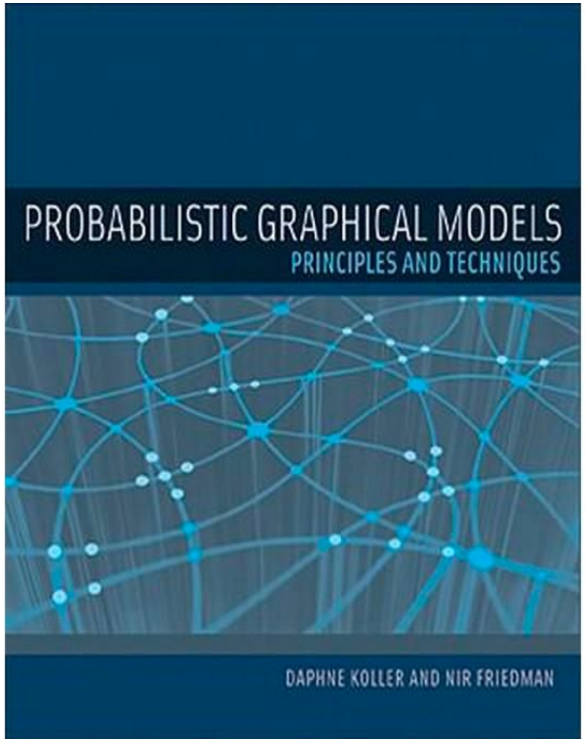
$$s = \int p(\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x} = E_p[f(\mathbf{x})] \tag{17.2}$$

$$\hat{s}_n = \frac{1}{n} \sum_{i=1}^{n} f(\boldsymbol{x}^{(i)}). \tag{17.3}$$

# Justification

- **Unbiased:**
  - **The expected value for finite *n* is equal to the correct value**
  - **The value for any specific *n* samples will have random error, but the errors for different sample sets cancel out**
- **Low variance:**
  - **Variance is $O(1/n)$**
  - **For very large *n*, the error converges "almost surely" to 0**

# For more information…



# Object Categorization

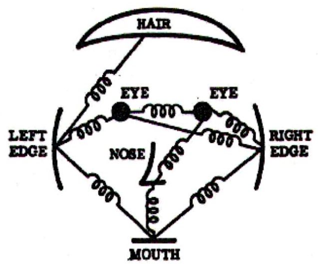
Lecture slides adapted from "Object Categorization an Overview and Two Models”
Fei Fei Li

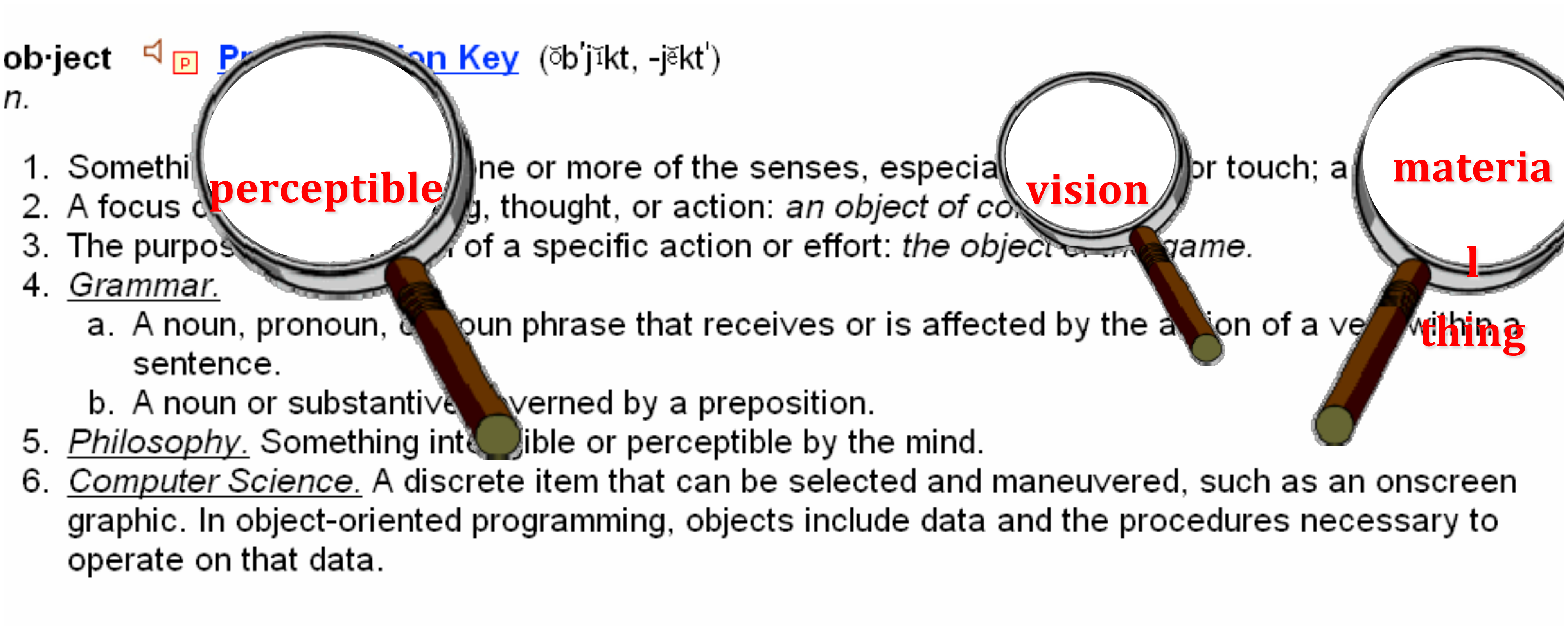
# Agenda

- Introduction to
  “Object Categorization”

- “Bag of Words” models

- Part-based models





**ob·ject** Pr[…]on Key (ŏb′jĭkt, -jĕkt′)
*n.*

1. Somethi[…]ne or more of the senses, especia[…]r touch; a[…]
2. A focus o[…]g, thought, or action: *an object of co[…]*
3. The purpos[…] of a specific action or effort: *the object […]game.*
4. Grammar.
   a. A noun, pronoun, o[…]oun phrase that receives or is affected by the a[…]ion of a ve[…] within a sentence.
   b. A noun or substantive […]verned by a preposition.
5. *Philosophy.* Something int[…]ible or perceptible by the mind.
6. *Computer Science.* A discrete item that can be selected and maneuvered, such as an onscreen graphic. In object-oriented programming, objects include data and the procedures necessary to operate on that data.

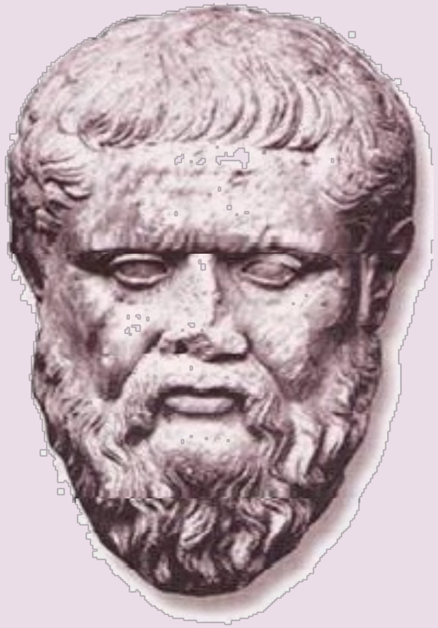
Plato said…

Ordinary objects are classified together if they `participate' in the same abstract Form, such as the Form of a Human or the Form of Quartz.

