Artificial Intelligence

Lecturer: Professor Pekka Toivanen

E-mail: Pekka.Toivanen@uef.fi

Tel: +358 40 543 9021

Exercises: Taha Nakabi and Saku ?

E-mail: tahanak@student.uef.fi

Artificial Intelligence

Lectures:

09.09.19	ma 12.00-14.00 F213
11.09.19	ke 12.00-14.00 F213
12.09.19	to 12.00-14.00 F213
18.09.19	ke 10.00-12.00 F213
19.09.19	to 10.00-12.00 F213
25.09.19	ke 12.00-14.00 F213
26.09.19	to 10.00-12.00 F213
02.10.19	ke 10.00-12.00 F213
03.10.19	to 12.00-14.00 F213
08.10.19	ti 14.00-16.00 F213
09.10.19	ke 12.00-14.00 F213
10.10.19	to 12.00-14.00 F213
15.10.19	ti 14.00-16.00 F213
16.10.19	ke 12.00-14.00 F213
17.10.19	to 12.00-14.00 F213
22.10.19	ti 12.00-14.00 F213
23.10.19	ke 12.00-14.00 F213

09.09.19	ma 12.00-14.00 TB180
11.09.19	ke 12.00-14.00 TB178
12.09.19	to 12.00-14.00 TB180
18.09.19	ke 10.00-12.00 TB180
19.09.19	to 10.00-12.00 TB180
25.09.19	ke 12.00-14.00 TB178
26.09.19	to 10.00-12.00 TB180
02.10.19	ke 10.00-12.00 TB180
03.10.19	to 12.00-14.00 TB180
08.10.19	ti 14.00-16.00 TB180
09.10.19	ke 12.00-14.00 TB180
10.10.19	to 12.00-14.00 TB180
15.10.19	ti 14.00-16.00 TB180
16.10.19	ke 12.00-14.00 TB178
17.10.19	to 12.00-14.00 TB180
22.10.19	ti 12.00-14.00 TB180
23.10.19	ke 12.00-14.00 TB180

Artificial Intelligence

Exercises:

Exercises:

13.09.19	pe 08.00-10.00	E24+25
17.09.19	ti 12.00-14.00	E24+25
24.09.19	ti 12.00-14.00	E24+25
30.09.19	ma 12.00-14.00	E24+25
08.10.19	ti 12.00-14.00	E24+25
15.10.19	ti 12.00-14.00	E24+25
21.10.19	ma 14.00-16.00	E24+25
24.10.19	to 12.00-14.00	E24+25

Kuopio: Taha Nakabi

12.09.19	to 10.00-12.00	TB247
17.09.19	ti 12.00-14.00	TB247
24.09.19	ti 12.00-14.00	TB247
01.10.19	ti 12.00-14.00	TB247
08.10.19	ti 12.00-14.00	TB247
15.10.19	ti 12.00-14.00	TB247
21.10.19	ma 14.00-16.00	TB247
24.10.19	to 12.00-14.00	TB178

Joensuu: Saku Kukkonen

Home assignment:

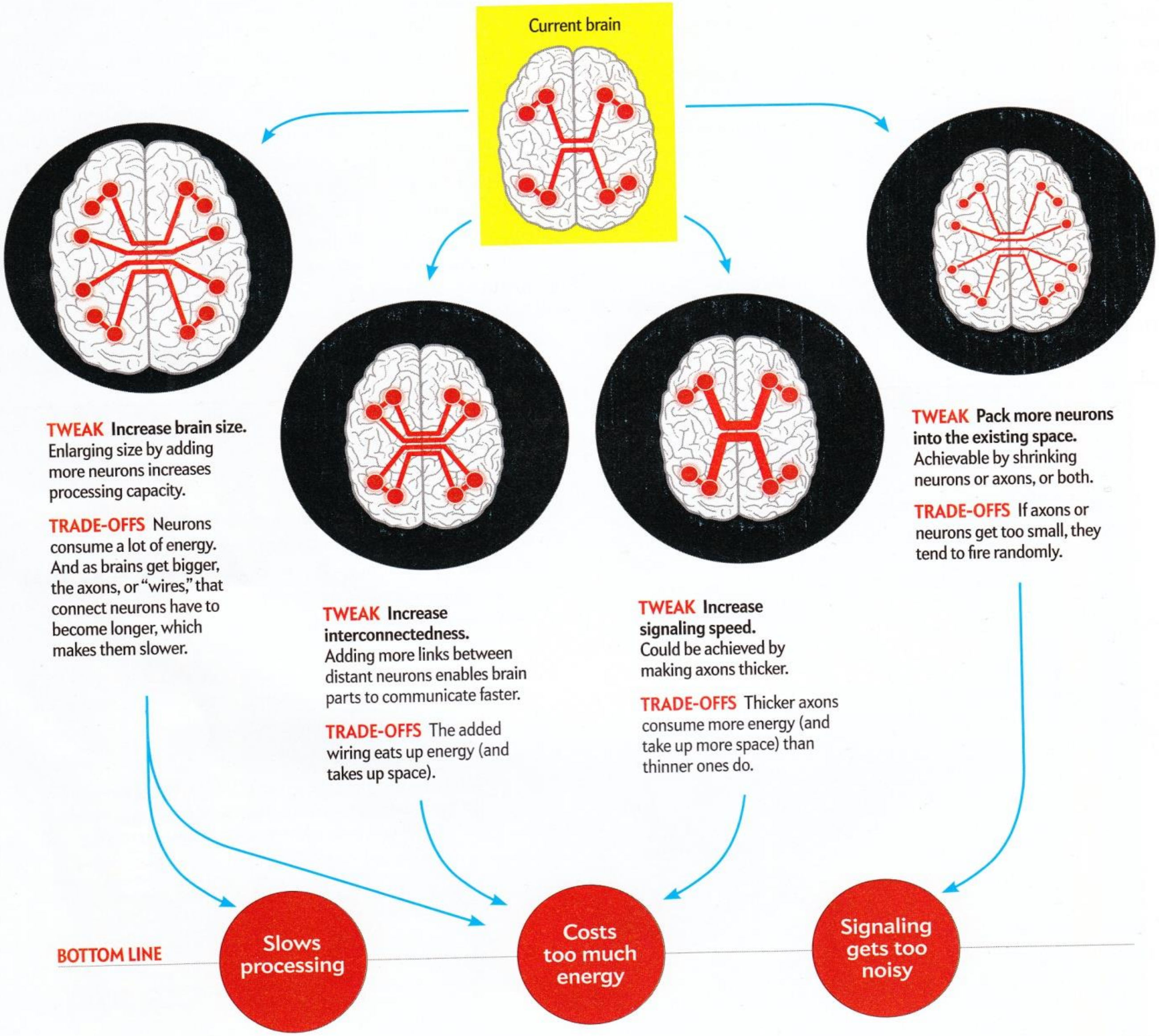
Grades: Rejected, Passed, Passed with honor

(+1 in final grade)

Examination: Friday 25.10.19 at 12.00-16.00, CA101, E100

Why We Probably Cannot Get Much Smarter

Miniaturization is just one of several evolutionary tweaks that could, in principle, enhance our intelligence and at the same time carry disadvantages and run into thermodynamic hurdles. Perhaps we are already close to being as smart as a neuron-based intelligence can be.



ARTIFICIAL

+

INTELLIGENCE

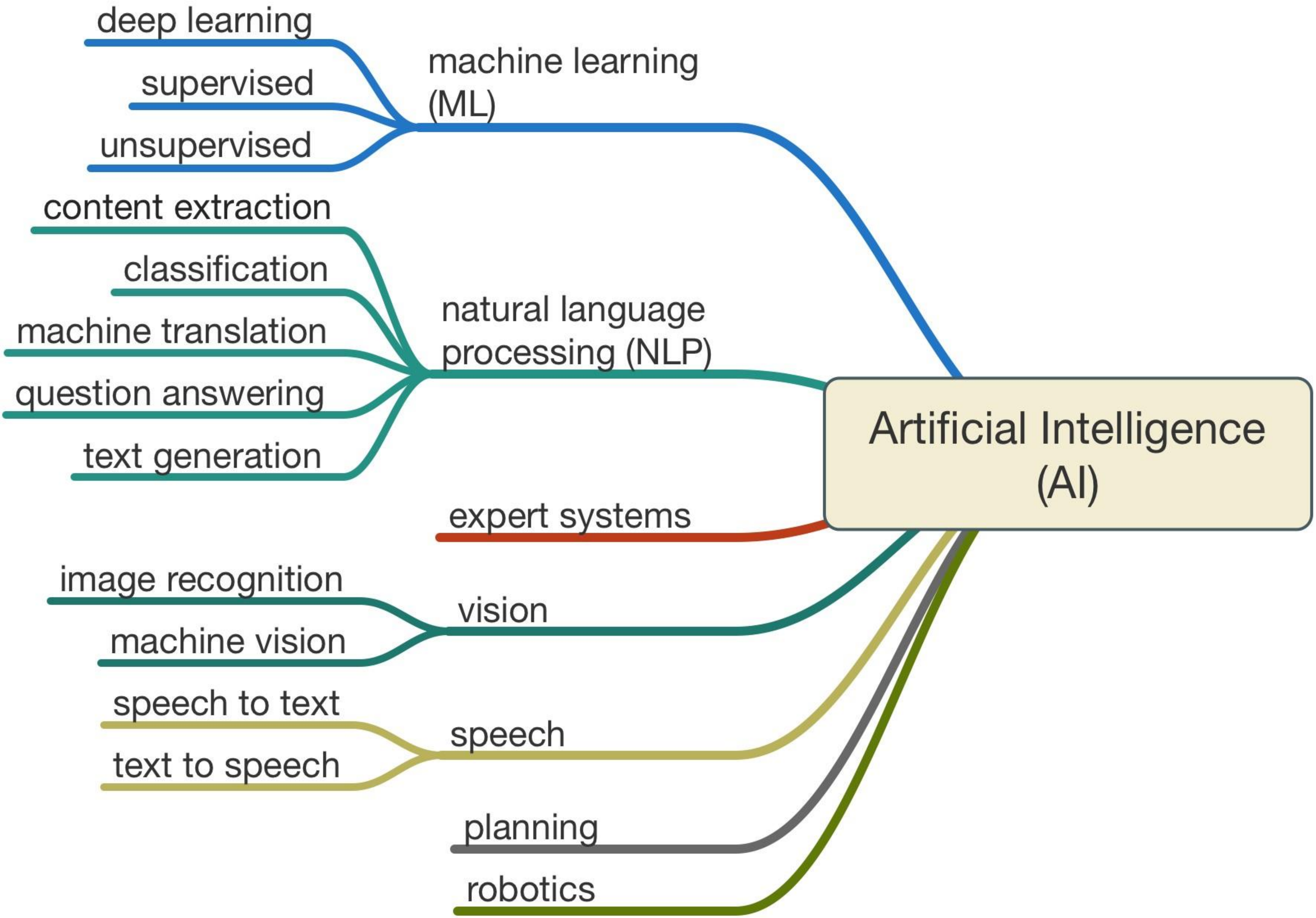
=

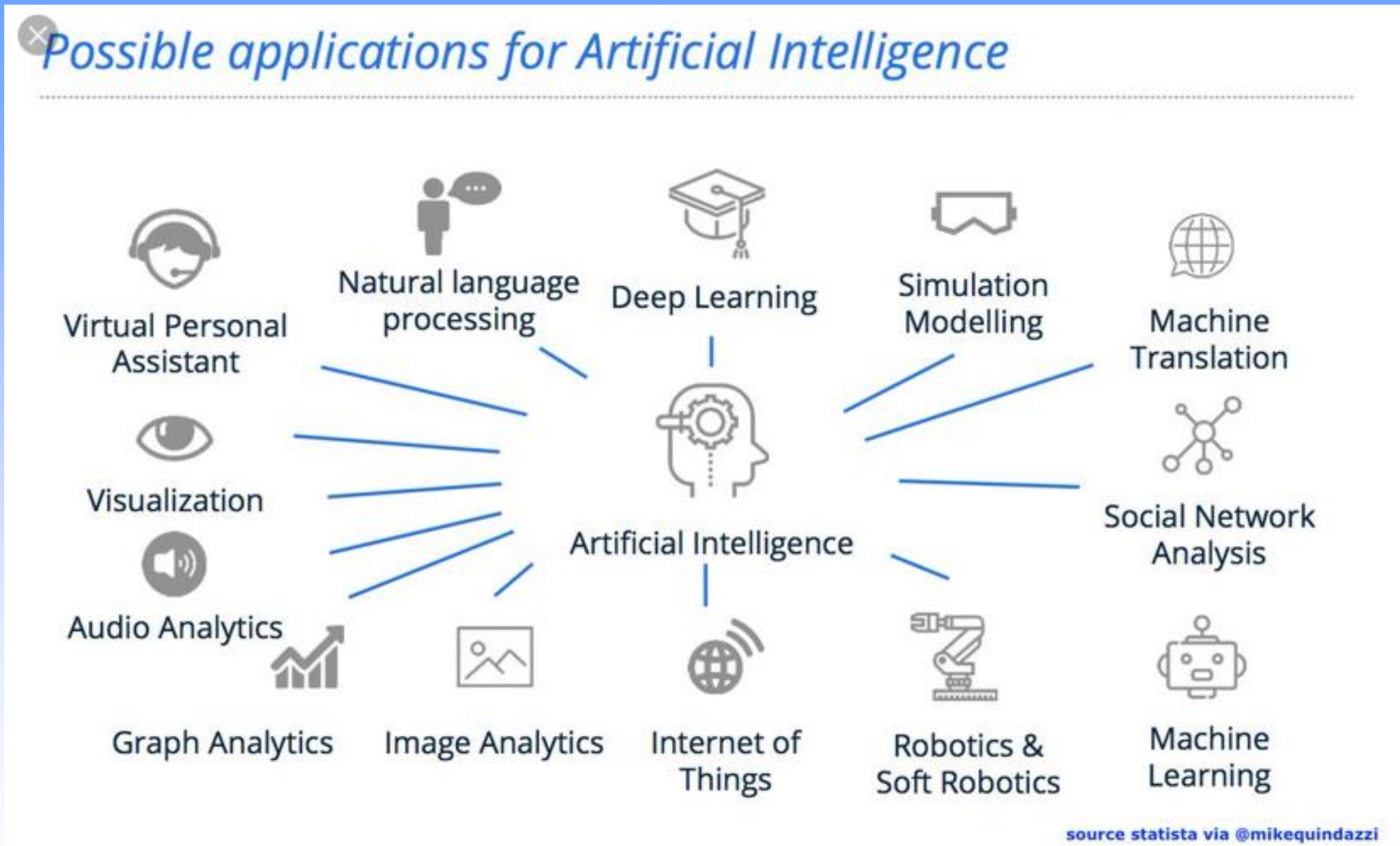
ARTIFICIAL INTELLIGENCE

- Intelligence: "The capacity to learn and solve problems"
- Artificial Intelligence: Artificial intelligence (AI) is the simulation of human intelligence by machines.
  - The ability to solve problems
  - The ability to act rationally
  - The ability to act like humans

Bio-inspired AI methods:

- Neural networks
  - Methods that have been inspired by the human brain
- Genetics algorithms and Evolutionary computing
- Swarm intelligence
- Emergent dynamics
- Etc.


