Machine Learning as a Motor for Deep Transformation?

Antoine Cornuéjols

AgroParisTech – INRA   MIA 518

EKINOCS research group

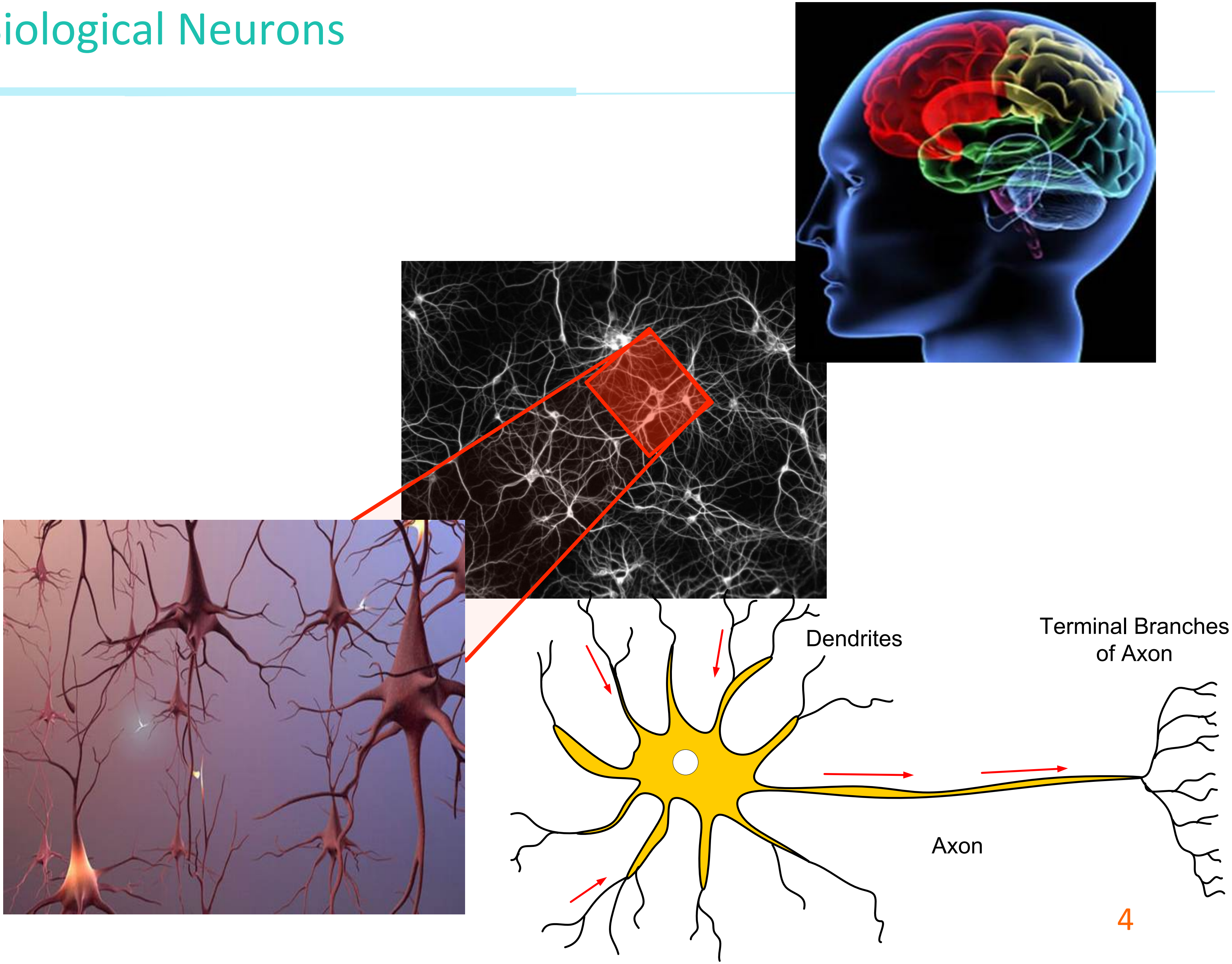
Artificial **Intelligence**

Dream of the pioneers

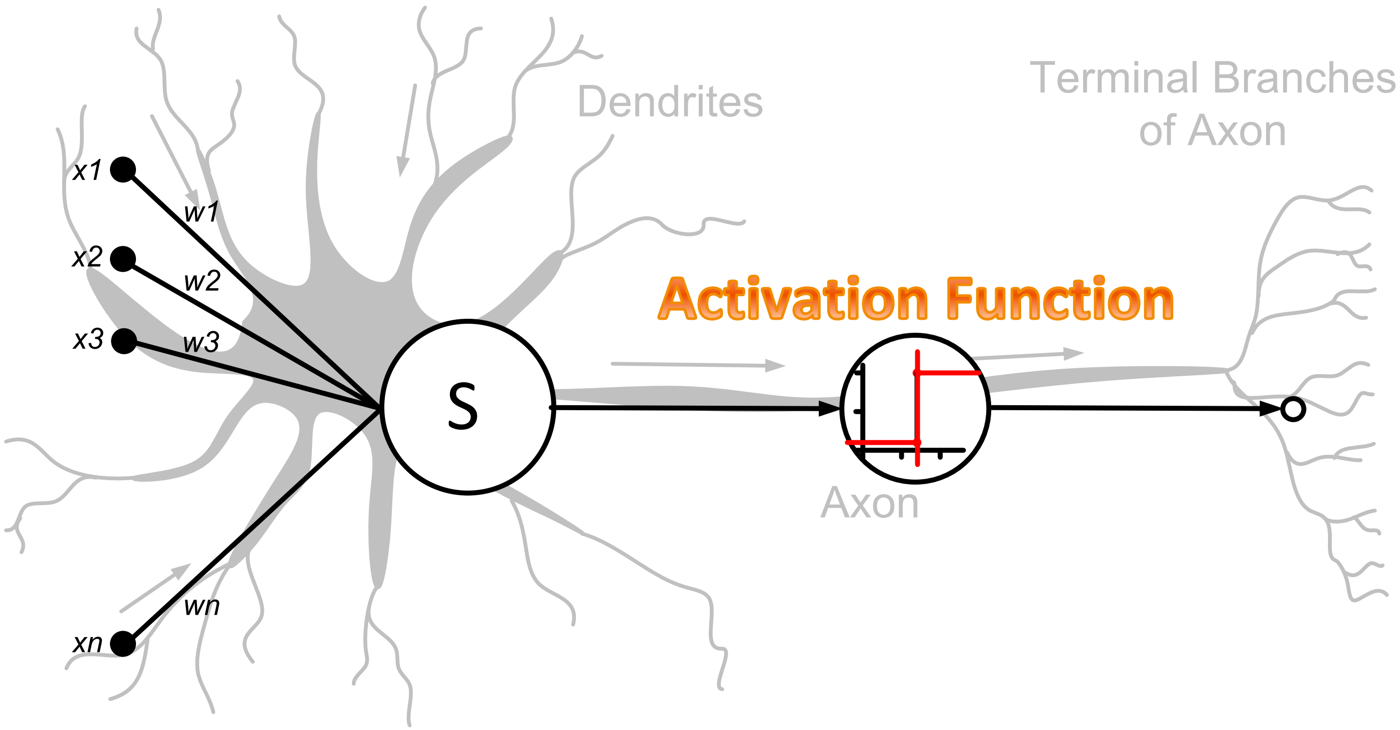
Computation

Information

Biological Neurons



1st formal model of the neuron



McCulloch & Pitts (1943)

Dream of the **pioneers**

1. **Understanding** intelligence
   - **Reasoning**: symbolic AI
   - Inspired by the **brain**

Computation

Information

2. Focused on **human performances**
   - **Playing** chess
   - **Reasoning** like humans: *planning*, *solving problems*, *analogy thinking*, ...
   - **Understanding** texts and discourses
   - Able to express itself using **natural language**

Outline

1. A brief history of AI
2. AI now: the triumph of deep neural networks
3. AI in the near future
4. There are limits
5. The case of XAI
6. Conclusion