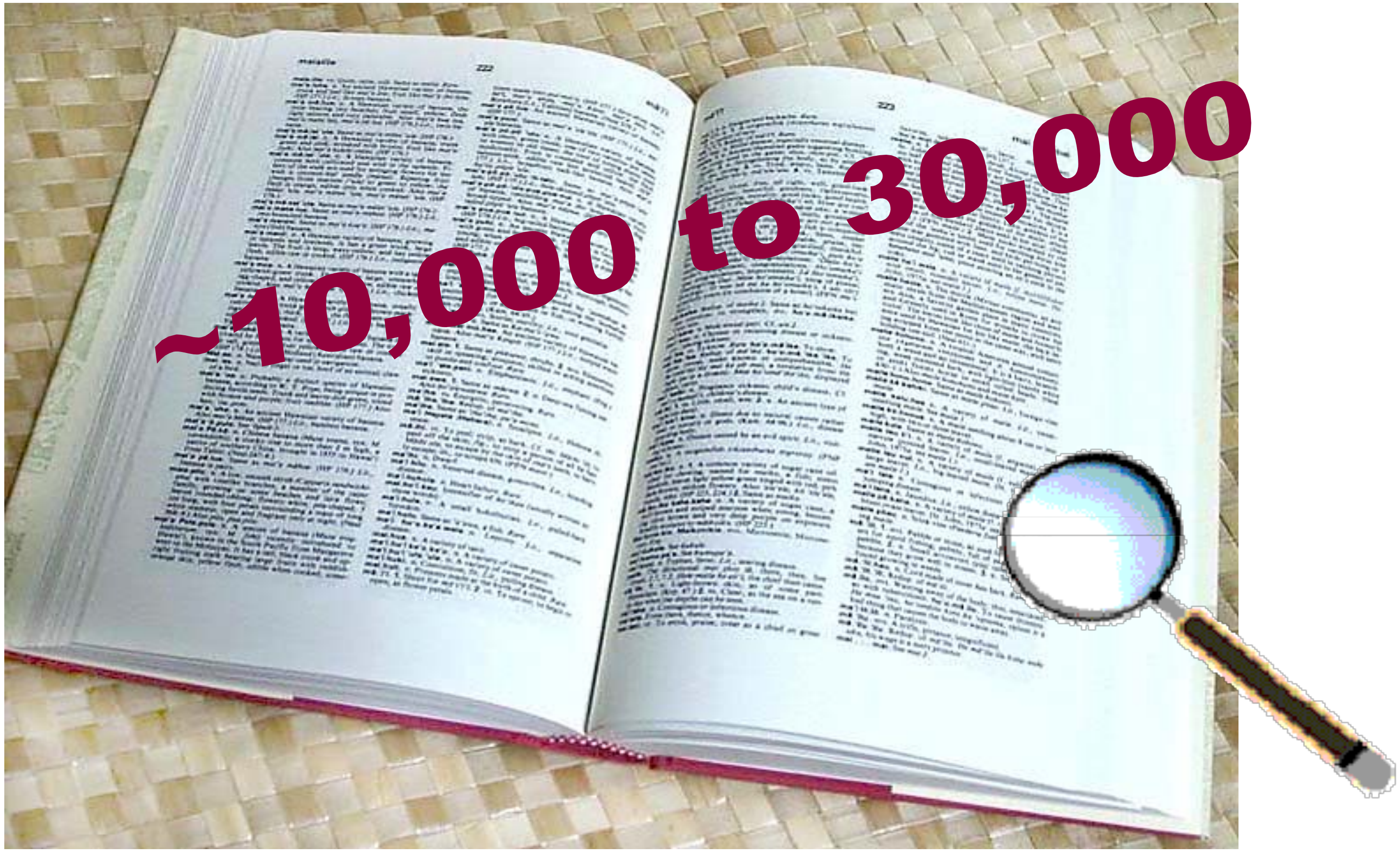
Forms are proper subjects of philosophical investigation, for they have the highest degree of reality.

Ordinary objects, such as humans, trees, and stones, have a lower degree of reality than the Forms.

Fictions, shadows, and the like have a still lower degree of reality than ordinary objects and so are not proper subjects of philosophical enquiry.

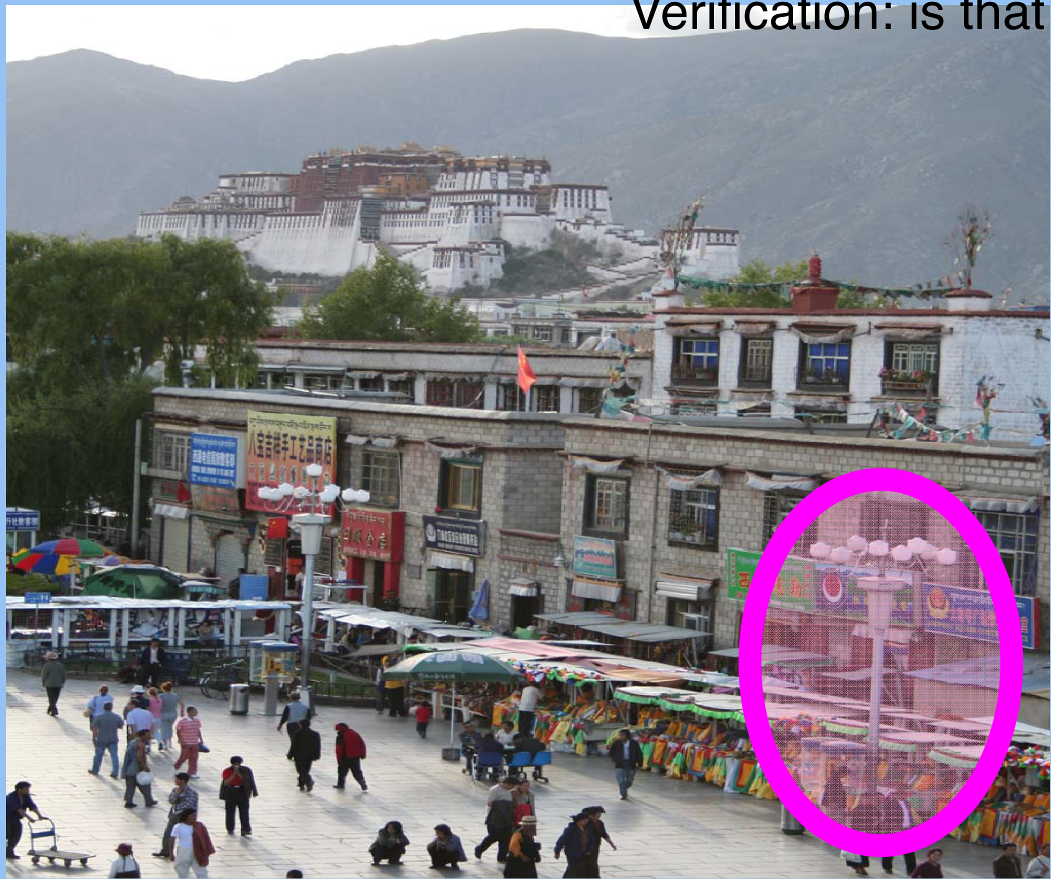


How many object categories are there?



Identification: is that Potala Palace?

Verification: is that a lamp?



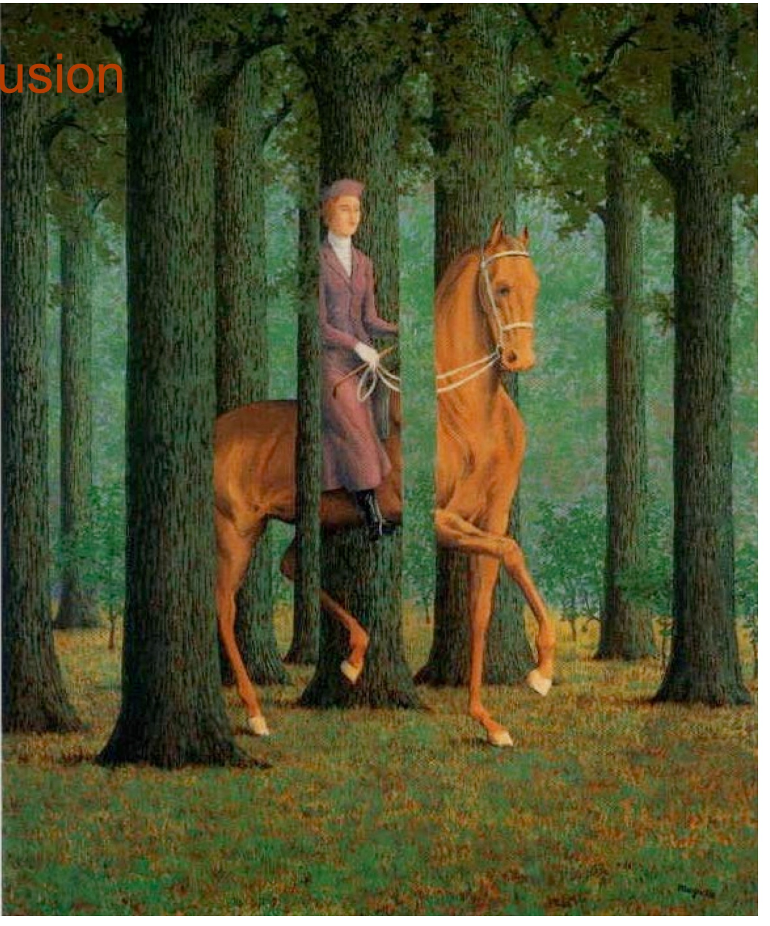
Detection: are there people?