Artificial Intelligence (AI)
machine learning (ML)
deep learning
supervised
unsupervised
natural language processing (NLP)
content extraction
classification
machine translation
question answering
text generation
expert systems
vision
image recognition
machine vision
speech
speech to text
text to speech
planning
robotics


Possible applications for Artificial Intelligence
Virtual Personal Assistant
Natural language processing
Deep Learning
Simulation Modelling
Machine Translation
Visualization
Artificial Intelligence
Social Network Analysis
Audio Analytics
Graph Analytics
Image Analytics
Internet of Things
Robotics & Soft Robotics
Machine Learning
source statista via @mikequindazzi

Definition of Intelligence

Webster's New Collegiate Dictionary defines intelligence as "**1a**

(1) : The ability to learn or understand or to deal with new or trying situations : REASON; *also* : the skilled use of reason

(2) : the ability to apply knowledge to manipulate one's environment or to think abstractly as measured by objective criteria (as tests)."

Another Definition of Intelligence

The capability of a system to adapt its behavior***** to meet its goals in a range of environments. It is a property of all purpose-driven decision makers.

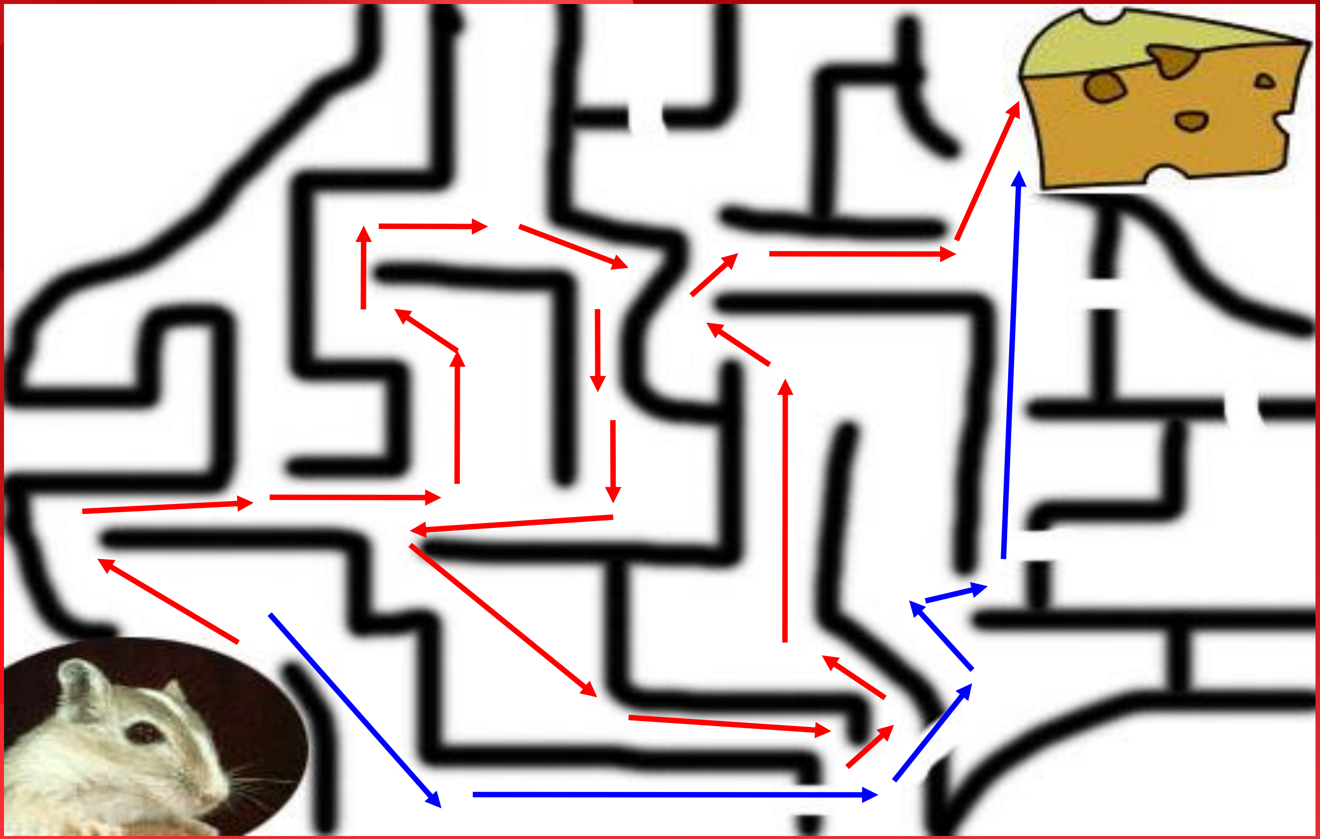
- David Fogel

***** implement decisions

Intelligence

- Are the things shown below, Intelligent?

Ex-1: Searching a path …



Different mice might follow different paths based to their intelligence
In other words: The problem can be solved in many ways

Ability to solve problems demonstrates Intelligence

Ex-2: Next number in the sequence …

- Consider the following sequence …

  1,3,7,13,21,___

  – What is the next number ?

- Key: Adding the next EVEN number …

1+2 = 3; 3+4 = 7; 7+6 = 13; 13+8 =21; 21+10 = 31

1,3,7,13,21,31

Ability to solve problems demonstrates Intelligence

Machine Translation

The spirit is willing but the flesh is weak. (2005/6/29)

Дух охотно готов но плоть слаба	Spirit is willingly ready but flesh it is weak
精神是愿意的但骨肉是微弱的	The spirit is wants but the flesh and blood is weak
精神は喜んでであるが、肉は弱い	Mind is rejoicing,, but the meat is weak
El alcohol está dispuesto pero la carne es débil	The alcohol is arranged but the meat is weak
الكحول مستعدّة غير أنّ اللحمة ضعيفة.	The alcohol is ready nevertheless the meat is weak.

- IBM statistical machine translation models
- US gov major consumer
  - Why Vodka (Russian)?

Question Answering

- What happened to Gagarin?
- Shallow natural language processing, heuristics

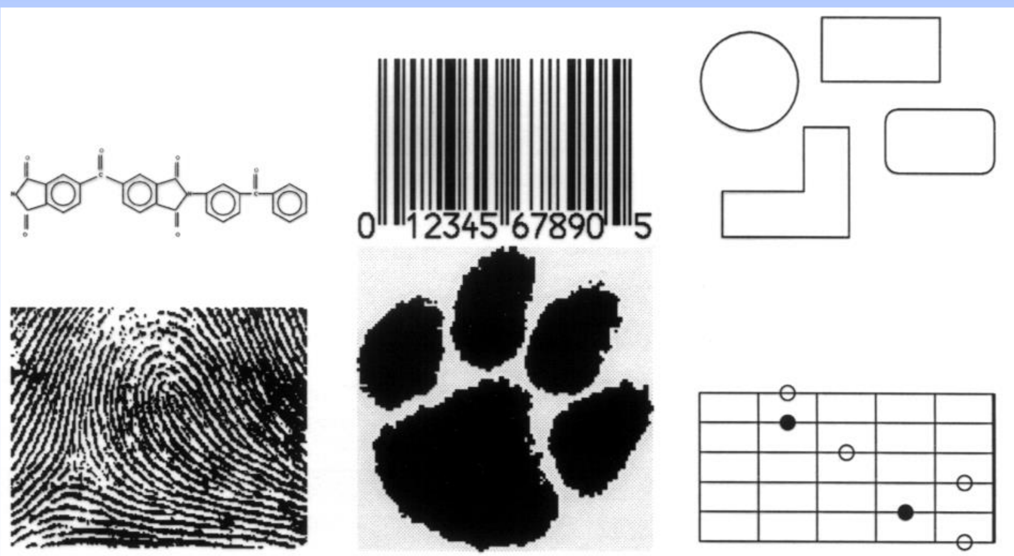
So, Let's Summarize…

- Ability to solve problems
- Ability to plan and schedule
- Ability to memorize and process information
- Ability to answer fuzzy questions
- Ability to learn
- Ability to recognize
- Ability to understand
- Ability to perceive
- And many more …

Food for thought: Can only humans beings and animals possess these qualities?

What is a Pattern?

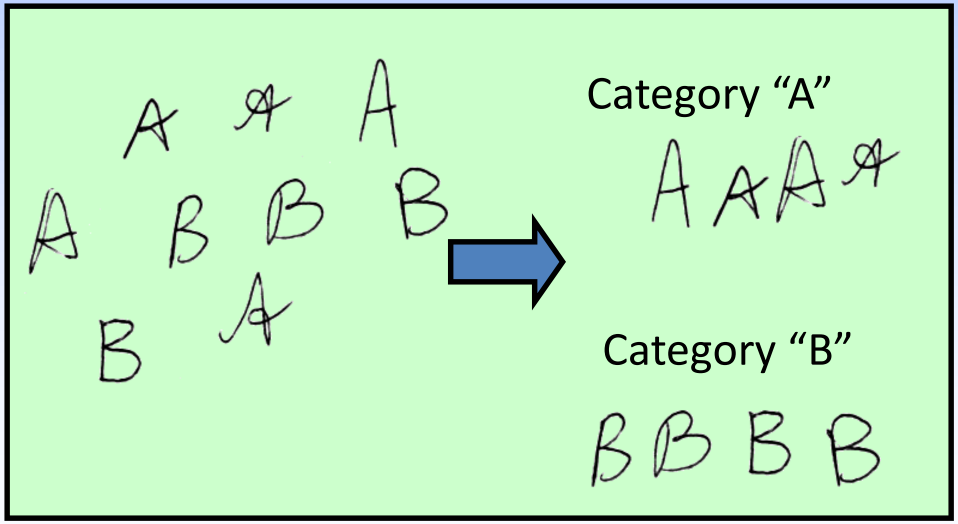
"A pattern is the opposite of a chaos; it is an entity vaguely defined, that could be given a name." (Watanabe)



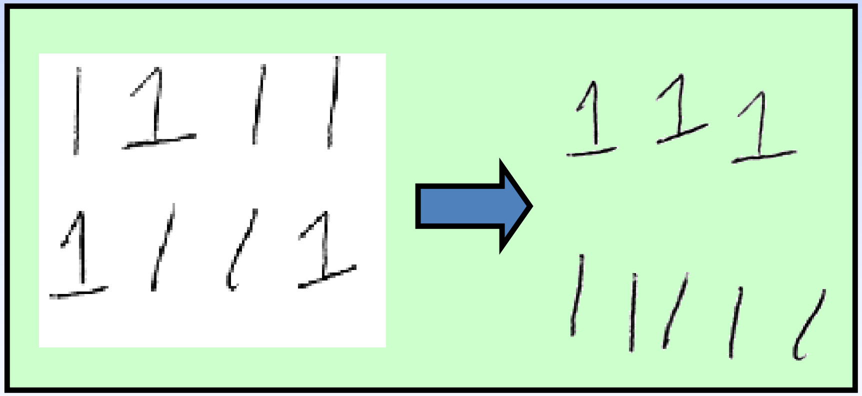
Recognition

Identification of a pattern as a member of a category we already know, or we are familiar with

- Classification (known categories)
- Clustering (learning categories)