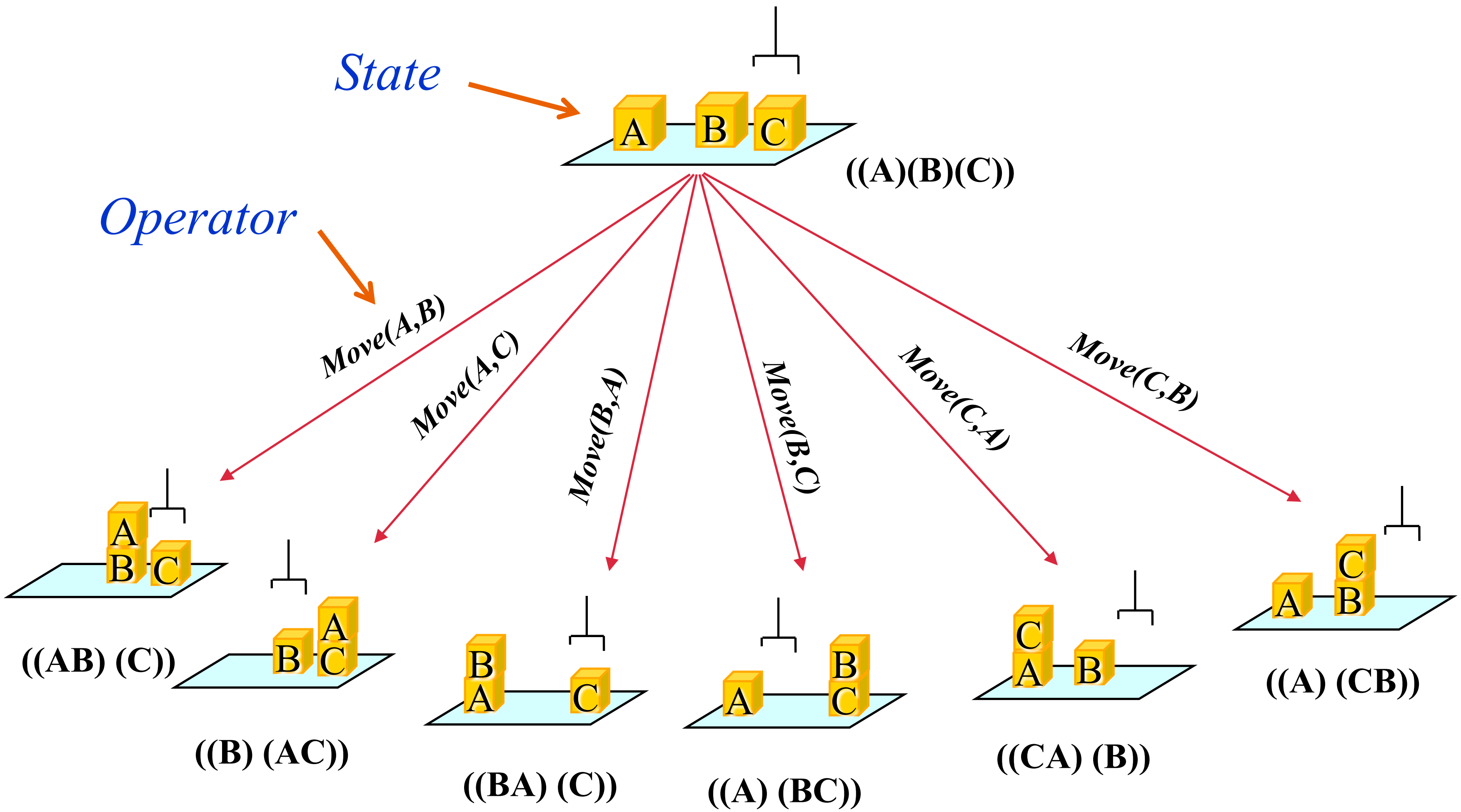
The **assumption**

Intelligence is

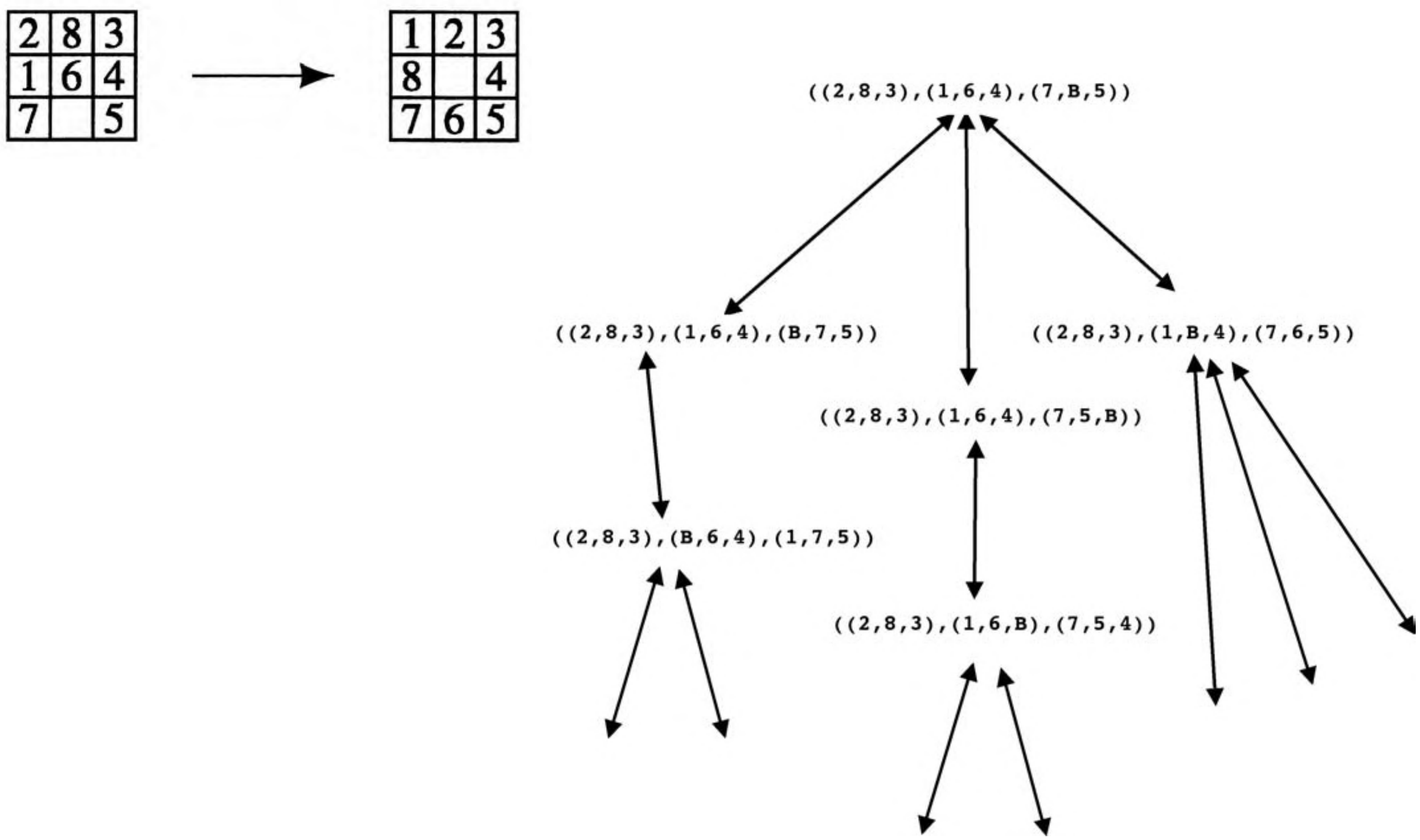
general reasoning processes

(~1956 – ~1969)

Reasoning / problem solving



Reasoning / problem solving



- Search in a graph

The first mobile robot: Shakey (SRI)