# Object categorization

# Challenges 1: view point variation

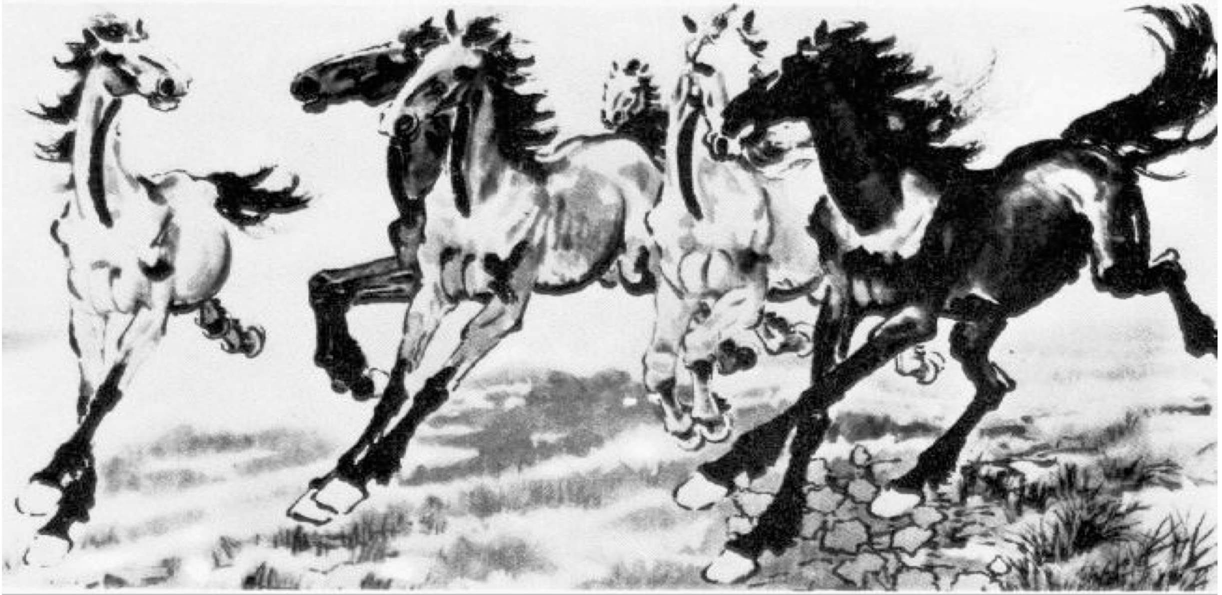
# Challenges 2: illumination

# Challenges 3: occlusion



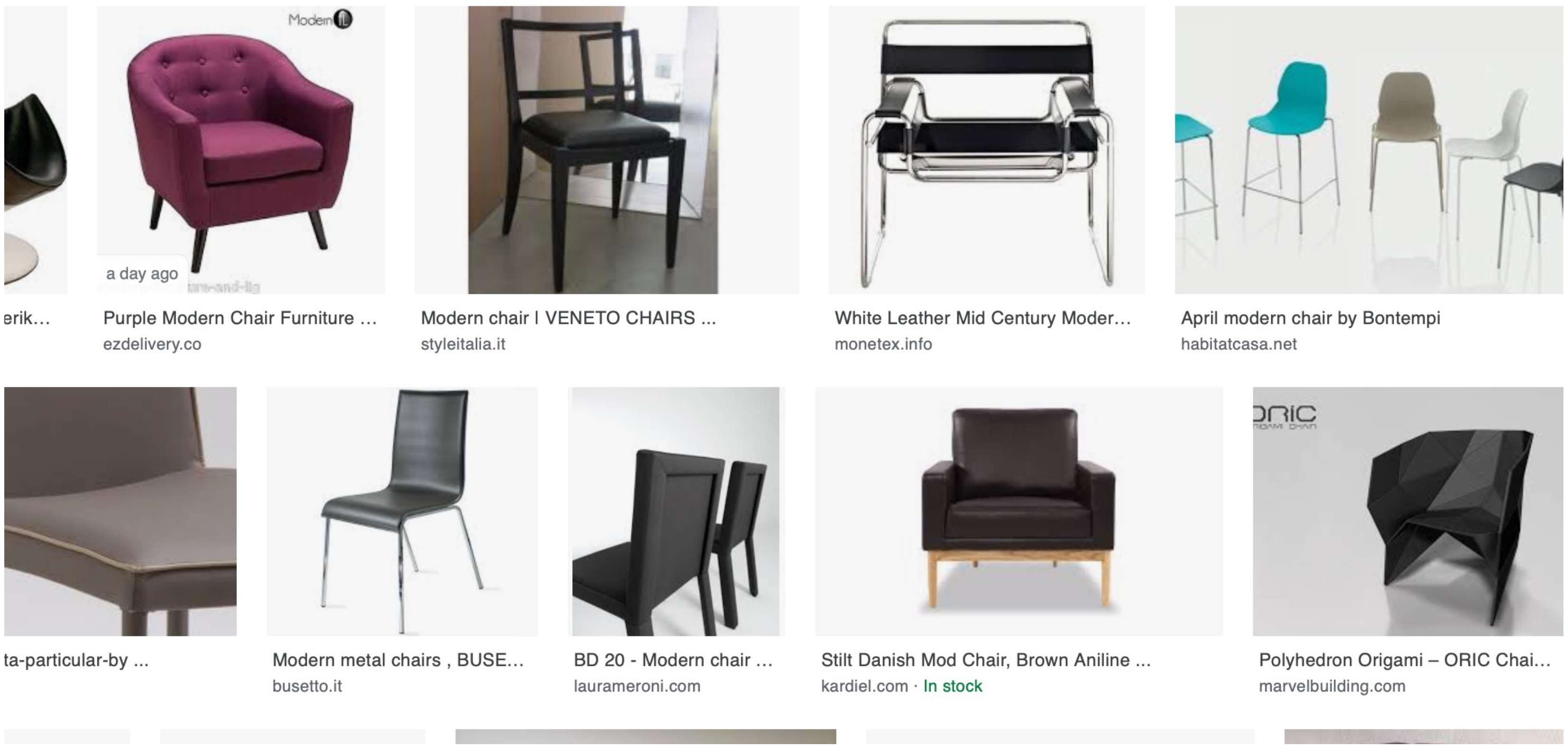
# Challenges 4: scale

# Challenges 5: deformation



# Challenges 6: background clutter

# Challenges 7: intra-class variation




~10,000 to 30,000

**Three main issues**

Representation

How to represent an object category

Learning

How to form the classifier, given training data

Recognition

How the classifier is to be used on novel data

“Bag-of-words” models

Object

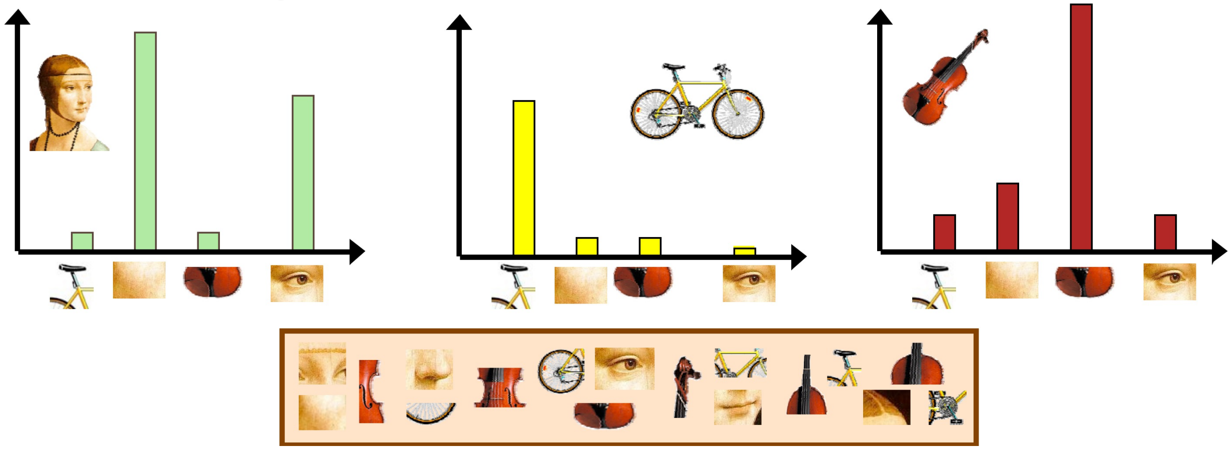
Bag of

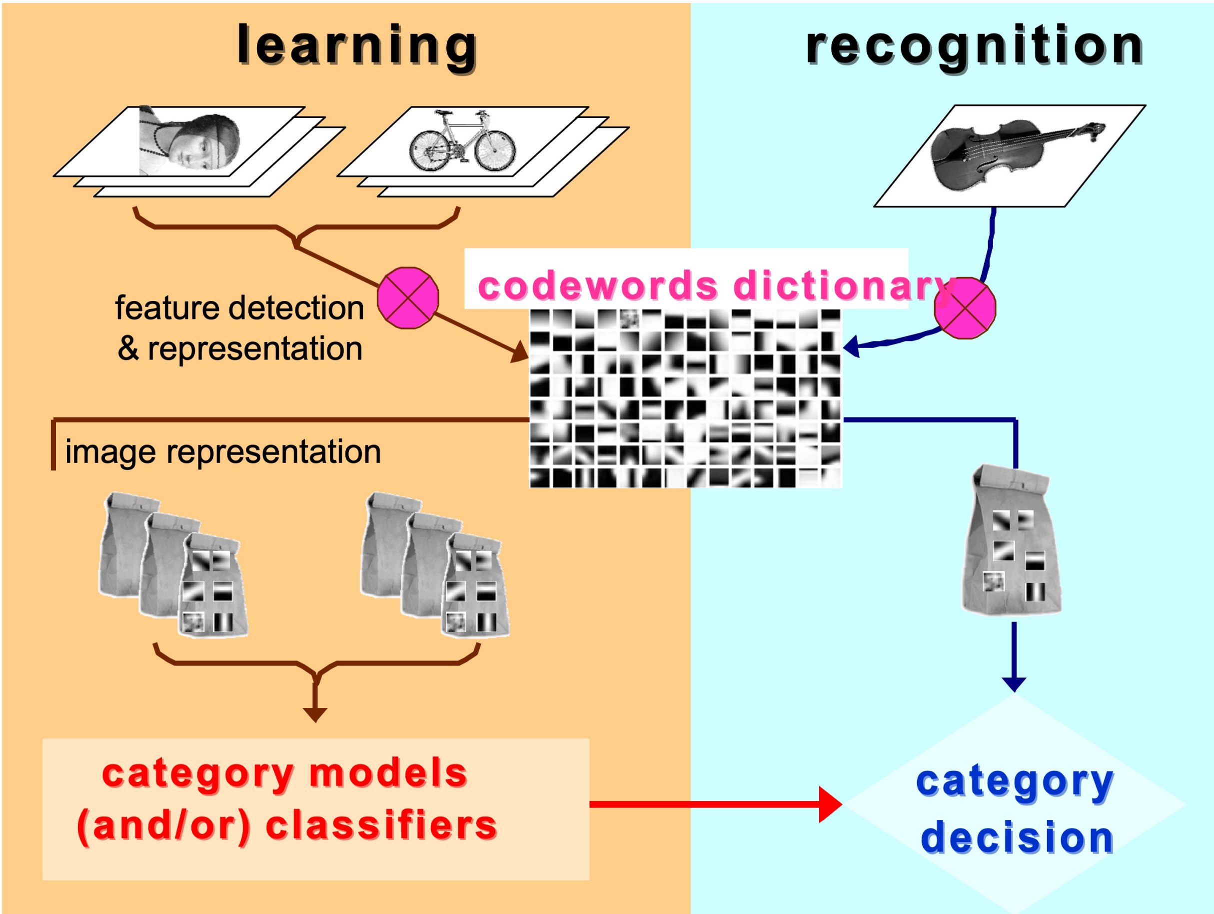
# A clarification: definition of “BoW”

- Looser definition
  - Independent features

# A clarification: definition of "BoW"

- Looser definition
  - Independent features
- Stricter definition
  - Independent features
  - histogram representation




learning
recognition
feature detection
& representation
codewords dictionary
image representation
category models
(and/or) classifiers
category
decision

# Hints that DL … *MUST* WORK

Towards a regularity theory for ReLU networks – chain rule and global error estimates

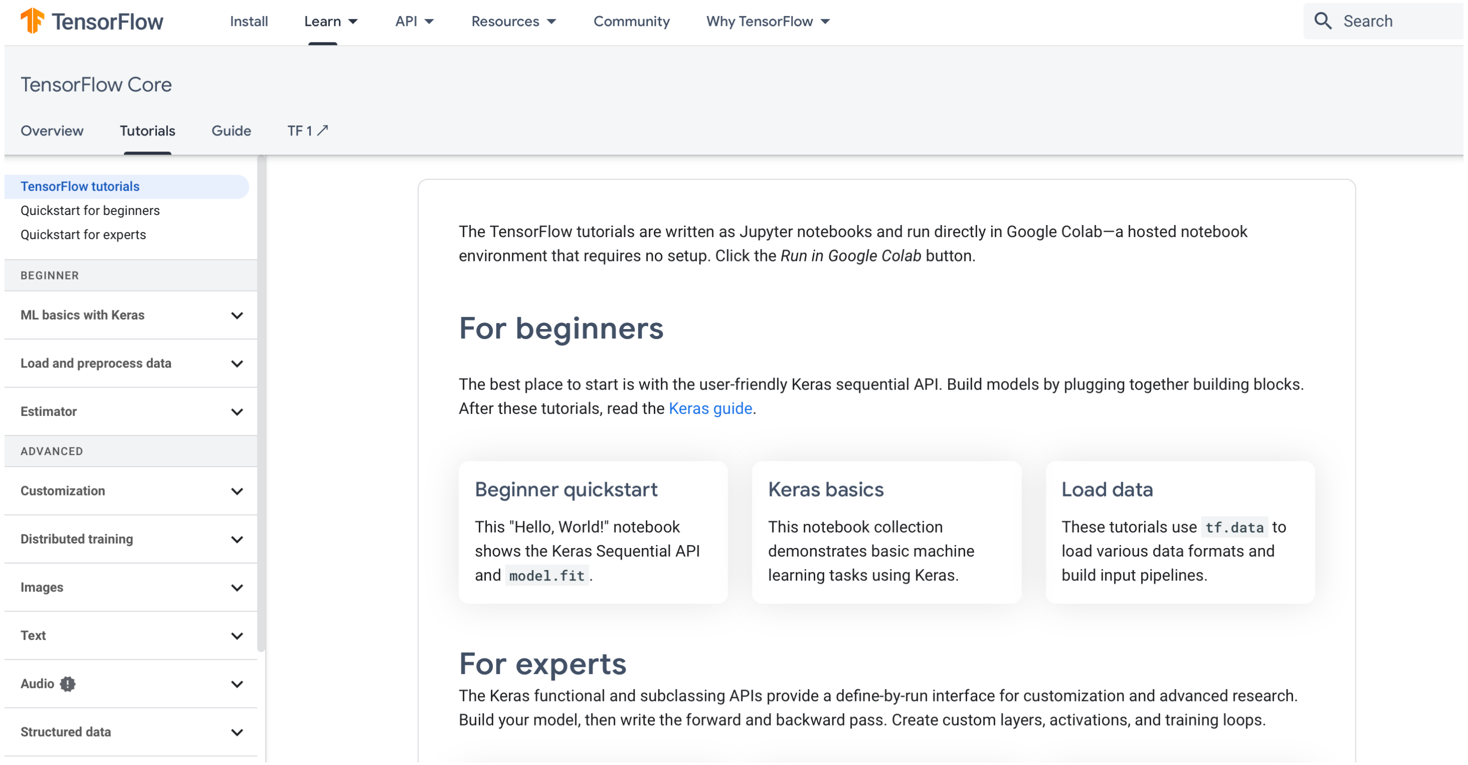
Julius Berner*, Dennis Elbrächter*, Philipp Grohs‡, Arnulf Jentzen§
*Faculty of Mathematics, University of Vienna
Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria
‡Faculty of Mathematics and Research Platform DataScience@UniVienna, University of Vienna
Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria
§Department of Mathematics, ETH Zürich
Rämistrasse 101, 8092 Zürich, Switzerland



***Abstract*—Although for neural networks with locally Lipschitz continuous activation functions the classical derivative exists almost everywhere, the standard chain rule is in general not applicable. We will consider a way of introducing a derivative for neural networks that admits a chain rule, which is both rigorous and easy to work with. In addition we will present a method of converting approximation results on bounded domains to global (pointwise) estimates. This can be used to extend known neural network approximation theory to include the study of regularity properties. Of particular interest is the application to neural networks with ReLU activation function, where it contributes to the understanding of the success of deep learning methods for high-dimensional partial differential equations.**

a way that admits a chain rule which is both rigorous as well as easy to work with. Chain rules for functions which are not everywhere differentiable have been considered in a more general setting in e.g. [16], [17]. We employ the specific structure of neural networks to get stronger results using simpler arguments. In particular it allows for a stability result, i.e. Lemma III.3, the application of which will be discussed in Section V. We would also like to mention a very recent work [18] about approximation in Sobolev norms, where they deal with the issue by using a general bound for the Sobolev norm of the composition of functions from the Sobolev space $W^{1,\infty}$.