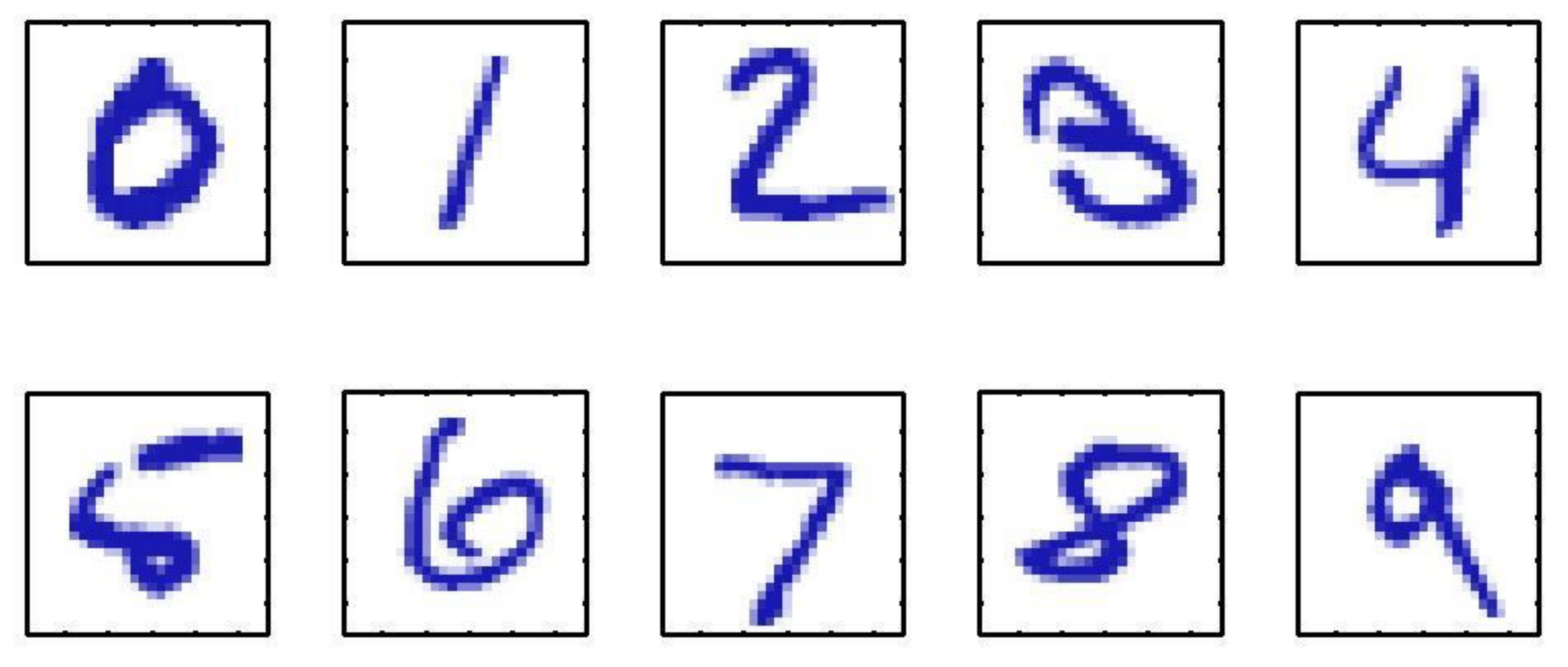


Classification



Clustering

Handwritten Digit Recognition



Cat vs. Dog

Supervised Classification

Training samples are labeled

Unsupervised Classification

Training samples are unlabeled

Pattern Class

- A collection of similar (not necessarily identical) objects
- A class is defined by class samples (paradigms, exemplars, prototypes, training/learning samples)
- **Intra-class variability**
- **Inter-class similarity**
- How do we define similarity?

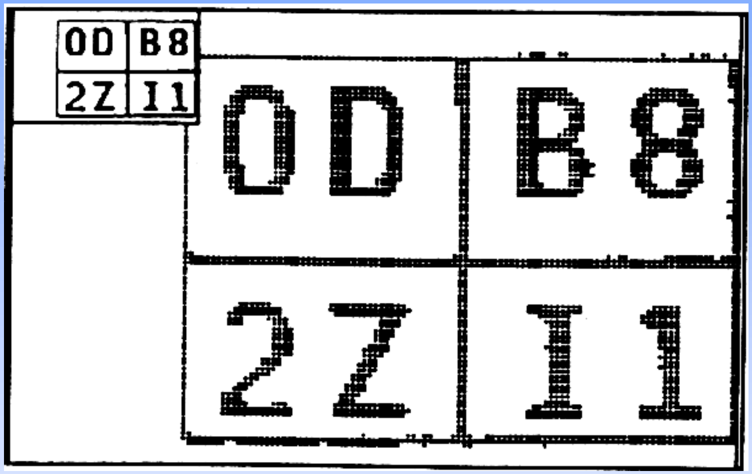
Intra-class Variability

The letter “T” in different typefaces

Same face under different expression, pose, illumination

Inter-class Similarity

Identical twins



Characters that look similar

Pattern Class Model

- A mathematical or statistical description for each pattern class (population); it is this class description that is learned from samples

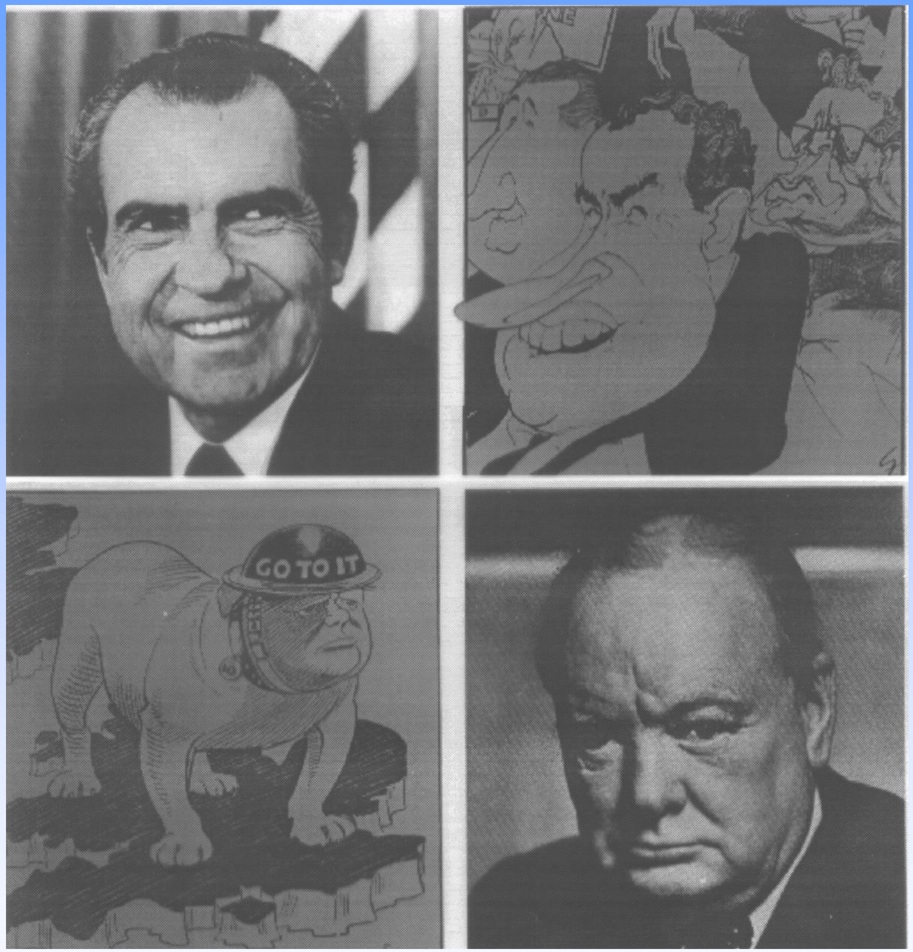
- Given a pattern, choose the best-fitting model for it; assign the pattern to the class associated with the best-fitting model

Pattern Recognition

- Having been shown a few positive examples (and perhaps a few negative examples) of a pattern class, the system learns to tell whether or not a new object belongs in this class (Watanabe)
- Inferring a generality from a few exemplars
- COGNITION = Formation of new classes

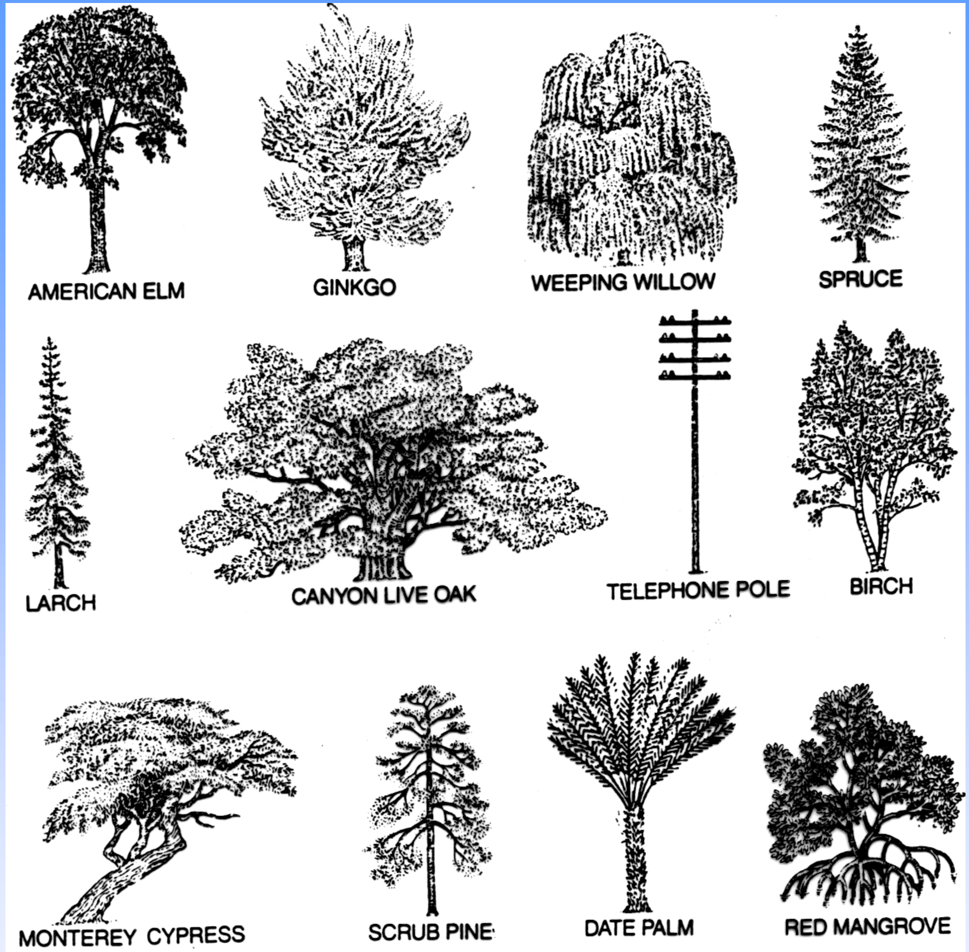
  RECOGNITION = known classes

Difficulties of Representation



How should we model a face to account for the large intra-class variability?

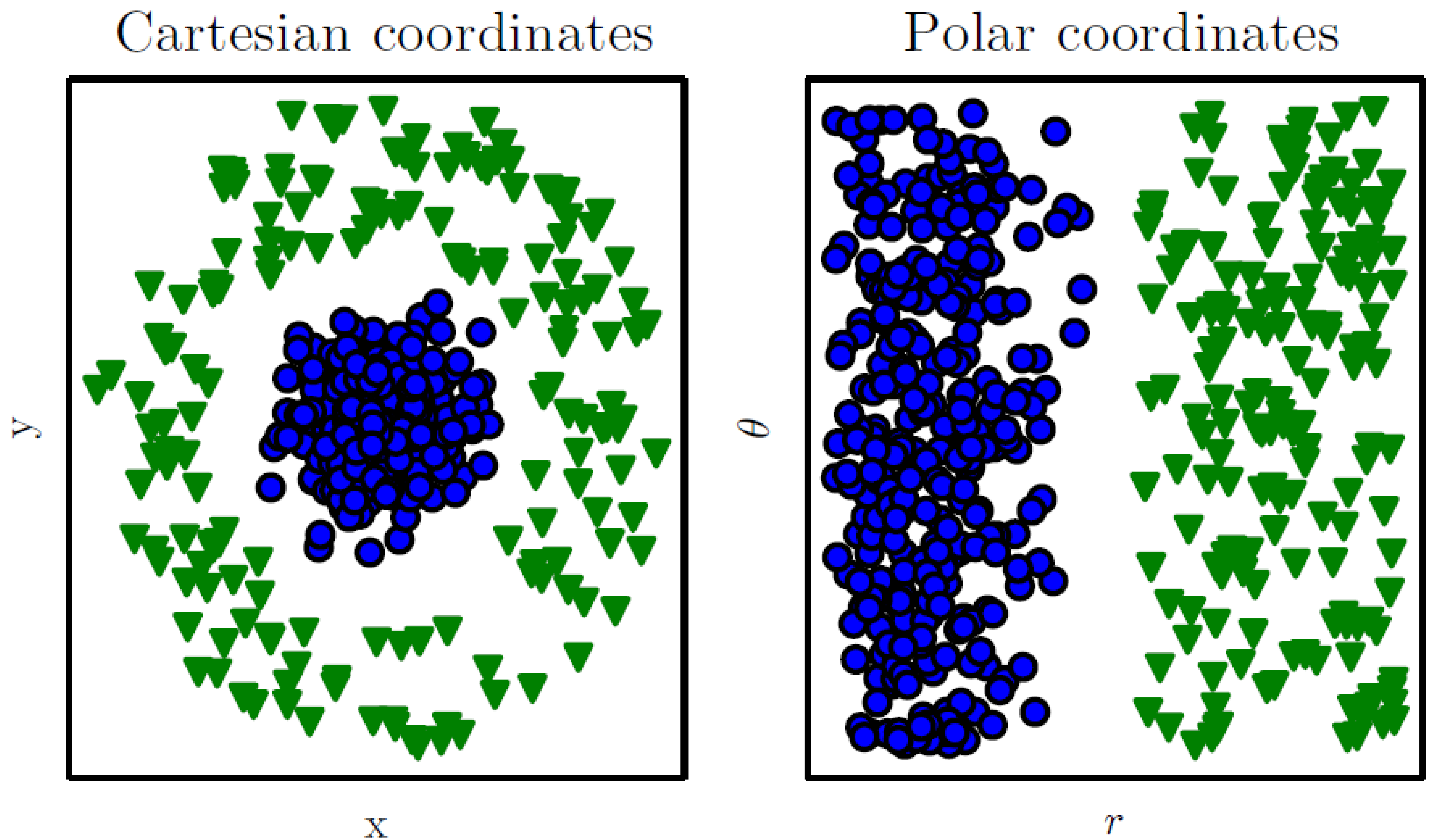
John P. Frisby, *Seeing. Illusion, Brian and Mind*, Oxford University Press, 1980