Vision

+ planning

+ interface through pseudo natural language

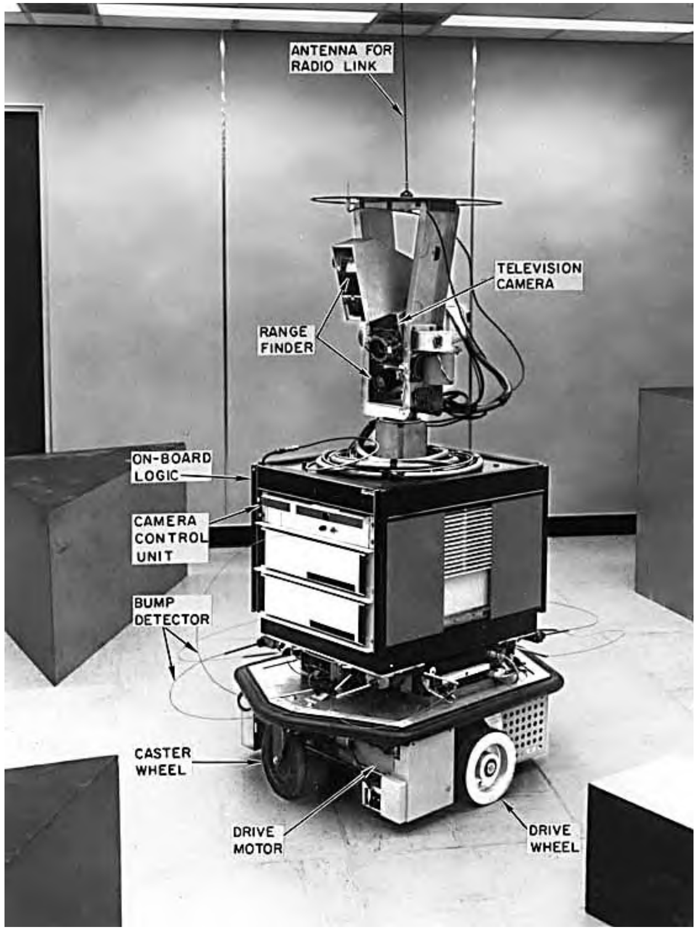
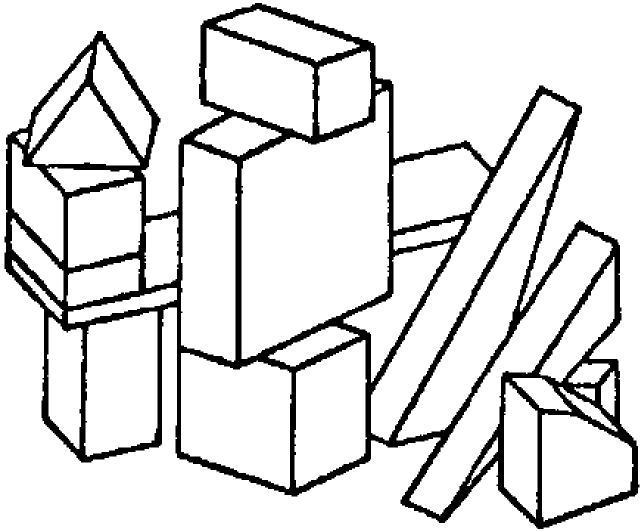


Figure 12.3: Shakey as it existed in November 1968 (with some of its components labeled). (Photograph courtesy of SRI International.)

Machine Vision

- Stanford AI Lab

Figure 8.1: Site of the Stanford AI Lab from 1966 until 1980. (Photograph courtesy of Lester Earnest.)



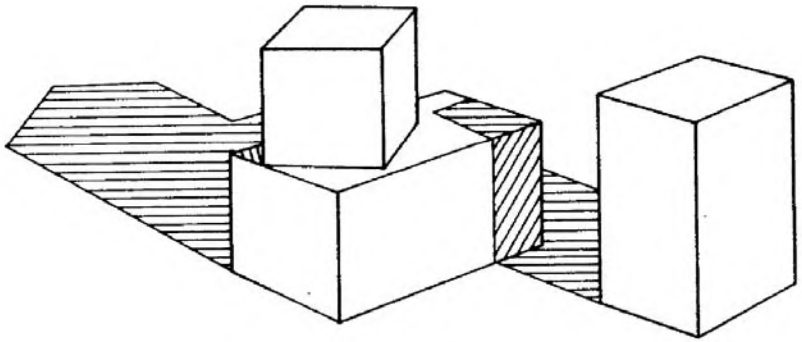


Figure 9.15: A scene with shadows analyzed by Waltz's program. (Illustration used with permission of David Waltz.)

Interface en langage pseudo naturel

```
1.  pick up a big red block.
In order to pick up the red block, it had to clear it
off by finding a space for the green one and moving the
green one away.

OK.
The system answers "OK" when it carries out a command.

2.  grasp the pyramid.
There are three pyramids in the scene, so the program
does nothing.
I DON'T UNDERSTAND WHICH PYRAMID YOU MEAN.
3.  find a block which is taller than the one you are
holding and put it into the box.

The system has to figure out what is meant by "one" and
"it," as well as using deduction to decide which block
it is holding.
BY "IT," I ASSUME YOU MEAN THE BLOCK WHICH IS TALLER
THAN THE ONE I AM HOLDING.
Printed out to keep us informed.
The hand has to set down the red block to pick up the
taller blue one.
OK.
4.  what does the box contain?
THE BLUE PYRAMID AND THE BLUE BLOCK.
```

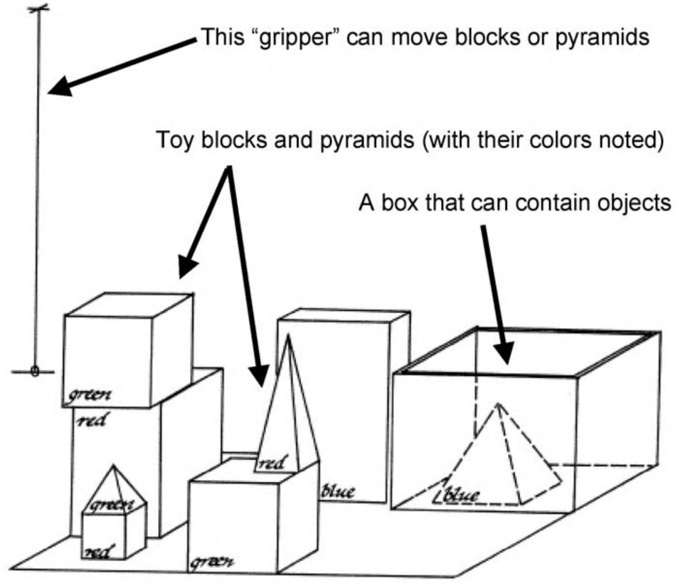
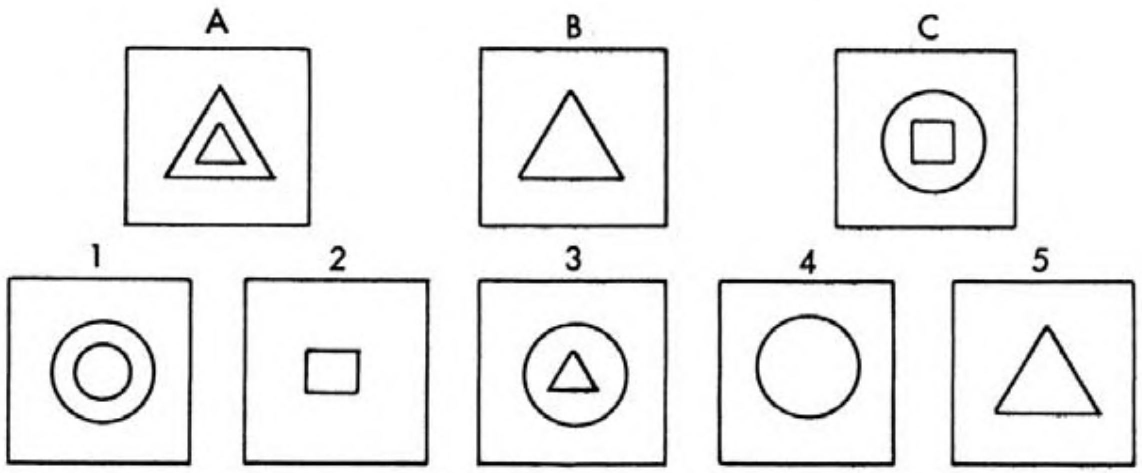


Figure 13.2: SHRDLU's world. (Illustration used with permission of Terry Winograd.)

1968

- **Analogy** making

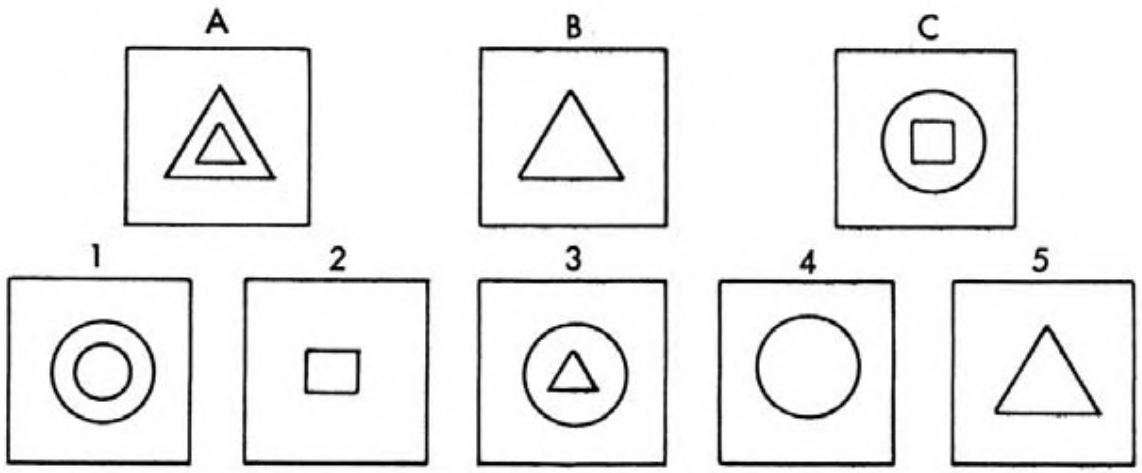


A is to B what C is to ?