# Try Deep Learning by yourself!




https://www.tensorflow.org/tutorials

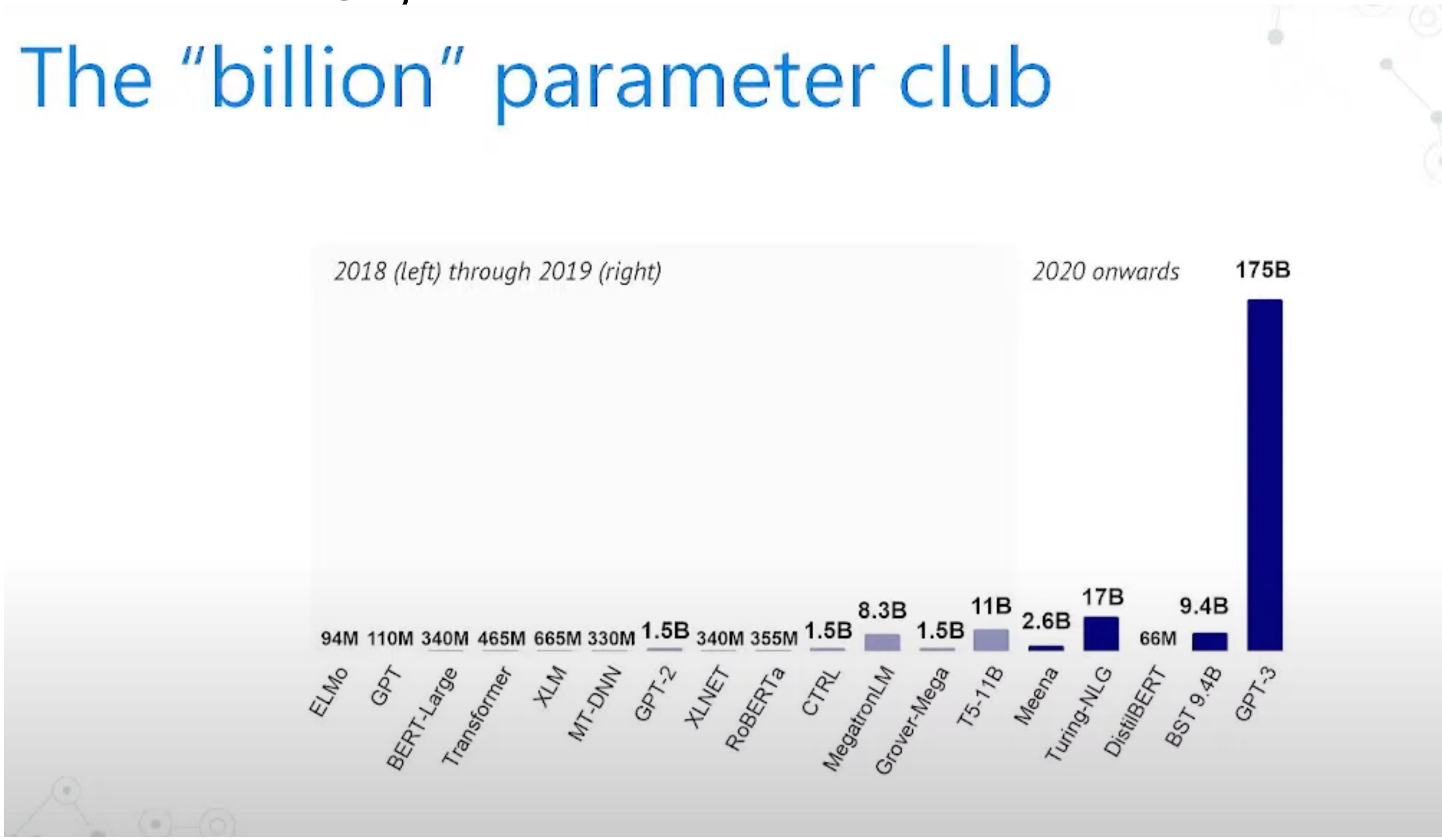
# Well,…

## State of AI – some considerations

- The "billion" parameters club → how large, up to 175B these days!
- Cost (more later), about $1 per 1000 parameters
- Interesting: outrageous cost for incremental improvement
  - <u>Need research and theory</u>
  - We can be <u>more efficient</u> in training algorithms
- Large models are driven by efficiency with small data
  - Sometimes… with transfer learning
- Power-law → parameter & computational power do not scale linearly (which is bad!)

(stolen from Giorgio Metta)

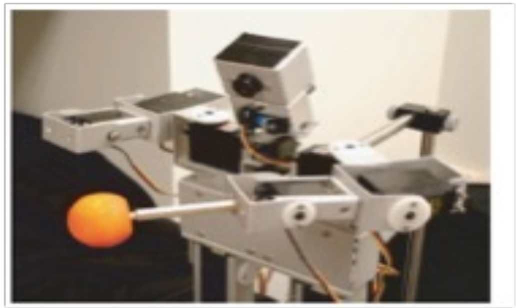
**Well,…**

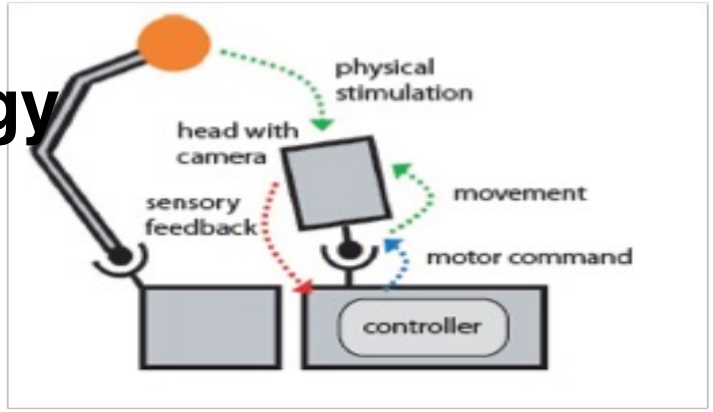


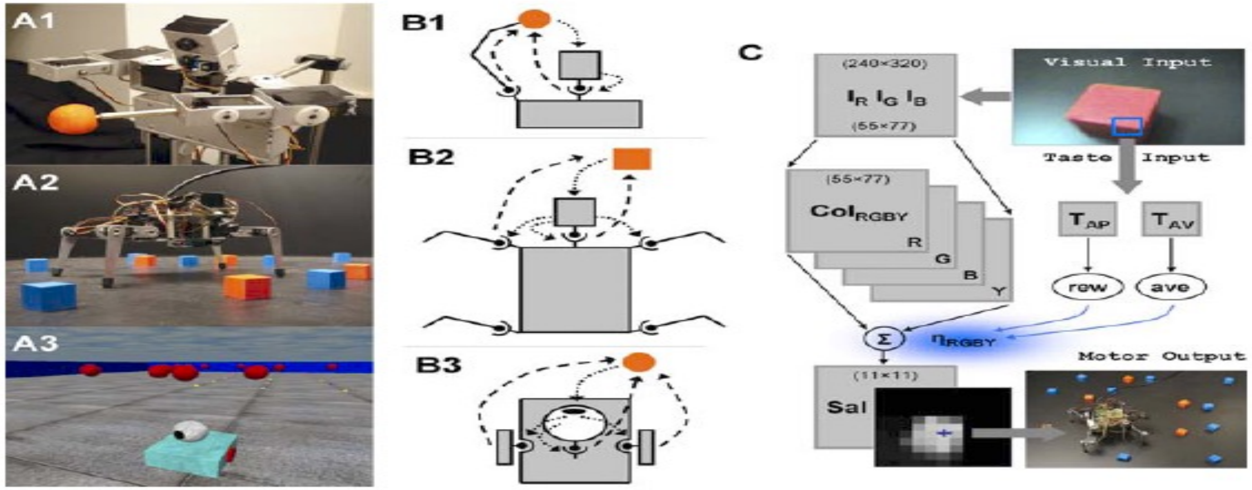

(stolen from Giorgio Metta)

# Looking for new paths forward...
# For example: Information self-structuring

- Experiments:
- Lungarella and Sporns, **2006 Mapping information flow in sensorimotor networks PLoS Computational Biology**








Lungarella, Sporns (2006)

**Figure 1.** Robots, Sensorimotor Interactions, and Neural Control Architecture

(A1) *Roboto* has a total of 14 DOF, five of which are used in the current set of experiments. Note the head-mounted CCD camera, the pan-tilt head system (2 DOF), and the moveable left arm with shoulder, elbow, and wrist joints (3 DOF). The object is a red ball (1.25 inches diameter) attached to the tip of the last joint.