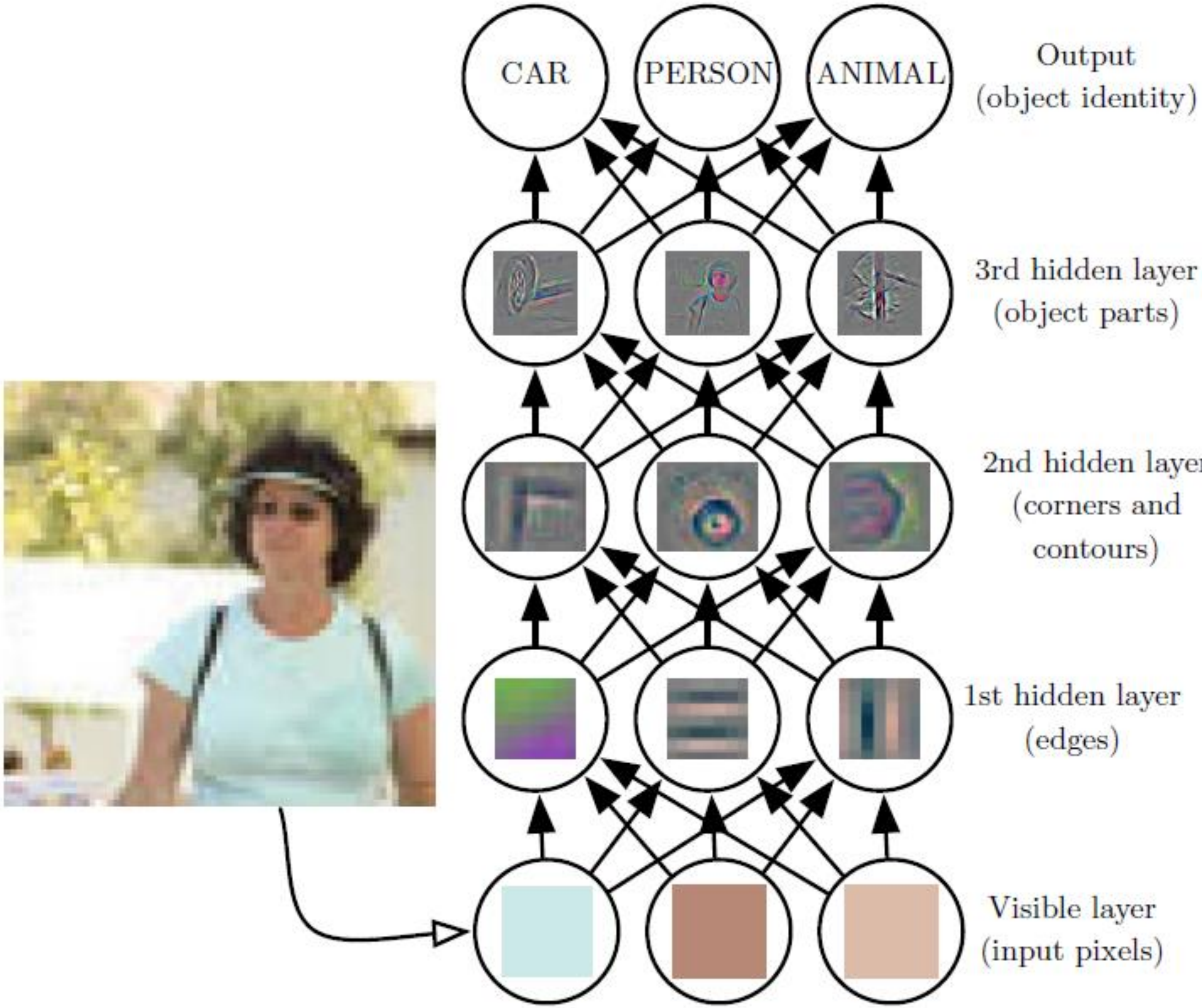
ARE ALL THESE OBJECTS TREES? Even a young child can answer correctly; a conventional computer, however, has enormous difficulty in doing so. Although there is a fair amount of regularity among the trees shown (each has a trunk and branches, for example), there is also a major component of arboreal irregularity among them. A generalized definition of a tree based on the underlying regularity could lead to erroneous identifications (such as mistaking a telephone pole, which has a "trunk" and "branches," for a tree). Hence any effective program designed to recognize trees would essentially have to be a list of all types of trees, which cannot be done in a few lines of computer code.


MO 200YRS 77 JAN
AK4-464
MONTANA 7 91
28-15521
BIG SKY '76
NEBRASKA
8-A9968
1776 BICENTENNIAL 1976
ALABAMA '77
CFZ 668
ALASKA
ARN 866
centennial
KR 3615
19 colorado
NEW HAMPSHIRE
5568
FIRST FOR INDEPENDENCE
CROSSROADS of the REVOLUTION
NJ
LAND OF ENCHANTMENT
EMT-1
19 NEW MEXICO USA 76
1776 BICENTENNIAL 1976
356 890
WASHINGTON D.C.
BICENTENNIAL
1776 1976
THE FIRST STATE
DELAWARE
ga.
1776 1976
1645
1540-1776-1976
NEW MEXICO BICENTENNIAL
SANTA FE
000 000
new mexico
PENNSYLVANIA
436 800
BICENTENNIAL STATE '76
HAWAII 76
2M-2799
ALOHA STATE
449 316
1976 INDIANA
96 C 3050
Heritage State
19 OKLAHOMA 76
PE-1373
1776 BICENTENNIAL 1976
1776 Bicentennial 1976
NCZ 198
SOUTH CAROLINA
76
VERMONT BICENTENNIAL
Maine
76
BICENTENNIAL
345 AEF
1776 MARYLAND 1976
MICHIGAN 78
VYR737
14th STATE
1791-1991
VERMONT BICENTENNIAL
OCT Virginia 76
135 709
1776 Independence Bicentennial 1976
1 3 7148
19 WYOMING

Representations Matter



Depth: Repeated Composition



Difficulties of Representation

- "How do you instruct someone (or some computer) to recognize caricatures in a magazine, let alone find a human figure in a misshapen piece of work?"
- "A program that could distinguish between male and female faces in a random snapshot would probably earn its author a Ph.D. in computer science." (Penzias 1989)
- A representation could consist of a vector of real-valued numbers, ordered list of attributes, parts and their relations….

Good Representation

- Should have some invariant properties (e.g., w.r.t. rotation, translation, scale…)
- Account for intra-class variations
- Ability to discriminate pattern classes of interest
- Robustness to noise, occlusion,..
- Lead to simple decision making strategies (e.g., linear decision boundary)
- Low measurement cost; real-time

Pattern Recognition in Practice

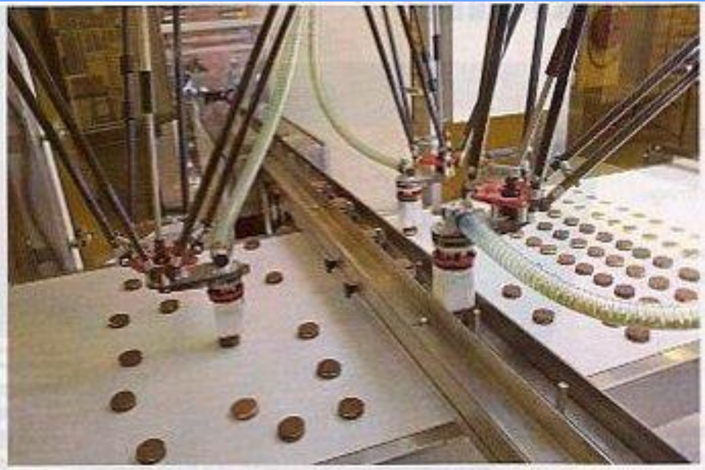
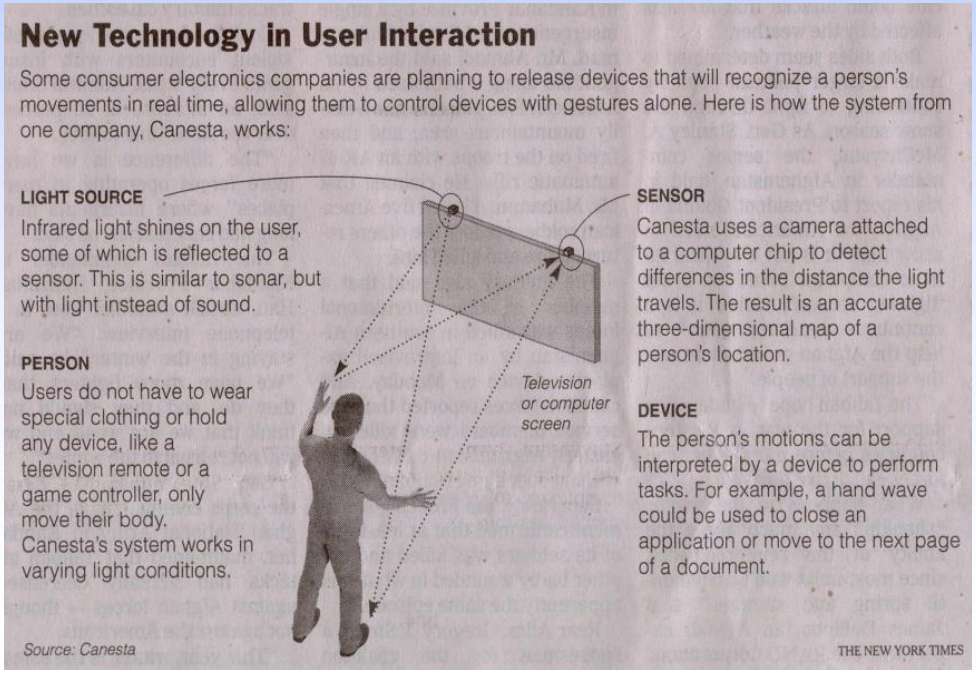


FIGURE 2. Single-camera solutions, such as this cookie packing system from Bosch Packaging Technology, use a geometric pattern-matching algorithm to determine the shape and location of cookies and then stack them in the appropriate section of the tray

FIGURE 3. To automate the process of orange picking, Vision Robotics has proposed the development of a stereo camera-based system that will use multiple cameras placed at the end of multiaxis arms to create a virtual 3-D image of the entire orange tree.

Vision System Design, Nov 2009



NY Times, Jan 12, 2010

Pattern Recognition Applications

Table 1: Example pattern recognition applications.