1968

- **Analogy** making

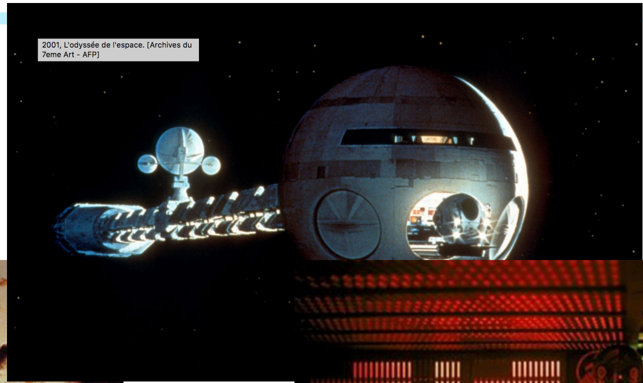
Pb: Find the best matching

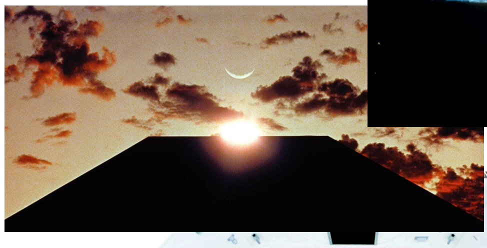


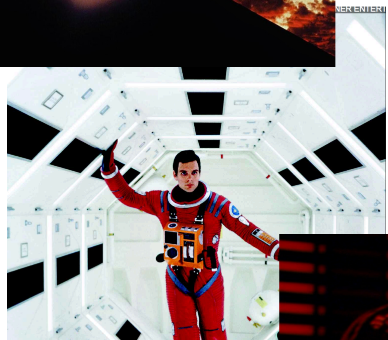
A is to B what C is to ?