(A2) *Strider* has a total of 14 DOF, with four legs of 3 DOF each and 2 DOF in the pan-tilt head system. Objects are red and blue blocks (1 inch cubes). *Strider* is situated in an environmental enclosure with black walls.

(A3) *Madame* has 4 DOF, with 2 DOF in the pan-tilt system and 2 DOF for the wheels, which are both located on an axis vertical to the main body axis. The environment is a square arena bounded by blue walls containing 20 red-colored floating spheres.

(B1) *Roboto* engages in sensorimotor interactions via the head system and arm movements; sensory → motor (dotted arrows), motor → sensory (dashed arrows).

(B2) *Strider* engages in sensorimotor interactions via the head system, as well as via steering signals generated by the head and transmitted to the four legs.

(B3) *Madame*'s behavior consists of a series of approaches to colored objects and ovations. Fixations to the objects are maintained by independent action of head and body.

(C) Neural control architecture. The architecture common to all robots is composed of color image arrays $I_R$, $I_G$, $I_B$, color- intensity map $\mathbf{Col}_{RGBY}$, and saliency map *Sal* (see text for details). The peak of the saliency map (blue cross) determines the pan-tilt camera motion and body steering. In addition, *Strider*'s neural system contains a value system with taste sensory inputs relayed via a virtual taste sensor (blue square in visual image) to taste neurons ($T_{AP,AV}$), which in turn generates reward and aversiveness signals (rew, ave). These signals are used to modulate the strengths of the saliency factors $\eta_{RGBY}$ (see text for details).

DOI: 10.1371/journal.pcbi.0020144.g001

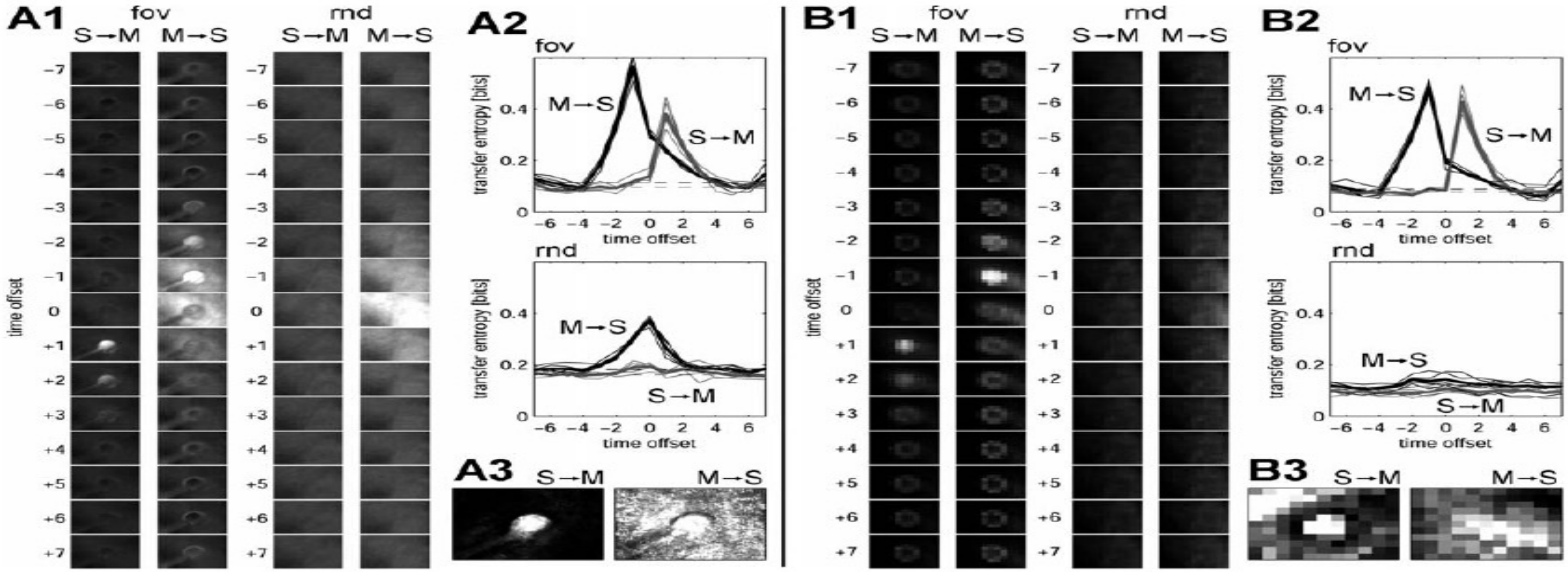



**Figure 3.** Information Flow (Transfer Entropy) between Sensory Input, Neural Representation of Saliency, and Motor Variables in *Roboto*

(A1) Transfer entropy between array $I_R$ (variable S) and pan-tilt amplitude (variable M). Series of plots show maps of transfer entropy from S to M (S $\rightarrow$ M) and from M to S (M $\rightarrow$ S) over visual space ($55 \times 77$ pixels), calculated for offsets between $-7$ ("M leading S") and $+7$ ("S leading M") time steps. Plots show data for conditions "fov" and "rnd." The gray scale ranges from 0.0 to 0.5 bits (for all plots in panels A1 and B1).

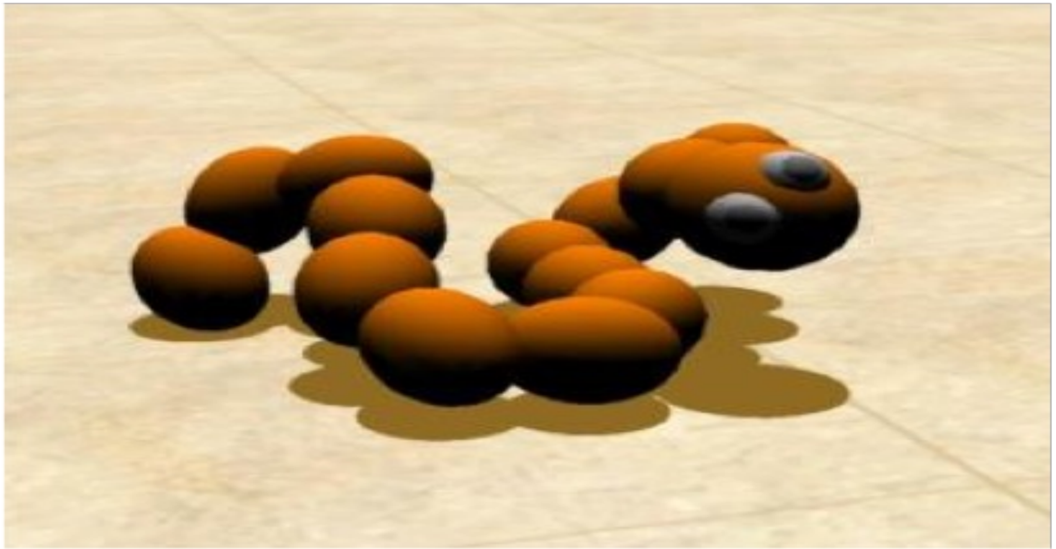
(A2) Curves show transfer entropy for five individual runs (thin lines) as well as the average over five runs (thick lines) between the single central pixel of array $I_R$ (S) and pan-tilt amplitude (M), for directions M $\rightarrow$ S (black) and S $\rightarrow$ M (gray).

(A3) z-Score maps of significant image regions (plotted between $z = 0$ and $z = 6$). The z-scores are expressed as number of standard deviations above background at time offset $+1$ (S $\rightarrow$ M) and $-1$ (M $\rightarrow$ S). Mean and standard deviation of background is calculated from transfer entropy values at maximal time delays ($-7$,$+7$ time steps).

(B) All three panels have the same format as (A), but the neural activations of the saliency map *Sal* are substituted as variable S ($11 \times 11$ neural units).