Problem Domain	Application	Input Pattern	Pattern Classes
Document image analysis	Optical character recognition	Document image	Characters, words
Document classification	Internet search	Text document	Semantic categories
Document classification	Junk mail filtering	Email	Junk/non-junk
Multimedia database retrieval	Internet search	Video clip	Video genres
Speech recognition	Telephone directory assistance	Speech waveform	Spoken words
Natural language processing	Information extraction	Sentences	Parts of speech
Biometric recognition	Personal identification	Face, iris, fingerprint	Authorized users for access control
Medical	Computer aided diagnosis	Microscopic image	Cancerous/healthy cell
Military	Automatic target recognition	Optical or infrared image	Target type
Industrial automation	Printed circuit board inspection	Intensity or range image	Defective/non-defective product
Industrial automation	Fruit sorting	Images taken on a conveyor belt	Grade of quality
Remote sensing	Forecasting crop yield	Multispectral image	Land use categories
Bioinformatics	Sequence analysis	DNA sequence	Known types of genes
Data mining	Searching for meaningful patterns	Points in multidimensional space	Compact and well-separated clusters

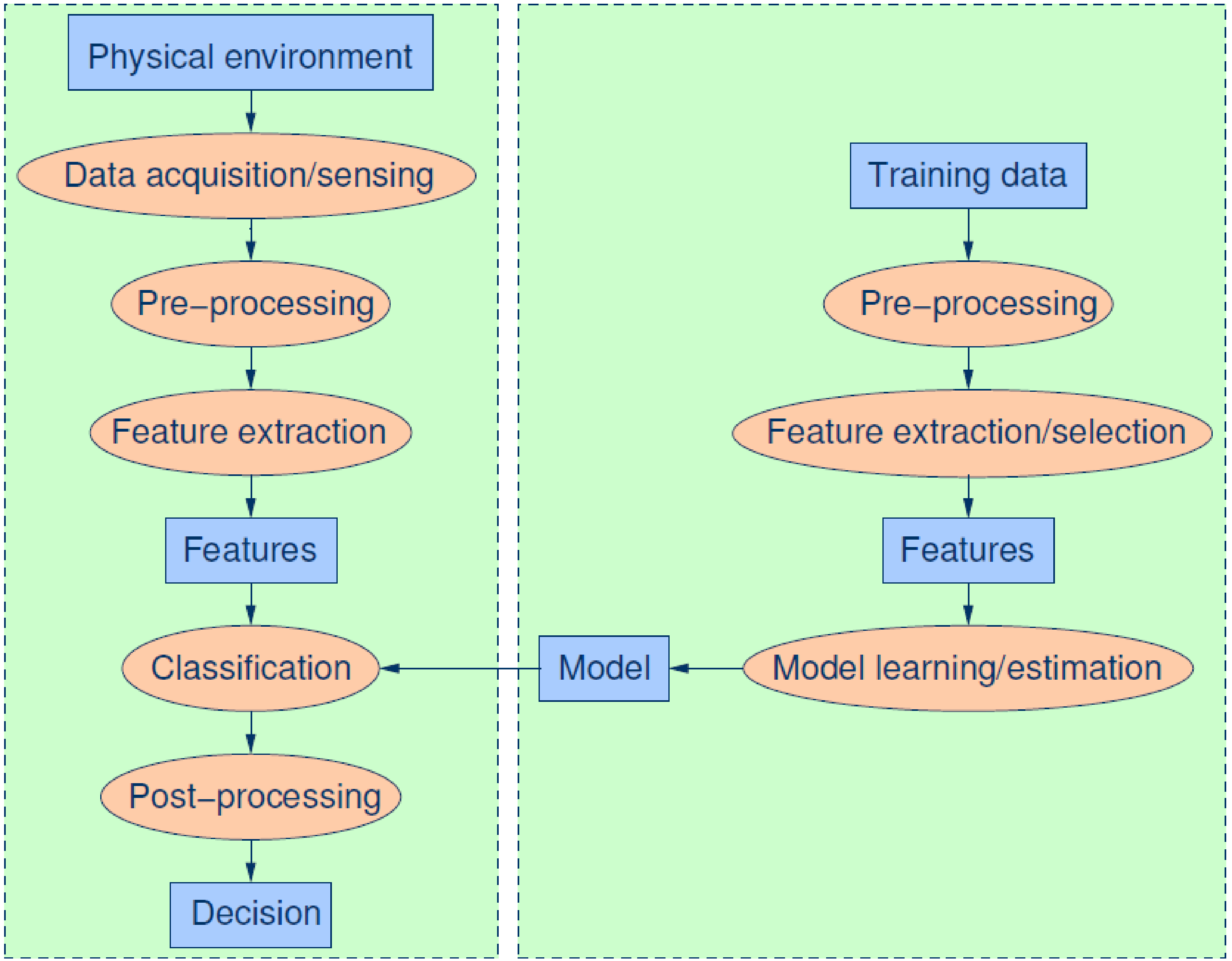
Pattern Recognition Applications

Problem	Input	Output
Speech recognition	Speech waveforms	Spoken words, speaker identity
Non-destructive testing	Ultrasound, eddy current, acoustic emission waveforms	Presence/absence of flaw, type of flaw
Detection and diagnosis of disease	EKG, EEG waveforms	Types of cardiac conditions, classes of brain conditions
Natural resource identification	Multispectral images	Terrain forms, vegetation cover
Aerial reconnaissance	Visual, infrared, radar images	Tanks, airfields
Character recognition (page readers, zip code, license plate)	scanned image	Alphanumeric characters

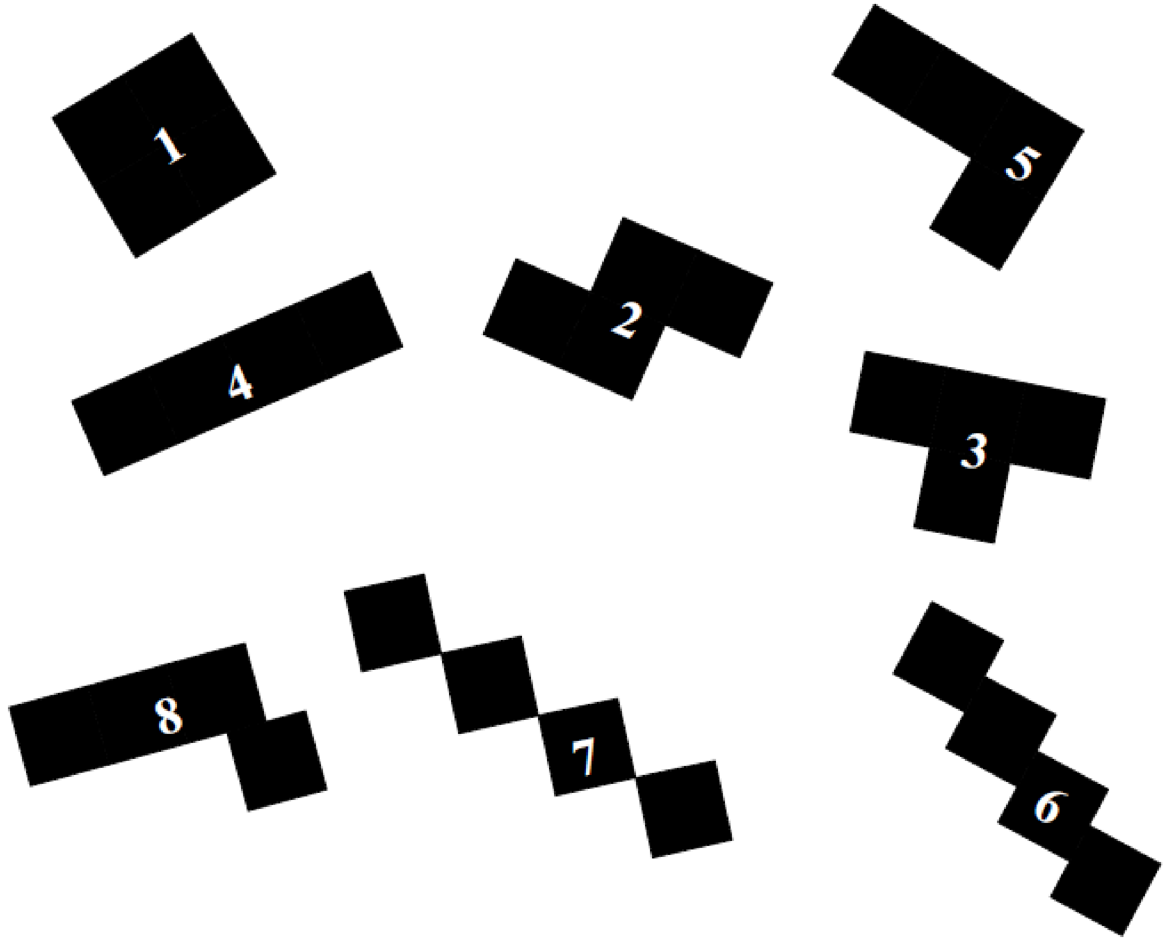
Pattern Recognition Applications

Problem	Input	Output
Identification and counting of cells	Slides of blood samples, micro-sections of tissues	Type of cells
Inspection (PC boards, IC masks, textiles)	Scanned image (visible, infrared)	Acceptable/unacceptable
Manufacturing	3-D images (structured light, laser, stereo)	Identify objects, pose, assembly
Web search	Key words specified by a user	Text relevant to the user
Fingerprint identification	Input image from fingerprint sensors	Owner of the fingerprint, fingerprint classes
Online handwriting retrieval	Query word written by a user	Occurrence of the word in the database

Pattern Recognition Systems

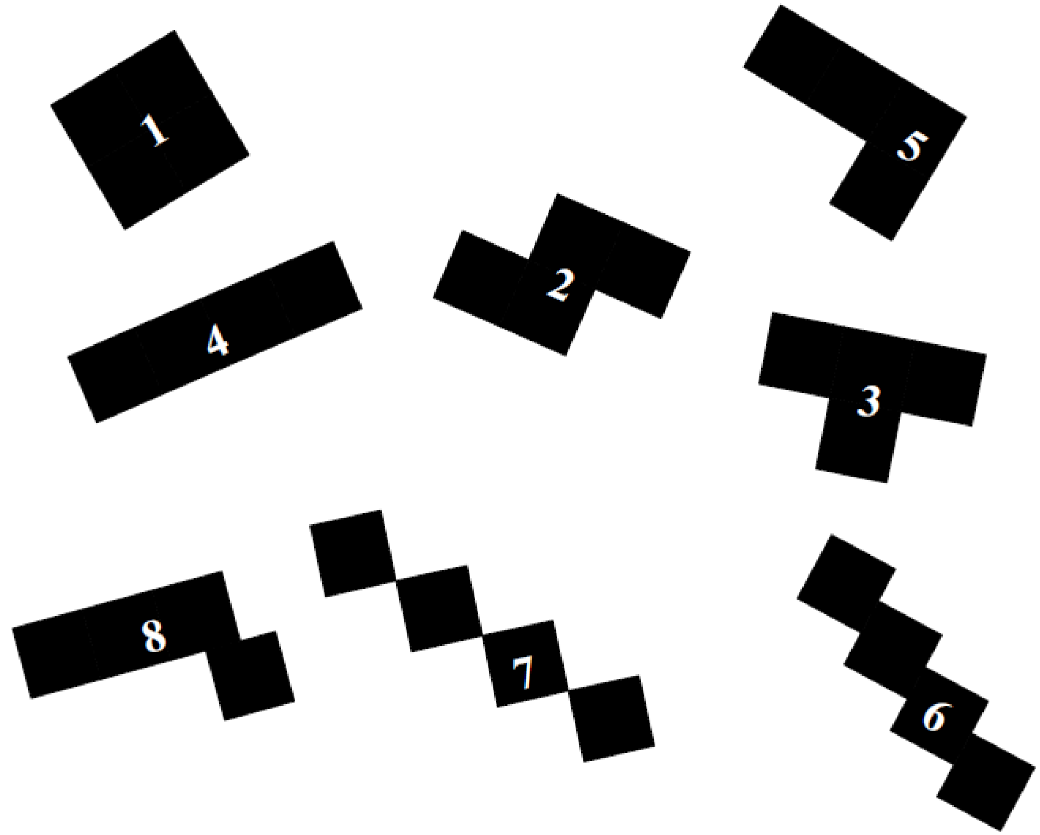


Simple recognition exercise: classification of objects



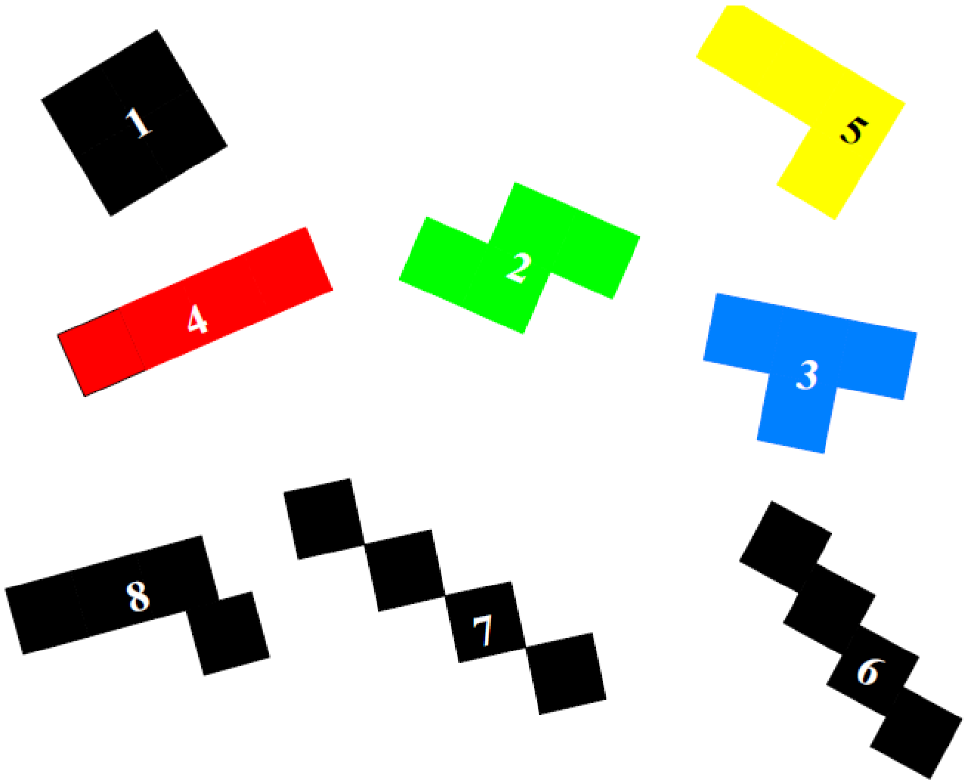
Classification of objects characterization with two features