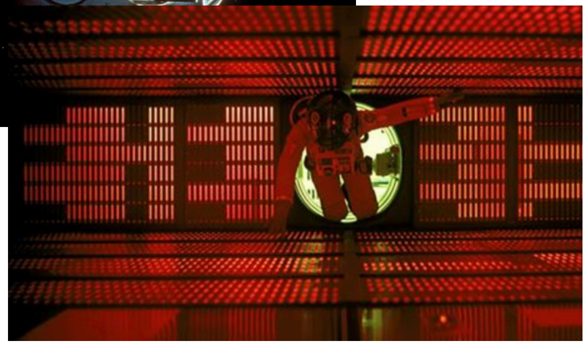
2001: A Space Odyssey

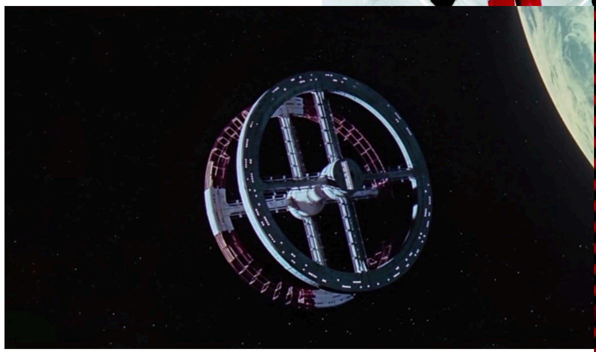
- 1968 …

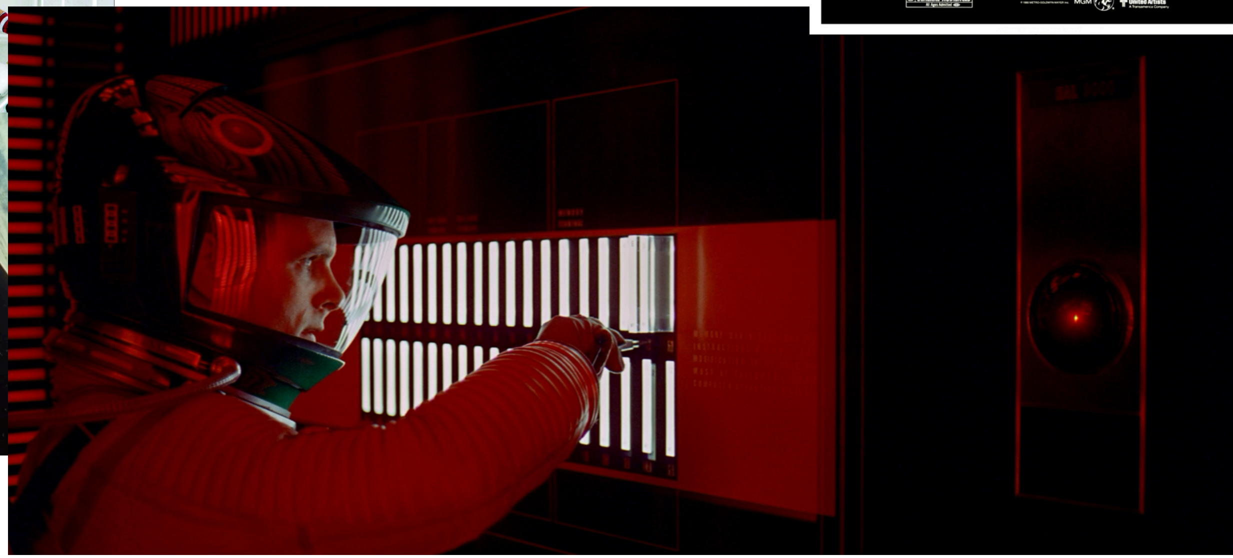






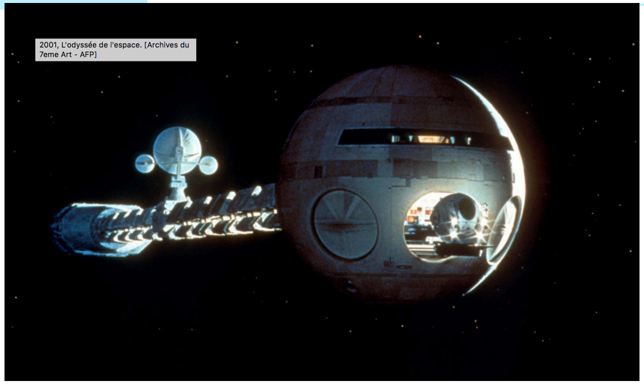


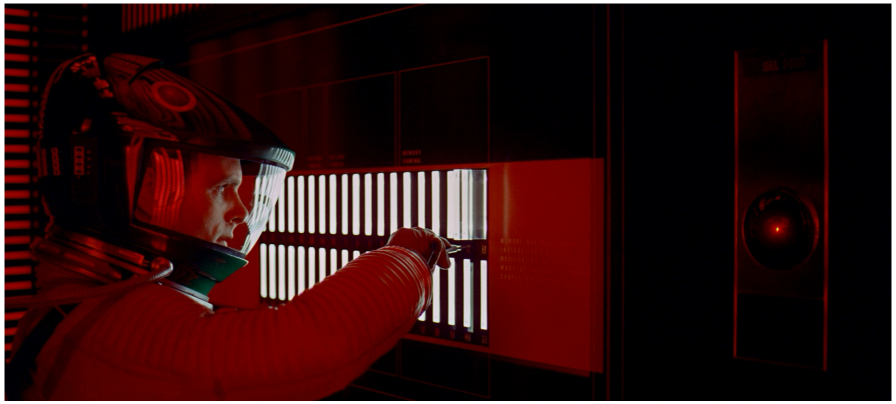




1968 -> 2001: A Space Odyssey

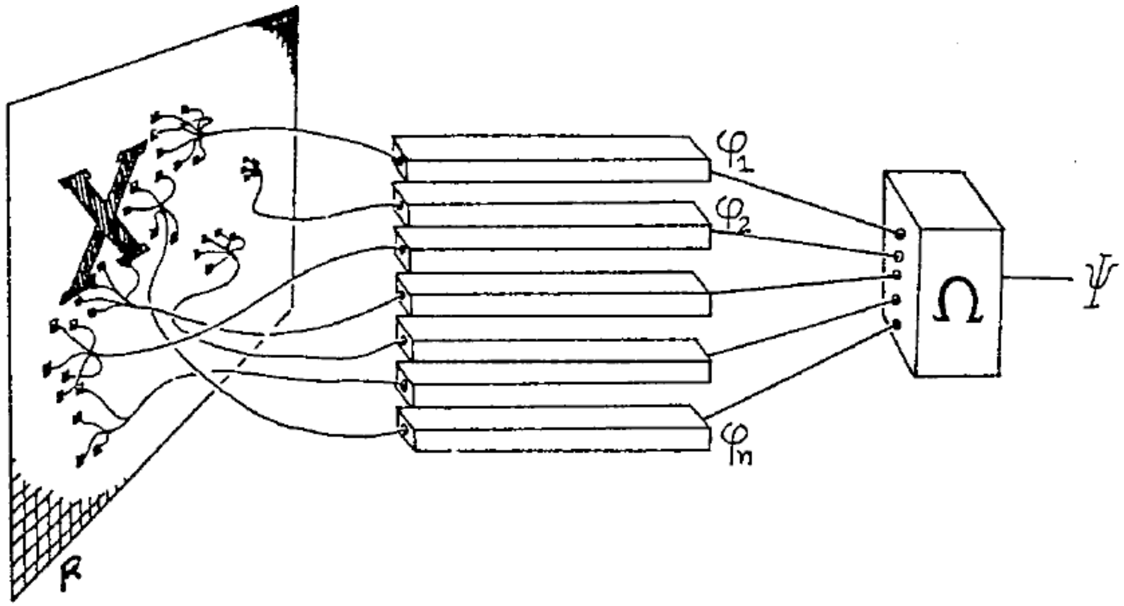
- Vision
- Communication
  - Lips reading
  - Conversation
- Planning complex tasks
- Reasoning
  - Plays (and wins) chess games
- Self-recoding
  - Kills the astronauts
- Emotion
  - Displays fear





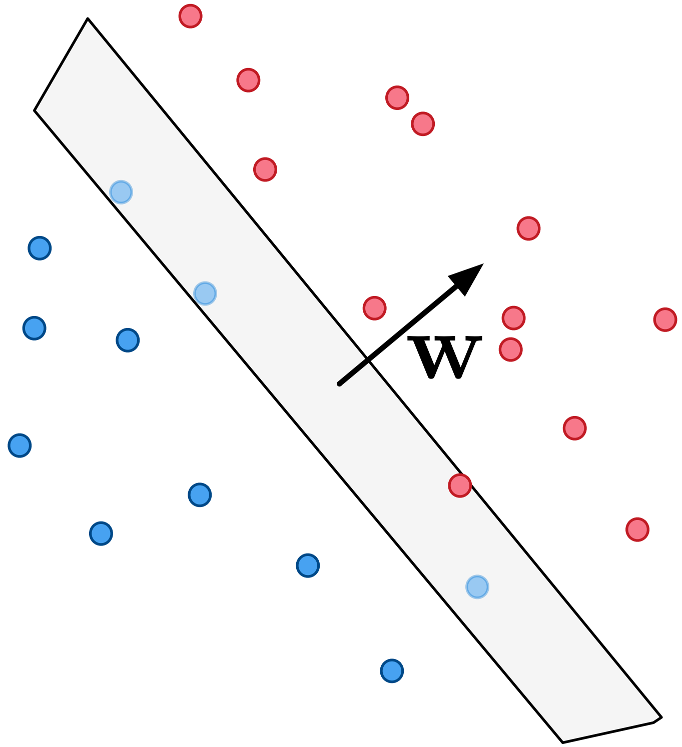
The perceptron

- Frank Rosenblatt (1958 – 1962)



$$\Psi(\mathbf{x}) = \sum_{i=1}^{n} w_i \, \phi_i(\mathbf{x})$$

The perceptron: a linear discriminant



But ... there are limits

Experts are **expert in their own domain**,

but not on **all domains**

Second assumption

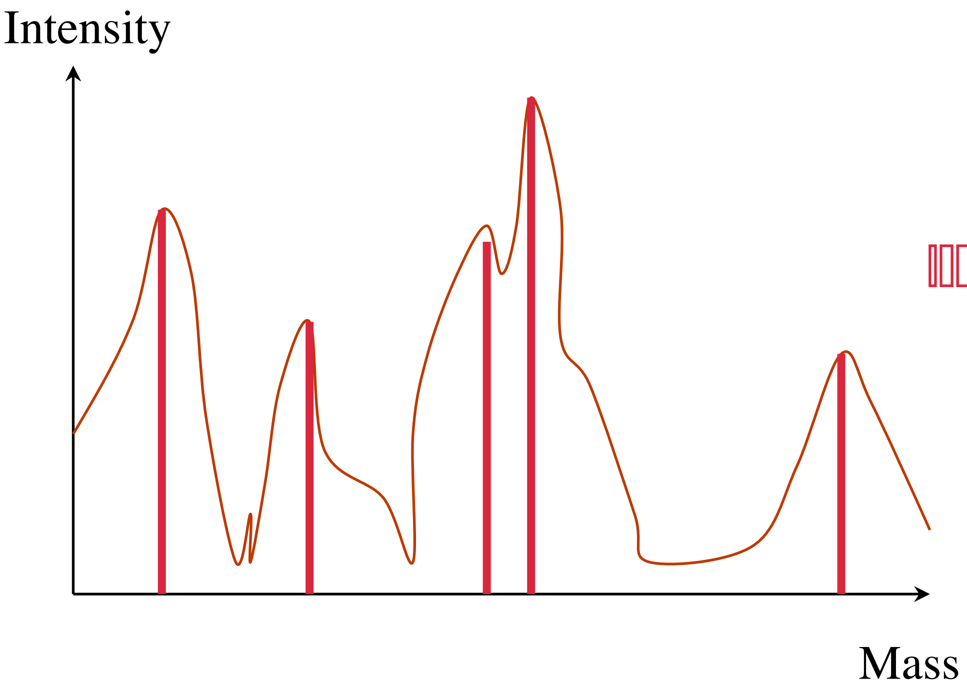
Knowledge is **power**

(~1970 – ~1985)

Expert Systems: DENDRAL

- A project of the NASA:
- Is there life on Mars?
- Mass spectrography

How does an expert performs this?



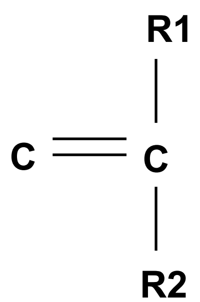
The developped formula of the molecules

- Examples of a piece of knowlege
  - Rule:

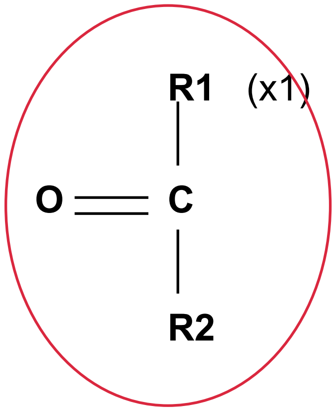
    If the spectrum of the molecule has two peaks x1 et x2 such that:

    1. x1 - x2 = M + 28
    2. x1 - 28 is a high peak
    3. x2 - 28 is a high peak
    4. At least one of the peaks x1 et x2 is high

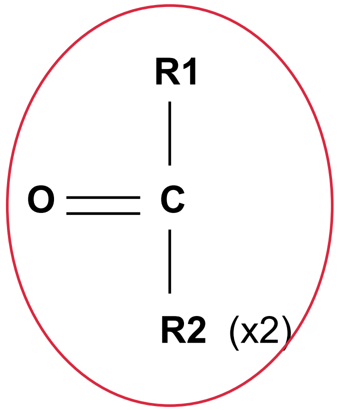
    Then the molecule contains a cetone group



breaks down in:



or in :



Expert Systems: MYCIN

MYCIN: My therapeutic recommendation will be first choice: to cover <1 2 3 5 6>, prescribe Gentamicin, dose 119 mg 98 h IV, for 10 days;Pour couvrir le point 4, prescrire de la Clindamycine, dose 595 mg 96 h IV, pendant 14 jours.

Do you want to eliminate certain medications?

Doctor: Yes, Clindamycine.

MYCIN: To cover <2 3 4 5 6> give Chloramphenicol, dose 1 g 96 h IV, for 14 days;

** Monitor the patient's white blood cell count.

To cover point 1, give gentamicin, dose 119 mg 98 h IV, for 10 days.