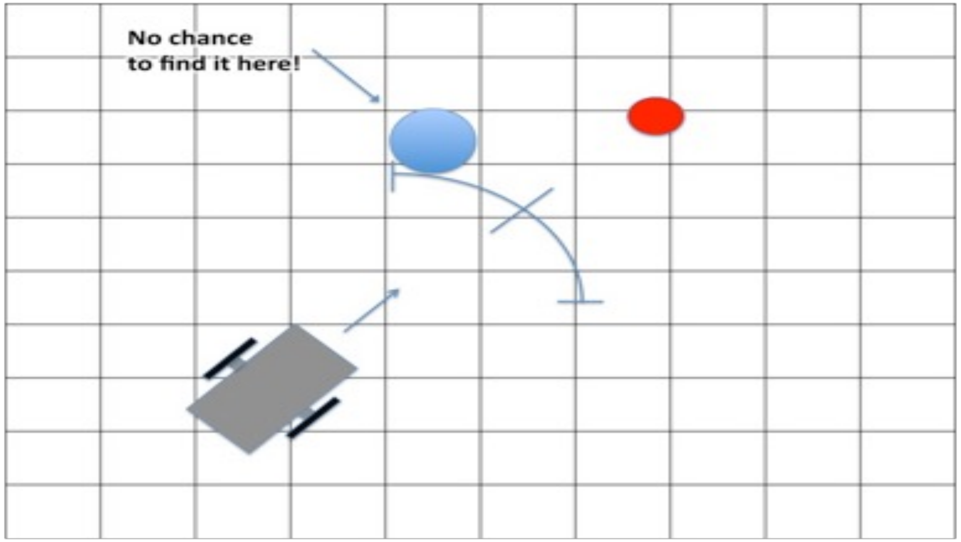
DOI: 10.1371/journal.pcbi.0020144.g003

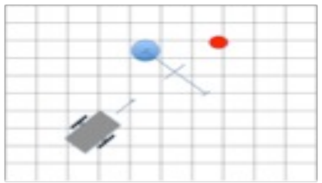
# Snakebot



see: **Tanev et. al, IEEE TRO, 2005**

# Maybe not GOF Euclidean space? :-)



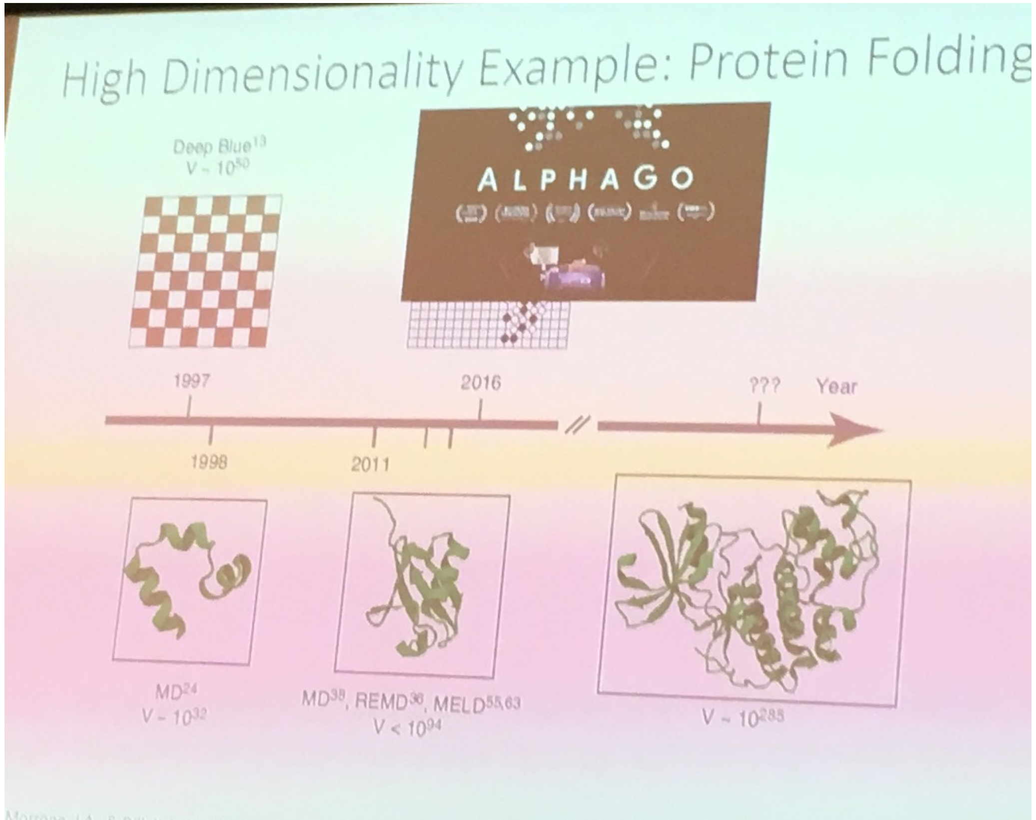



see: **Bonsignorio, Artificial Life, 2013**

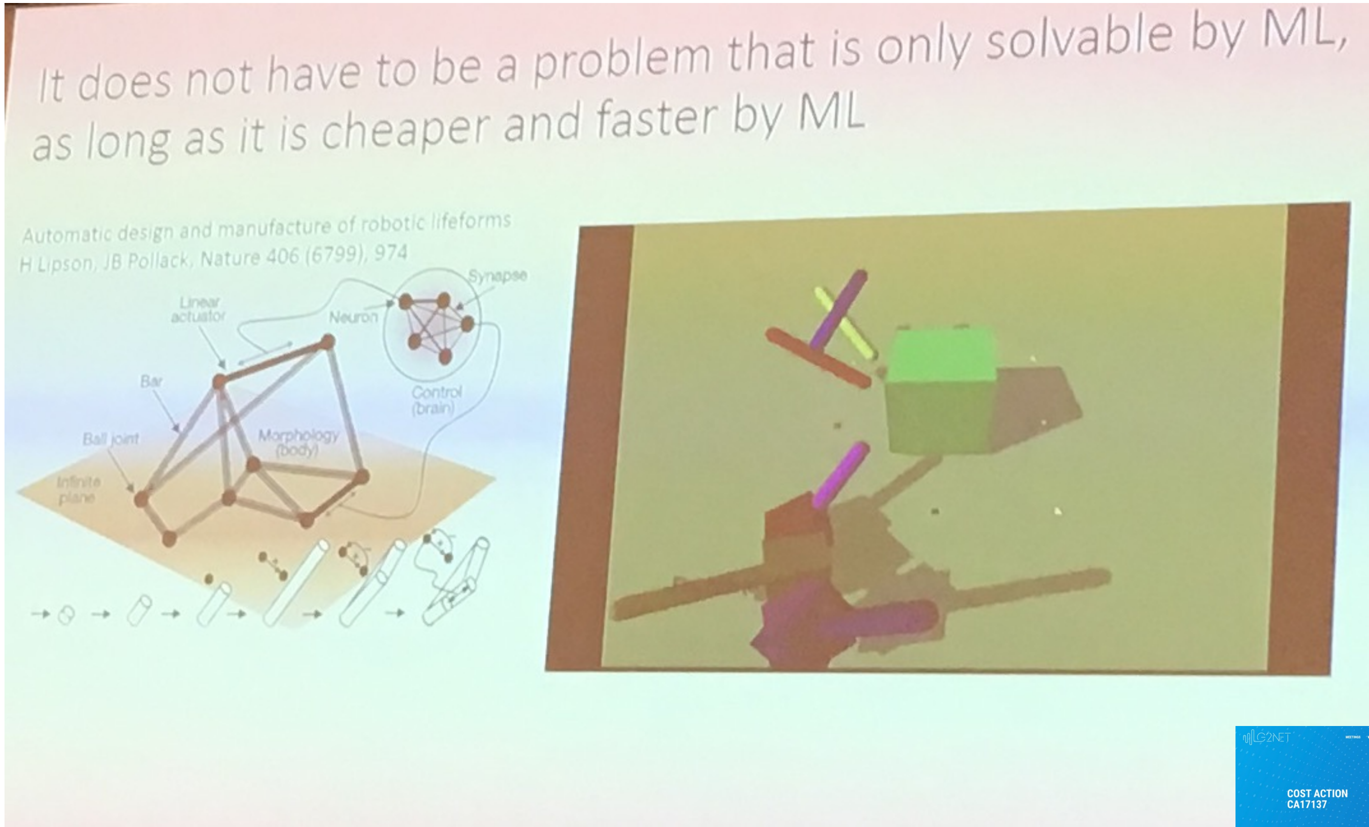
# Bottom Line: Physics Matters!