object	area	perimeter
1	4	8
2	4	10
3	4	10
4	4	10
5	4	10
6	4	13
7	4	16
8	4	11

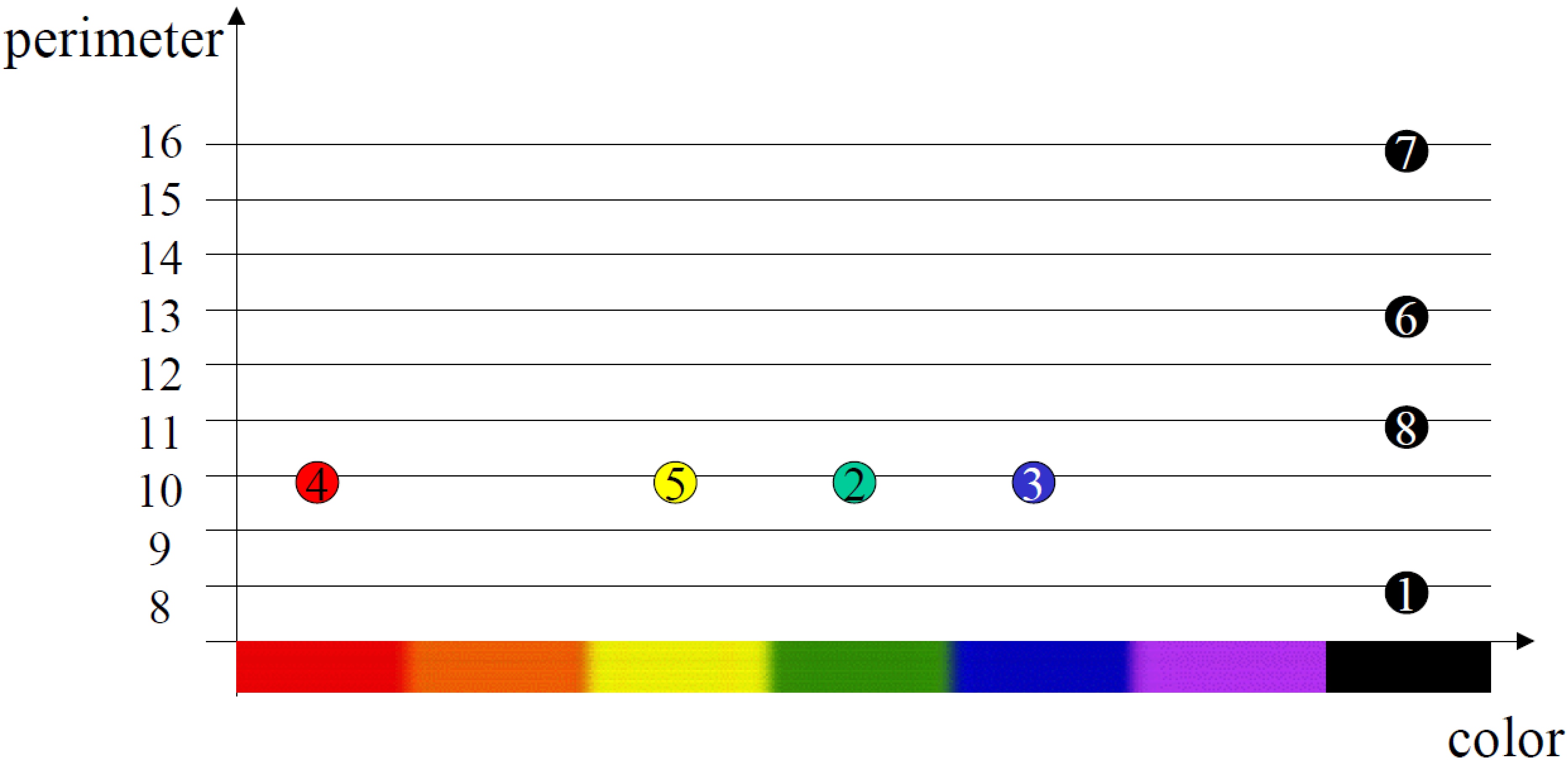


Classification of objects characterization with three features

object	area	perimeter	colour
1	4	8	black
2	4	10	green
3	4	10	blue
4	4	10	red
5	4	10	yellow
6	4	13	black
7	4	16	black
8	4	11	black



Two-dimensional feature space



A Case Study: Fish Classification

- Problem:
  - sort incoming fish on a conveyor belt according to species
  - Assume only two classes exist:
    - Sea Bass and Salmon

A Case Study: Fish Classification

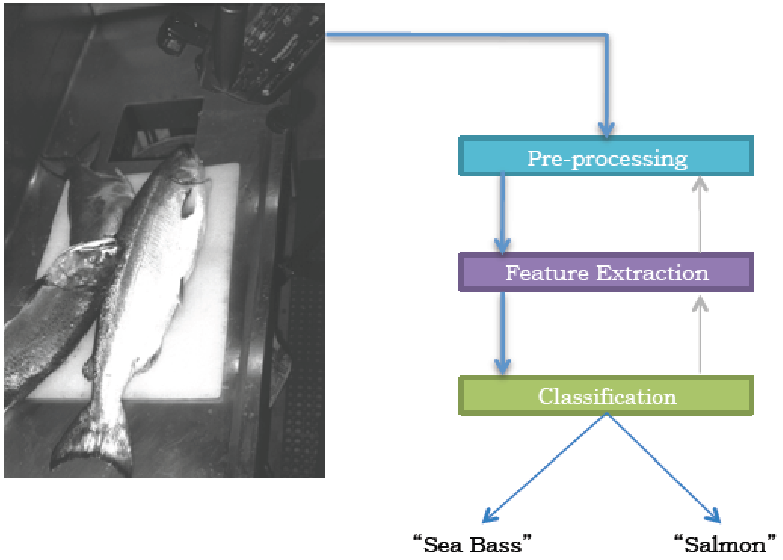
- What kind of information can distinguish one species from the other?

– length, width, weight, number and shape of fins, tail shape, etc.

- What can cause problems during sensing?

– lighting conditions, position of fish on the conveyor belt, camera noise, etc.

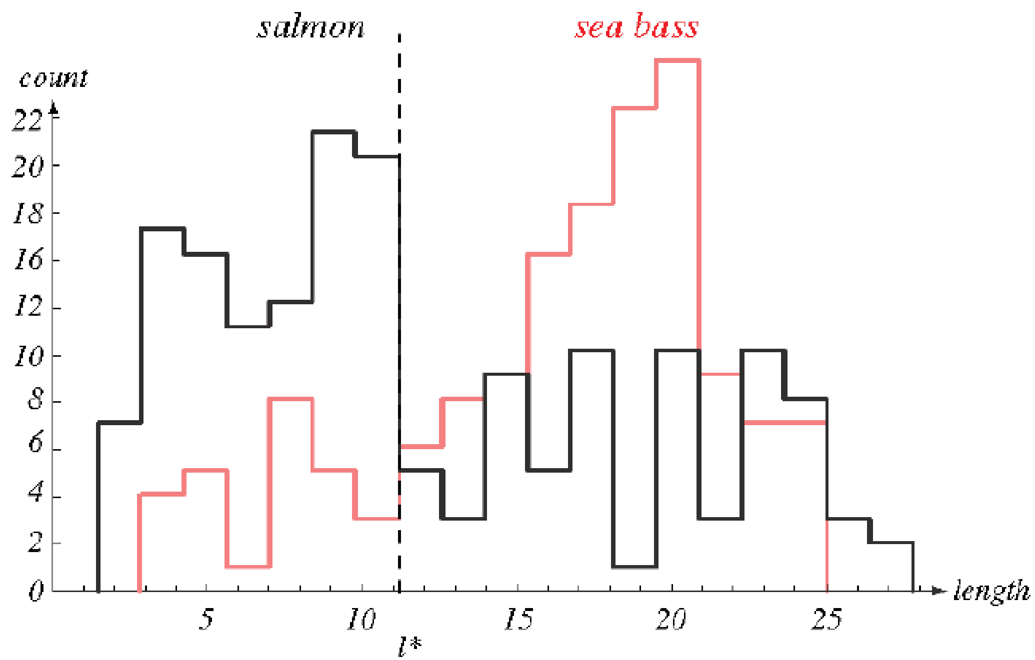
- What are the steps in the process?
    1. Capture image.
    2. Isolate fish
    3. Take measurements
    4. Make decision



A Case Study: Fish Classification

- Selecting Features
  - Assume a fisherman told us that a <u>sea bass is generally longer than a salmon</u>.
  - We can use <u>length as a feature</u> and decide between sea bass and salmon according to a threshold on length.
  - How can we choose this threshold?

Histograms of the length feature for two types of fish in training samples. How can we choose the threshold l* to make a reliable decision?

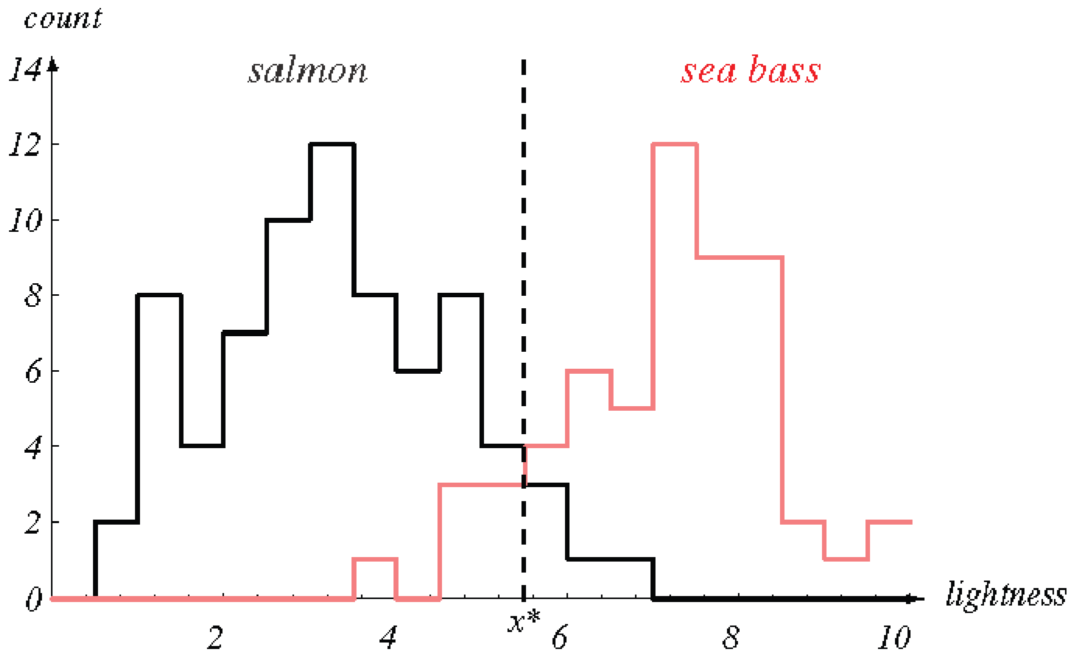


Even though "sea bass" is longer than "salmon" on the average, **there are many examples of fish where this observation does not hold...**

A Case Study: Fish Classification

- Selecting Features
  - Let's try another feature and see if we get better discrimination
    - ➡Average Lightness of the fish scales

Histograms of the lightness feature for two types of fish in training samples.



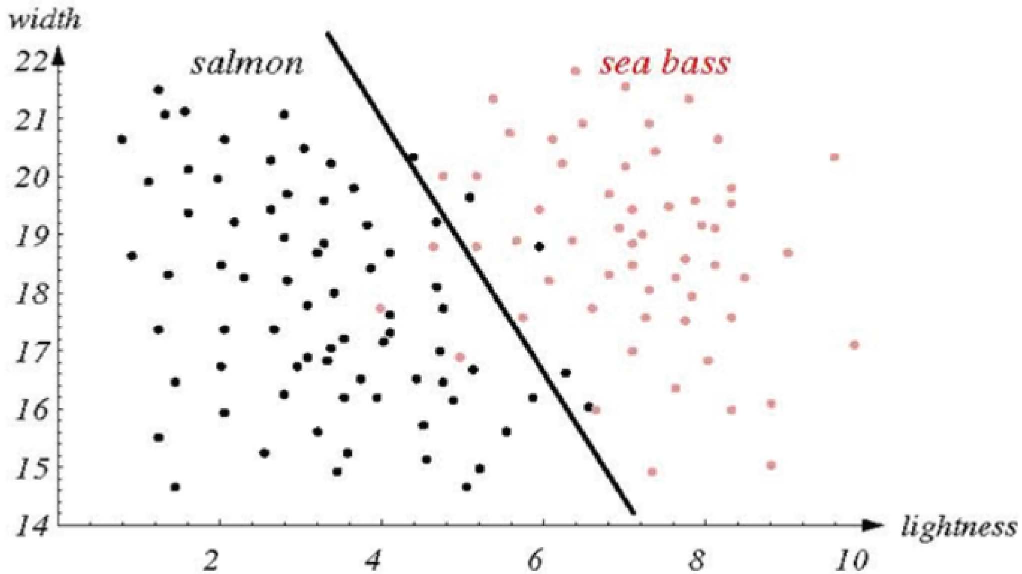
It looks easier to choose the threshold x* but we still cannot make a perfect decision.

A Case Study: Fish Classification

- Multiple Features
  - Single features might not yield the best performance.
  - To improve recognition, we might have to use more than one feature at a time.
  - Combinations of features might yield better performance.
  - Assume we also observed that sea bass are typically wider than salmon.