** Modify the dose in case of rhenal weakness.

Explanations in Expert Systems

- Mycin: Is meningitis a hospital acquired illness?
- Doctor: Why?
- Mycin: **This will aid in** determining the organism that might be causing the infection

  It has already be established that

  - [2.1] the infection that requires therapy is meningitis
  - [2.2] organisms were not seen on the stain of this pending csf culture
  - [2.3] the type of infection is bacterial
  - [2.4] neurosurgery is not one of the relevant items of patient 100
  - [2.5] the age of patient 100 is greater than 25

  therefore if [2.6] the infection was acquired while the patient was in hospital then there is evidence that the organisms that might be causing the infection are

  - E coli(0.75) staphylococcus-coag-pos(0.3)
  - pseudomonas-aeruginosa(0.3)
  - klebsiella-pneumoniae(0.5)

Explanations in Expert Systems

- Doctor: Why didn't you consider the possibility that the organism be a streptocoque?
- Mycin:

  Rule 33 should have been used to identify the organism as streptococcus,

  but the second clause "the morphology is shell-like" was already known to be wrong and rule 33 was never tried.

BORIS [Dyer,82]

"Richard hadn't heard from his college roomate Paul for years.

Richard had borrowed money from Paul which was never paid back.

But now he had no idea where to find his old friend. When a letter finally arrived from San Francisco, Richard was anxious to find out how Paul was."

~1970 ... ~1985

- "Knowledge is power"

Everything is OK

  - Explicit Reasoning
  - Explications
  - Dialogue

BUT ...

The truth…

- Necessitates Knowledge Engineering
  - A **painful** and **lengthy** process
  - Quite **artisanal**
  - **Difficult to maintain** Expert Systems

Third assumption (~1985 - ...)

Intelligence involves a lot of **knowledge**

that can be obtained through ...

Third assumption (~1985 - ...)

Intelligence involves a lot of **knowledge**

that can be obtained through **general learning processes**

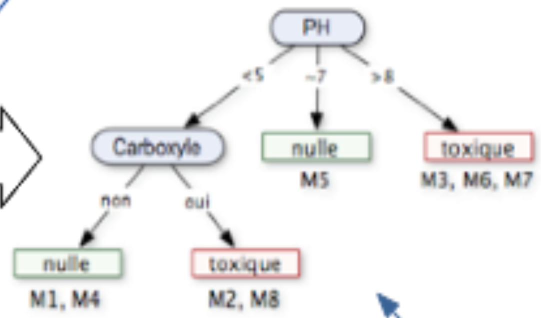
Why not learn everything from data?

Supervised Induction

Descriptors

Labels

	#Cycles	Masse	PH	Carboxyle	Activité
M1	1	faible	<5	non	nulle
M2	2	moyen	<5	oui	toxique
M3	0	moyen	>8	oui	toxique
M4	0	moyen	<5	non	nulle
M5	1	lourd	~7	non	nulle
M6	2	lourd	>8	non	toxique
M7	1	lourd	>8	non	toxique
M8	0	faible	<5	oui	toxique



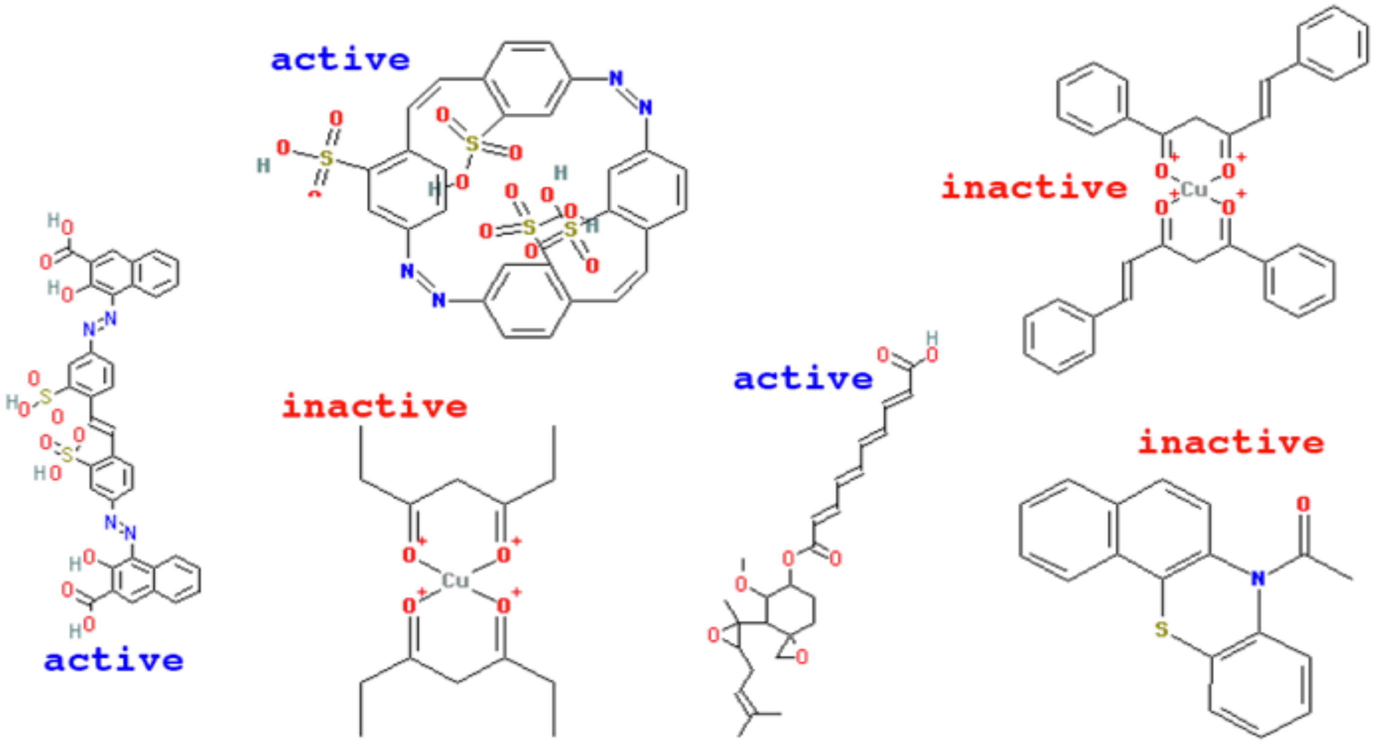
R1: Si pH > 8
alors toxique
R2 : Si pH < 5 & Carboxyle = oui
alors toxique
....

Example

Values of the descriptors

Concept
or
hypothesis

Supervised learning

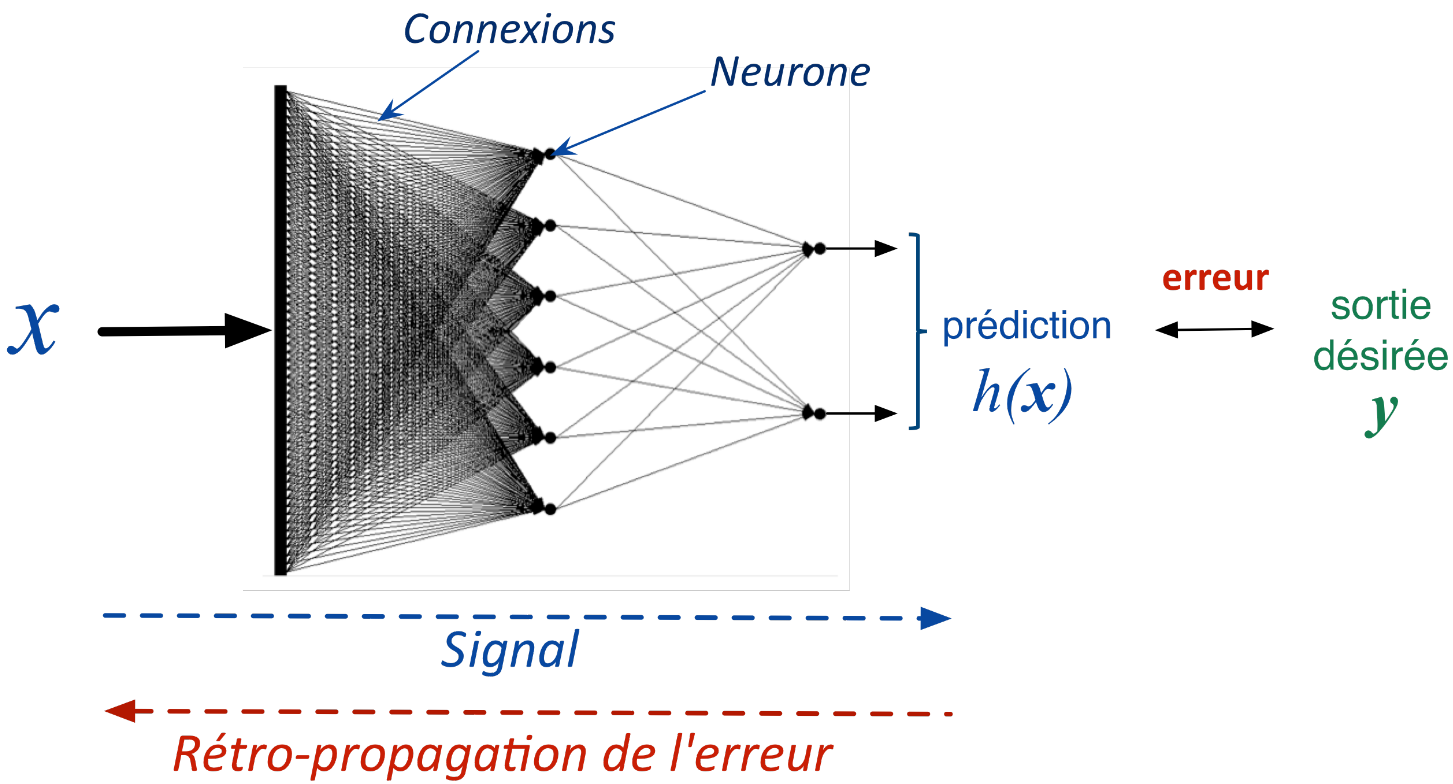


NCI AIDS screen results (from http://cactus.nci.nih.gov).