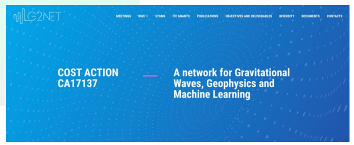
**Coping with the common underlying theoretical issues implied by the application of ML and DL to physical systems might have deep and wide scientific and technological impact**

# Bottom Line: Physics Matters!



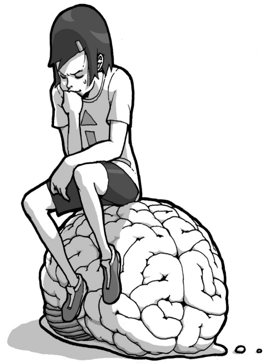
# Bottom Line: Physics Matters!



# Two views of intelligence

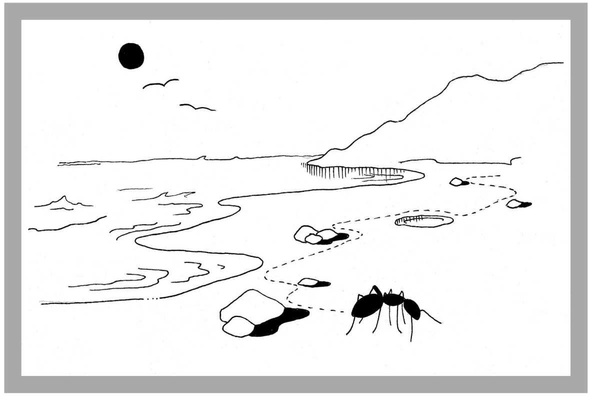
classical:
**cognition as computation**



embodiment:
**cognition emergent from sensory-motor and interaction processes**

# “Frame-of-reference”
# Simon’s ant on the beach

- **simple behavioral rules**
- **complexity in interaction, not — necessarily — in brain**



thought experiment:
**increase body by factor of 1000**
**everything else the same**

# The “symbol grounding” problem

**real world: doesn’t come with labels …**

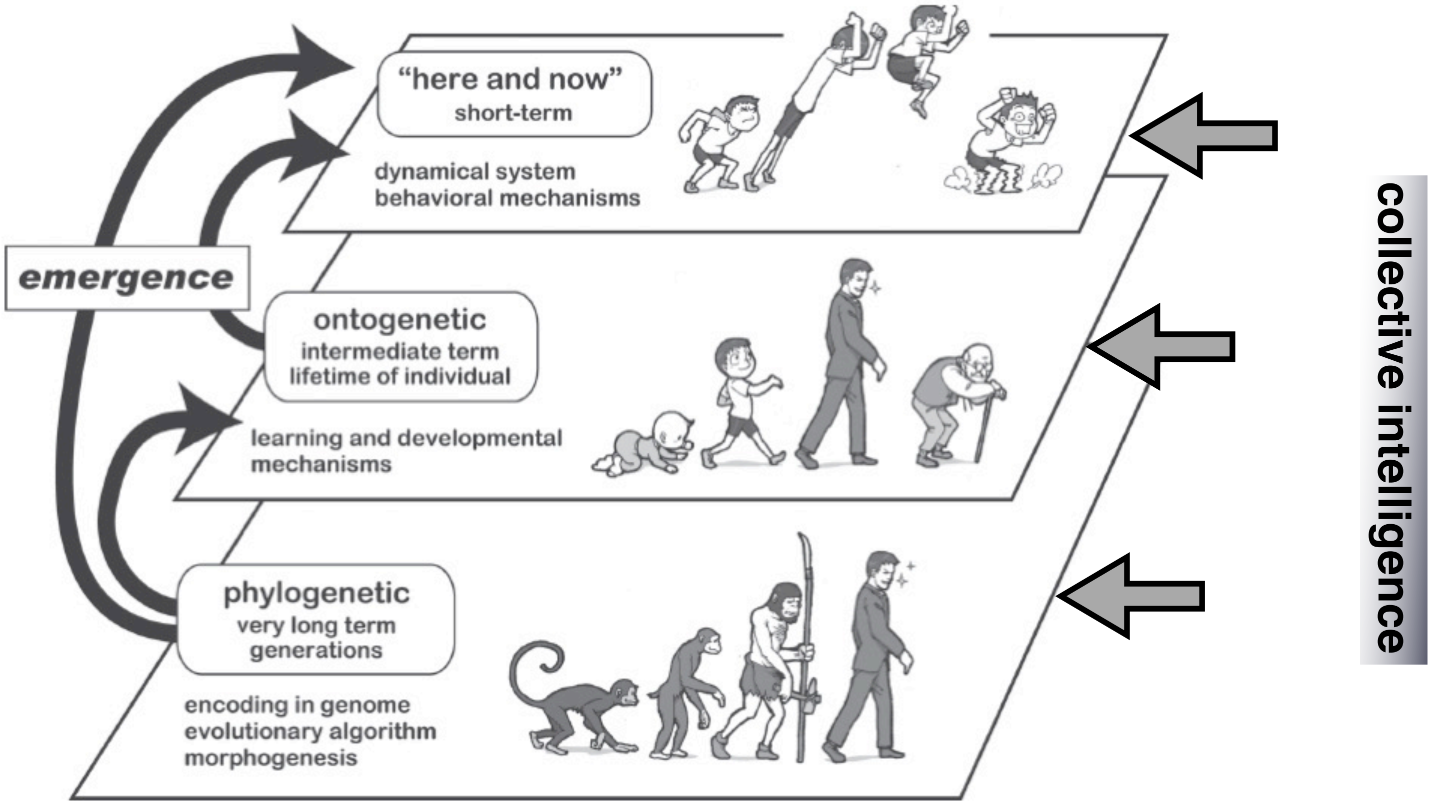
**How to put the labels??**


Gary Larson

# Self-organization and emergence at many levels

- molecules
- cells
- organs
- individuals
- groups of individuals

# Time perspectives



# Today's Guest Lecture

**10:10 Yulia Sandamirskaya,**

**Intel Germany, Munich, Germany**

**«Neuromorphic Computing»**

**Stay tuned!**

# Assignments for next week

- Check “How the body…” for self-study
- Think about how … deep learning…vs. evolutionary programming… vs … symbolic approaches…play with tensorflow (or similar)STUDENTS’ PREZ ALWAYS WELCOME.

# End of lecture 5

## Thank you for your attention!

## stay tuned for lecture 6

## “Embodied Intelligence: principles and open issues»

# Short Bio

## The ShanghAI Lectures 2013-

Prof. Fabio Bonsignorio is **ERA Chair in AI for Robotics** at FER, University of Zagreb, Croatia. He is **Founder and CEO of Heron Robots (advanced robotics solutions),** see **www.heronrobots.com**. He has been visiting professor at the **Biorobotic Institute of the Scuola Superiore Sant'Anna in Pisa from 2014 to 2019**. He has been a professor in the Department of System Enginering and Automation at the **University Carlos III of Madrid until 2014**. In 2009 he got the **Santander Chair of Excellence in Robotics** at the same university. He has been working for some 20 years in the high tech industry before joining the research community.

He is a **pioneer and has introduced the topic of Reproducibility of results in Robotics and AI**. He is a **pioneer in the application of the blockchain to robotics and AI (smart cities, smart land, smart logistics, circular economy). He coordinates the Topic Group of euRobotics about Experiment Replication, Benchmarking, Challenges and Competitions. He is co-chair of the IEEE Robotics & Automation Society (RAS) Technical Committee, TC-PEBRAS (PErformance and Benchmarking of Robotics and Autonomous Systems)**.

He is a **Distinguished Lecturer** for **IEEE Robotics and Automation Society,** Senior Member of IEEE and member of the Order of the Engineers of Genoa, Italy.

He coordinates the task force robotics, in the G2net,an EU network studying the application of **Machine Learning and Deep Learning (Apprendimento Profondo) to Gravitational wave research, la Geophysics and Robotics**.

**Has given invited seminars and talks in many places: MIT Media Lab, Max Planck Institute, Imperial College, Politecnico di Milano in Shenzhen, London, Madrid, Warsaw, San Petersburg, Seoul, Rio Grande do Sul.…**

# Thank you!

fabio.bonsignorio@fer.hr
fabio.bonsignorio@heronrobots.com
fabio.bonsignorio@gmail.com

University of Zagreb
Faculty of Electrical Engineering and Computing
Laboratory for Autonomous Systems and Mobile Robotics

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No. 952275

www.heronrobots.com