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \begin{array}{l} x_1 : lightness \\ x_2 : width \end{array}$$

Each fish image is now represented by a point in this 2D feature space

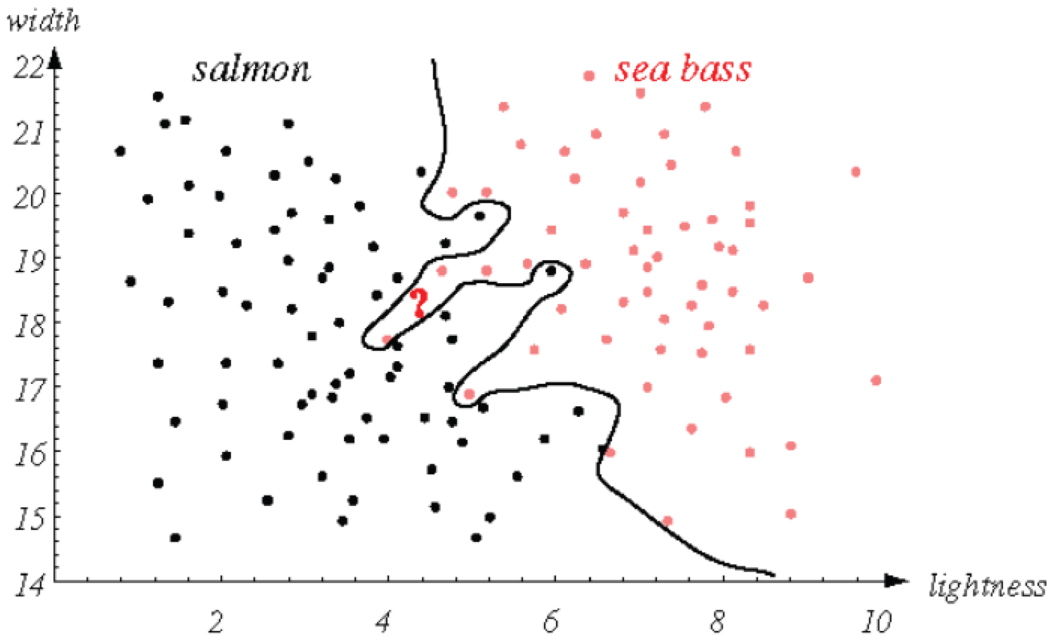


Scatter plot of lightness and width features for training samples. **We can draw a decision boundary to divide the feature space into two regions. Does it look better than using only lightness?**

A Case Study: Fish Classification

- Designing a **Classifier**
  - Can we do better with another decision rule?
  - More complex models result in more complex boundaries.

We may distinguish training samples perfectly but **how can we predict how well we can generalize to unknown samples?**



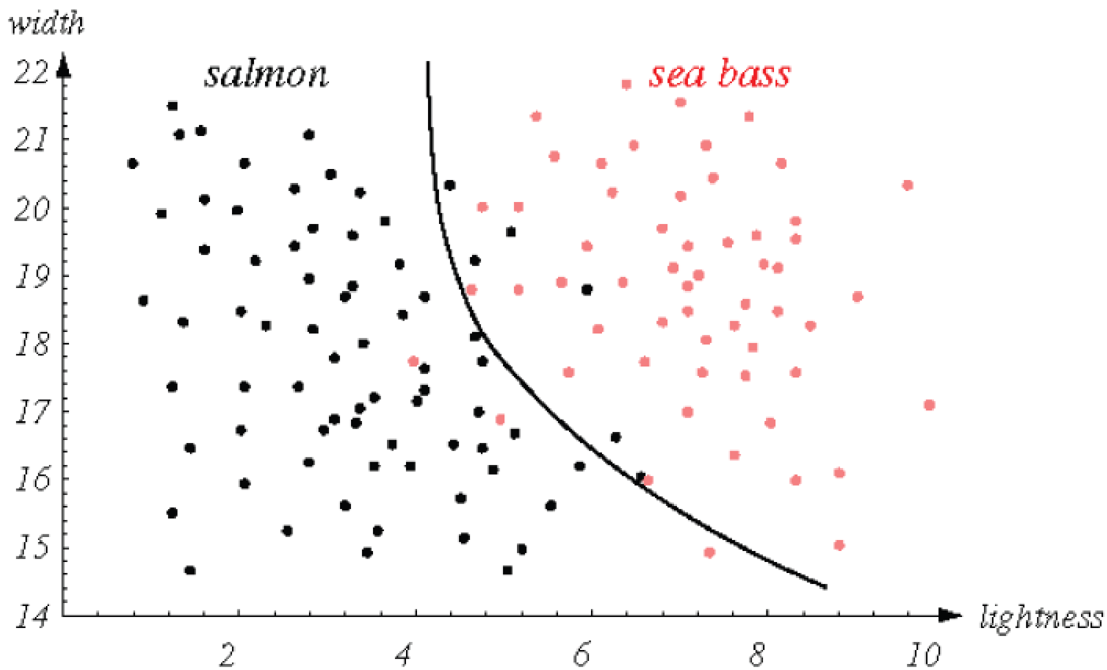
DANGER OF **OVER FITTING!!**

CLASSIFIER WILL FAIL TO GENERALIZE TO NEW DATA...

A Case Study: Fish Classification

- Designing a **Classifier**
  - How can we manage the tradeoff between complexity of decision rules and their performance to unknown samples?

Different criteria lead to different decision boundaries



Feature Extraction

- Designing a **Feature Extractor**
  - Its design is **problem specific** (e.g. features to extract from graphic objects may be quite different from sound events...)
  - The ideal feature extractor would produce the same feature vector X for all patterns in the same class, and different feature vectors for patterns in different classes.
  - In practice, different inputs to the feature extractor will always produce different feature vectors, but we hope that the **within-class variability** is small relative to the **between-class variability**.
- **Designing a good set of features is sometimes "more of an art than a science"...**

Feature Extraction

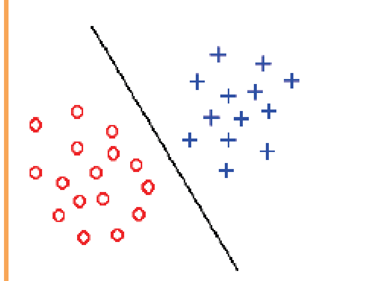
- Multiple Features
  - Does adding more features always improve the results?
    - No!! So we must:
      - Avoid unreliable features.
      - Be careful about correlations with existing features.
      - Be careful about measurement costs.
      - Be careful about noise in the measurements.
  - **Is there some curse for working in very high dimensions?**
    - YES THERE IS! ==> **CURSE OF DIMENSIONALITY**

      ➡**thumb rule: $n >= d(d-1)/2$**

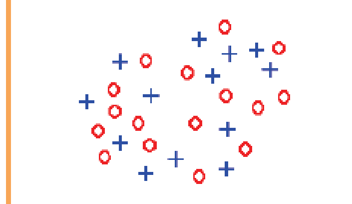
      n = nr of examples in training dataset
      d = nr of features

Feature Extraction

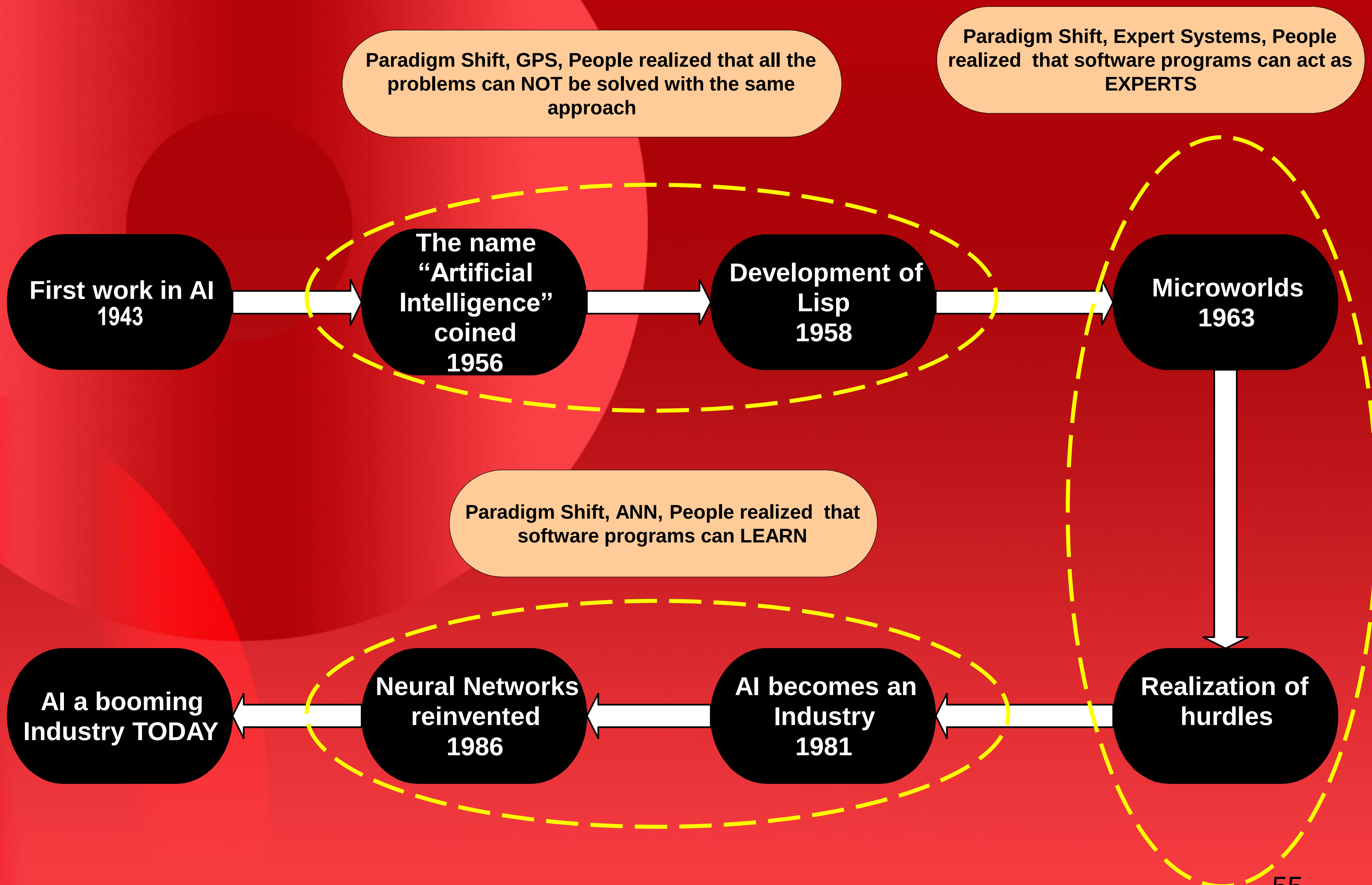
- Problem: Inadequate Features
  - features simply do not contain the information needed to separate the classes, it doesn't matter how much effort you put into designing the classifier.
  - **Solution**: go back and design better features.



"Good" features



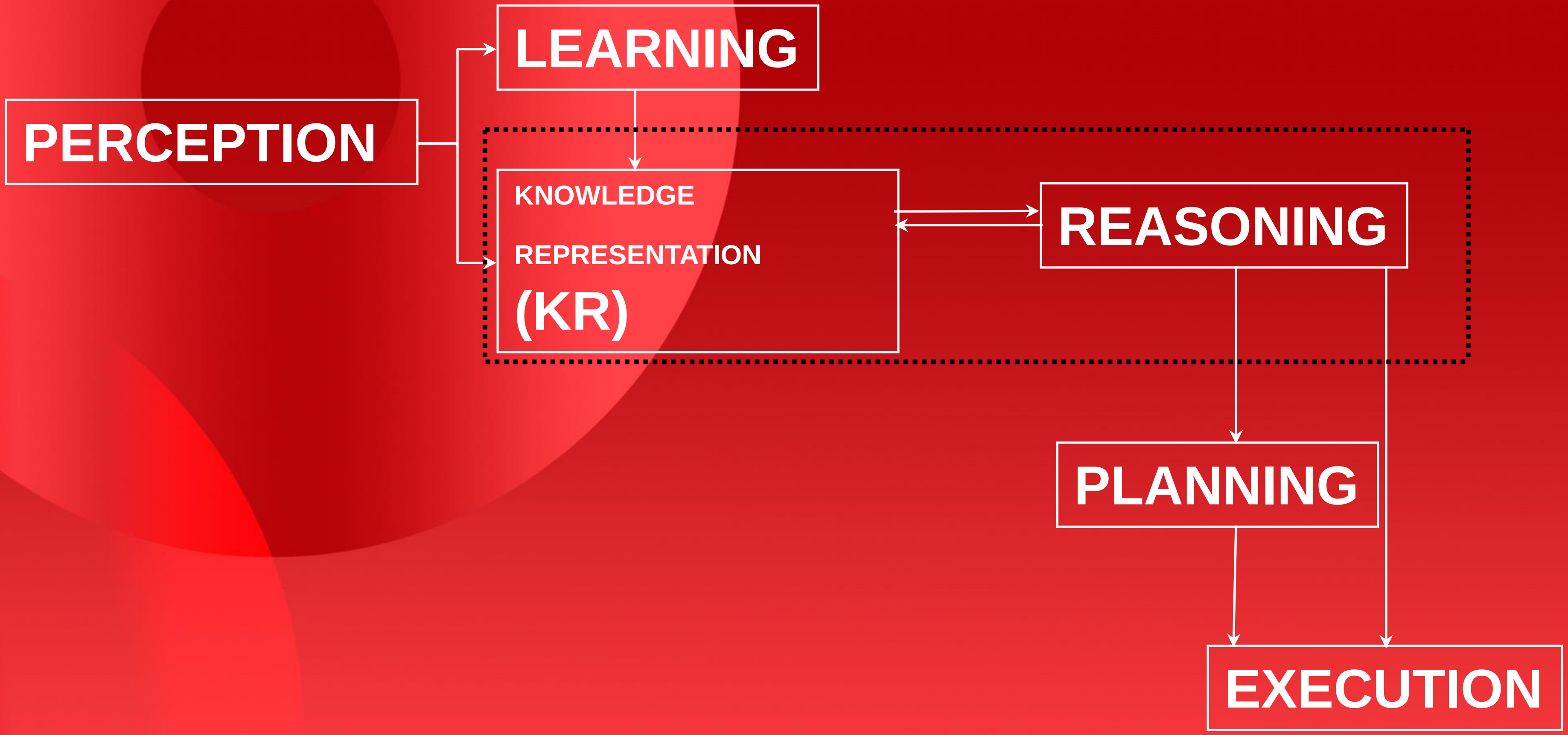
"Bad" features


Paradigm Shift, GPS, People realized that all the problems can NOT be solved with the same approach
Paradigm Shift, Expert Systems, People realized that software programs can act as EXPERTS
First work in AI 1943
The name "Artificial Intelligence" coined 1956
Development of Lisp 1958
Microworlds 1963
Paradigm Shift, ANN, People realized that software programs can LEARN
AI a booming Industry TODAY
Neural Networks reinvented 1986
AI becomes an Industry 1981
Realization of hurdles