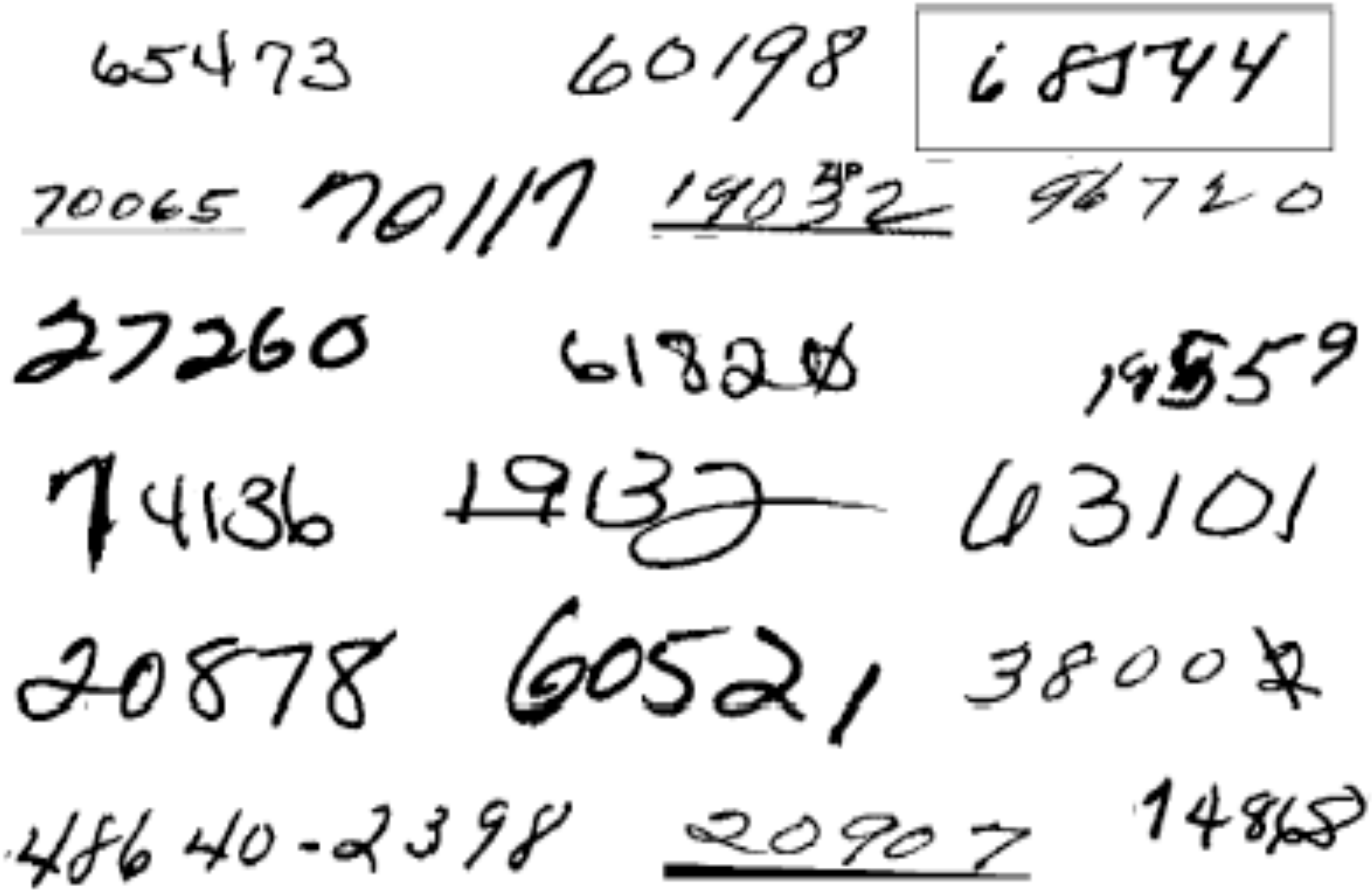
Learning with Multi-Layer Perceptrons

Performs magic!

- Automatically **self-adapt** from the data
- And **resistant** to **noisy data**