Lecture Contents

- Types of Knowledge:
  - procedural,
  - declarative,
  - meta,
  - heuristic,
  - structural
- Knowledge Representation Techniques
  - Facts
  - Object-attribute Value Triplets
  - Semantic Networks
  - Frames
  - Logic:
    - Propositional Calculus
    - Predicate Calculus

The AI Cycle


LEARNING
PERCEPTION
KNOWLEDGE
REPRESENTATION
(KR)
REASONING
PLANNING
EXECUTION

The Dilemma

- We do not know how the KR and reasoning components are implemented in humans, even though we can see their manifestation in the form of intelligent behavior.
- Hence, the need for a synthetic (artificial) way to model the knowledge representation and reasoning capability of humans in computers.

The Simple Approach

- Instead of focusing on how knowledge is acquired, we will assume for now that knowledge is externally injected into the system.
- For now, we focus on how to **represent** some given knowledge and then how to **reason** about that knowledge for the purpose of inference
- **Knowledge acquisition and learning will be discussed later**

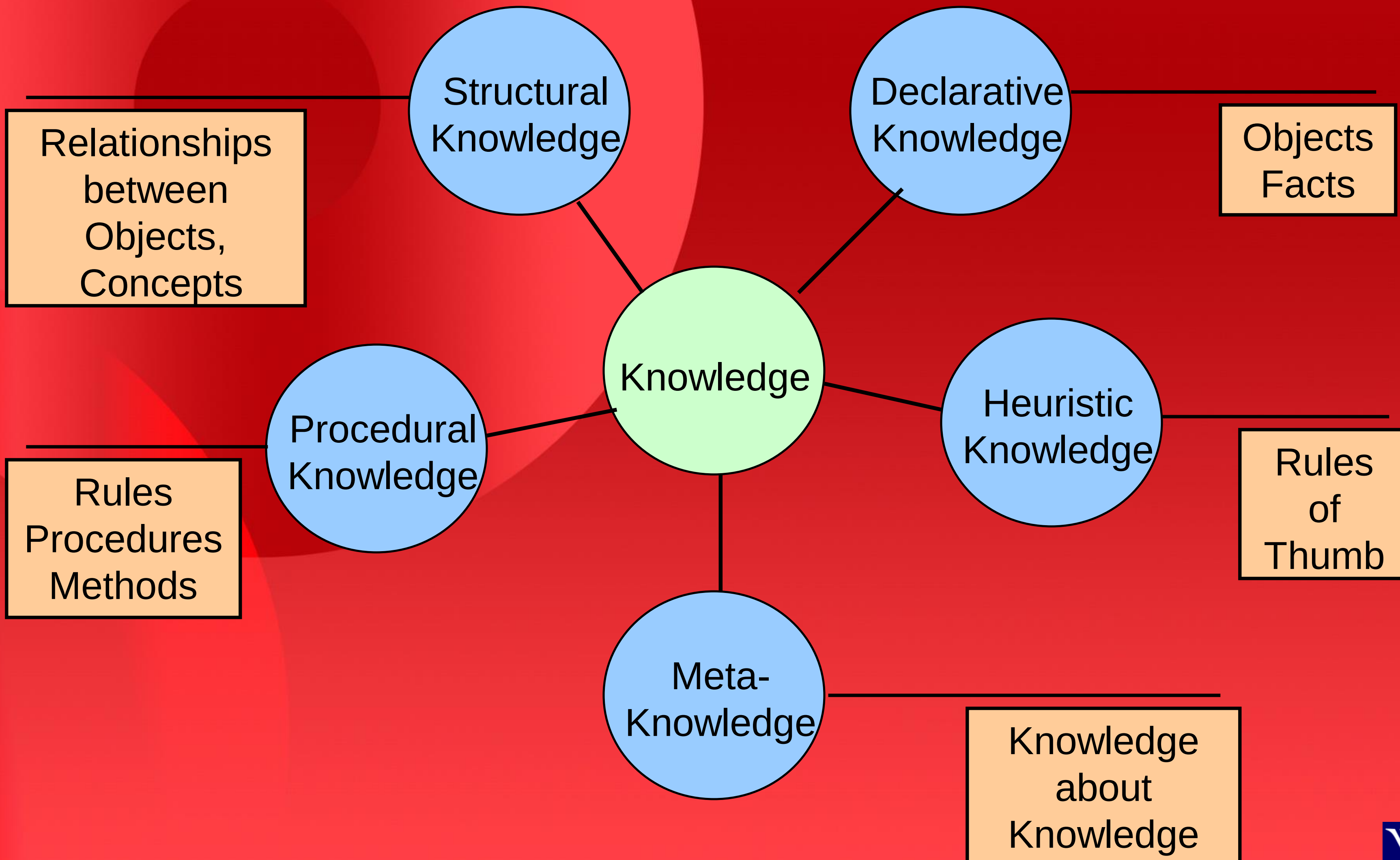
What is Knowledge

- “Understanding of a subject area” Durkin
- Domain: A well-focused subject area

Types of Knowledge

- **Procedural knowledge**: Describes how to do things, provides a set of directions of how to perform certain tasks, e.g., how to drive a car
- **Declarative knowledge**: It describes objects, rather than processes. What is known about a situation. e.g. it is sunny today, cherries are red
- **Meta knowledge**: Knowledge about knowledge, e.g., the knowledge that blood pressure is more important for diagnosing a medical condition than eye color.
- **Heuristic knowledge**: Rule-of-thumb. e.g. if I start seeing shops, I am close to the market.
  - Sometimes called shallow knowledge.
  - Empirical as opposed to deterministic
- **Structural knowledge**: Describes structures and their relationships. e.g. the how the various parts of the car fit together to make a car, or knowledge structures in terms of concepts, sub concepts, and objects.

Types of Knowledge



Towards Representation

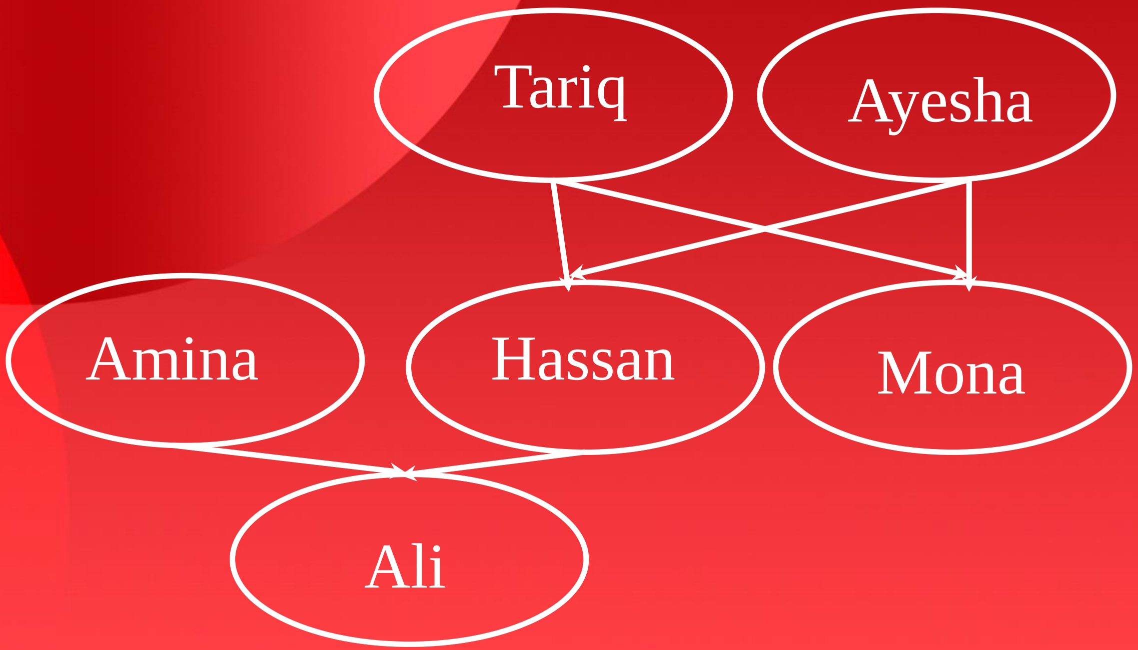
- There are multiple approaches that come to mind
  - Pictures and symbols. This is how the earliest humans represented knowledge when sophisticated linguistic systems had not yet evolved
  - Graphs and Networks
  - Numbers

Representation: Pictures

- What types of knowledge is best represented using pictures? e.g. can we represent the relationship between individuals in a family using a picture?
- To store procedural knowledge, we could use a series of pictures. e.g. how to boil an egg. A **series** of pictures showing the process.
  - Pictures are best suited for recognition tasks
  - Structural information
- How useful is such a representation for a computer?
  - Not very easily translated to useful information because computers cannot interpret pictures automatically.
  - But useful for human understanding because they allow a high level view of a concept to be obtained readily.

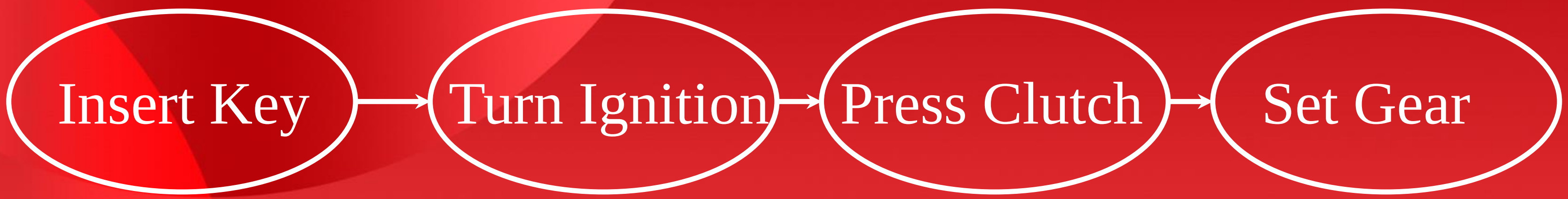
Representation: Graphs & Networks

- Graphs and Networks allow ***relationships*** between entities to be incorporated, e.g., to show family relationships, now we can use a graph.



Graphs and Networks

- May be used to represent procedural knowledge.
- e.g.How to start a car?

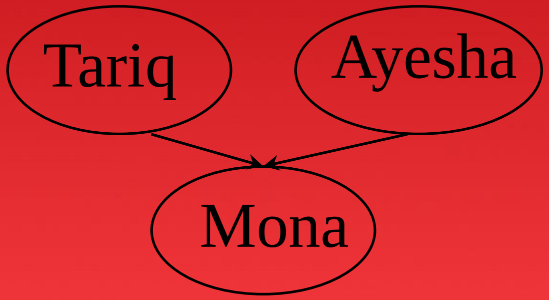


Representation: Numbers

- Numbers are an integral part of the knowledge representation used by humans.
- Translate easily to computer representation.
- Eventually, every representation we use gets translated to numbers in the computers internal representation.

Knowledge of a Family

- By a picture
- By a graph
- Description in words
  - Tariq is Mona's Father
  - Ayesha is Mona's Mother
  - Mona is Tariq and Ayesha's Daughter



Formal KR Techniques

- Now, we will discuss some formal methods of knowledge representation in AI.
- Each method is suited to representing a certain type of knowledge.
- Choosing the proper representation is important because it must facilitate reasoning. As the saying goes ‘Knowledge is Power’.

Facts

- Facts are a basic block of knowledge (the atomic units of knowledge)
- They represent declarative knowledge.
- A Proposition is the statement of a fact. Each proposition has an associated truth value. It may be true or false.
- In AI, to represent a fact, we use a proposition and its associated truth value
- e.g.
  - Proposition A: It is raining
  - Proposition B: I have an umbrella
  - Proposition C: I will go to school