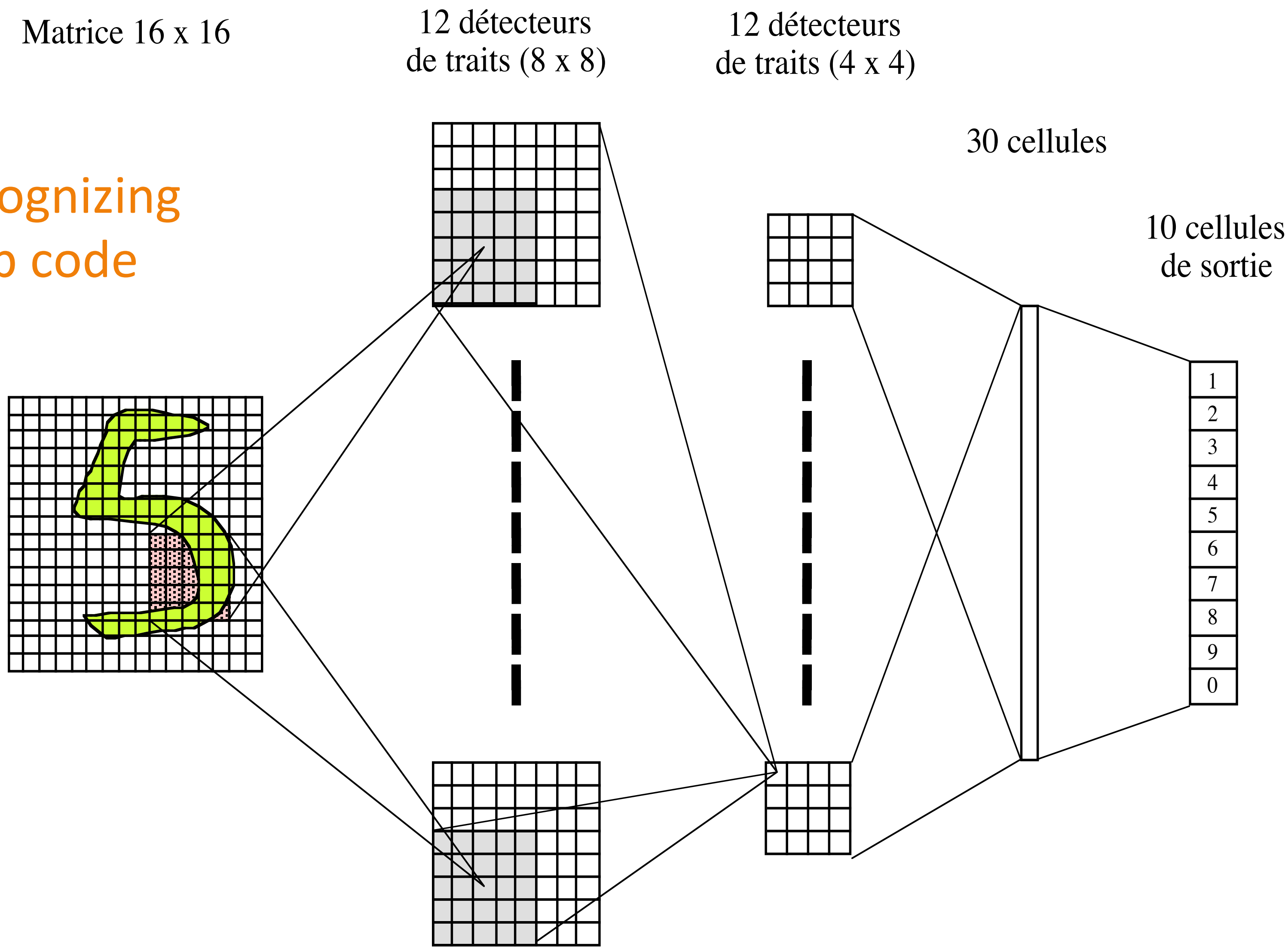


The database



Convolutionnal Neural Networks: the ancestor

Recognizing zip code

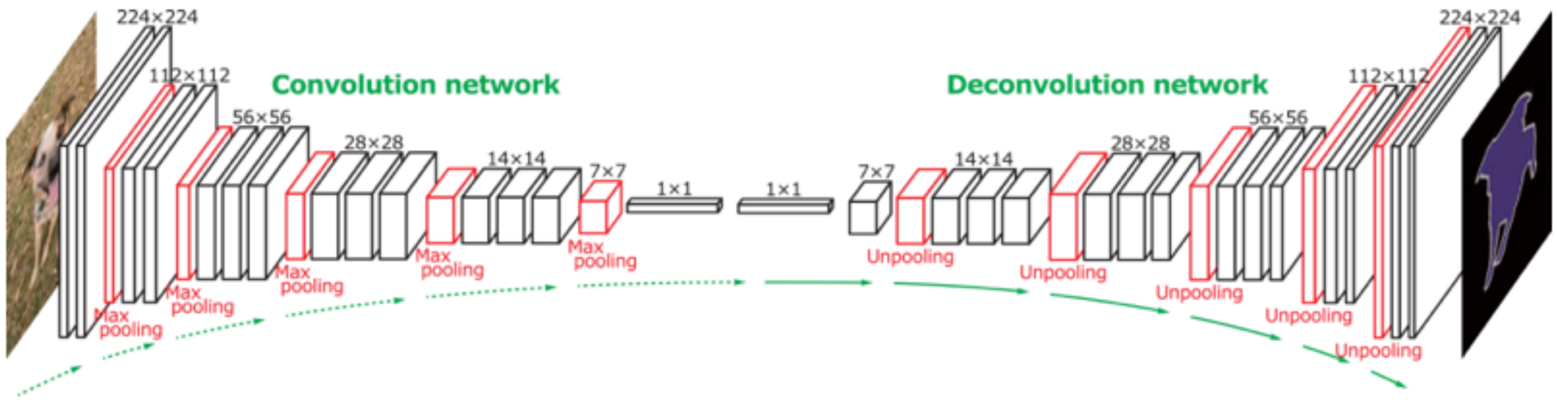


Outline

1. A brief history of AI
2. AI now: the triumph of **deep neural networks**
3. AI in the near future
4. There are limits
5. The case of XAI
6. Conclusion

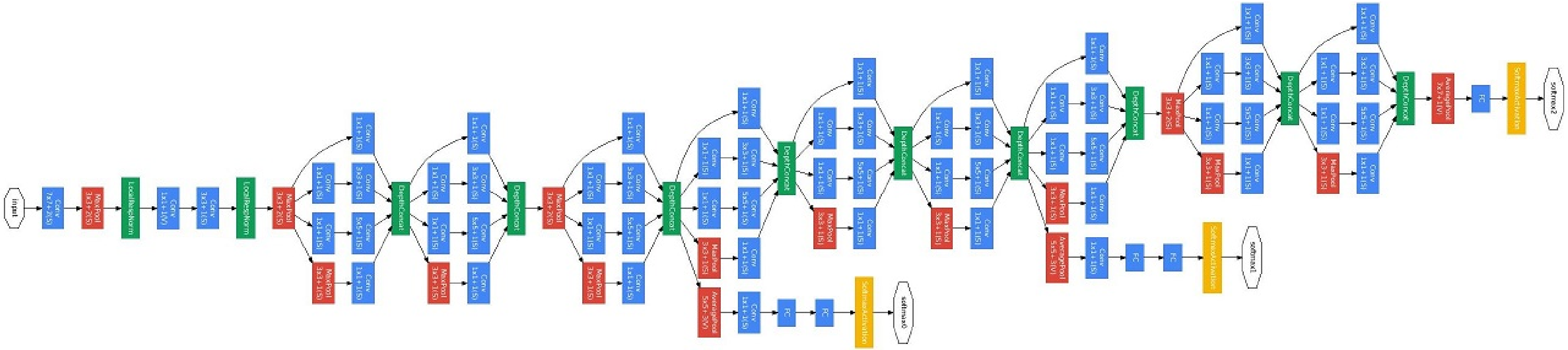
"Deep Neural Networks"

- Artificial Neural Networks with
  - A large number of **layers** (possibly > 100)
  - A very large number of **parameters** ($10^7 - 10^9$ parameters)



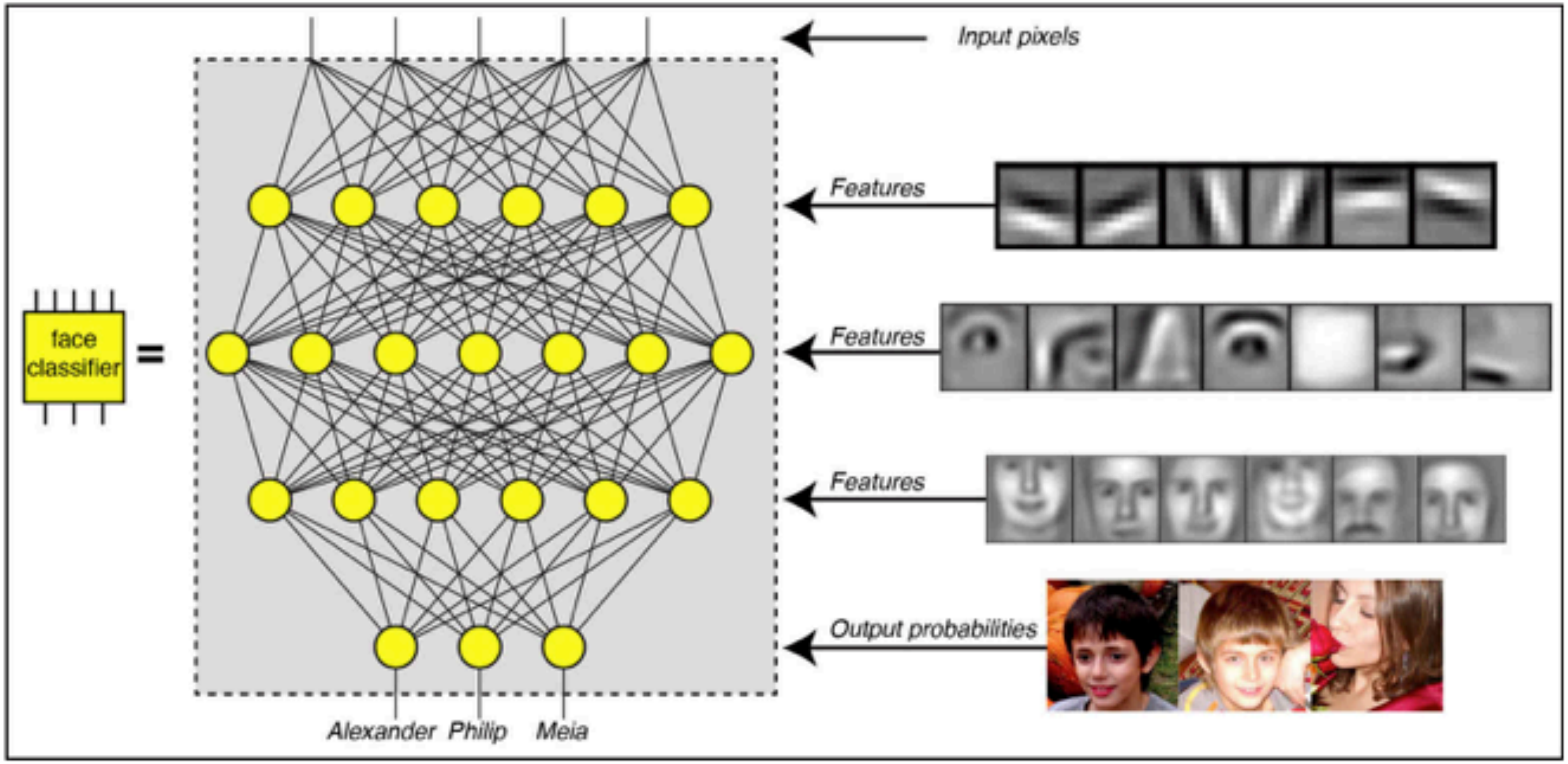
GoogleNet

- Illustration



Face recognition

…

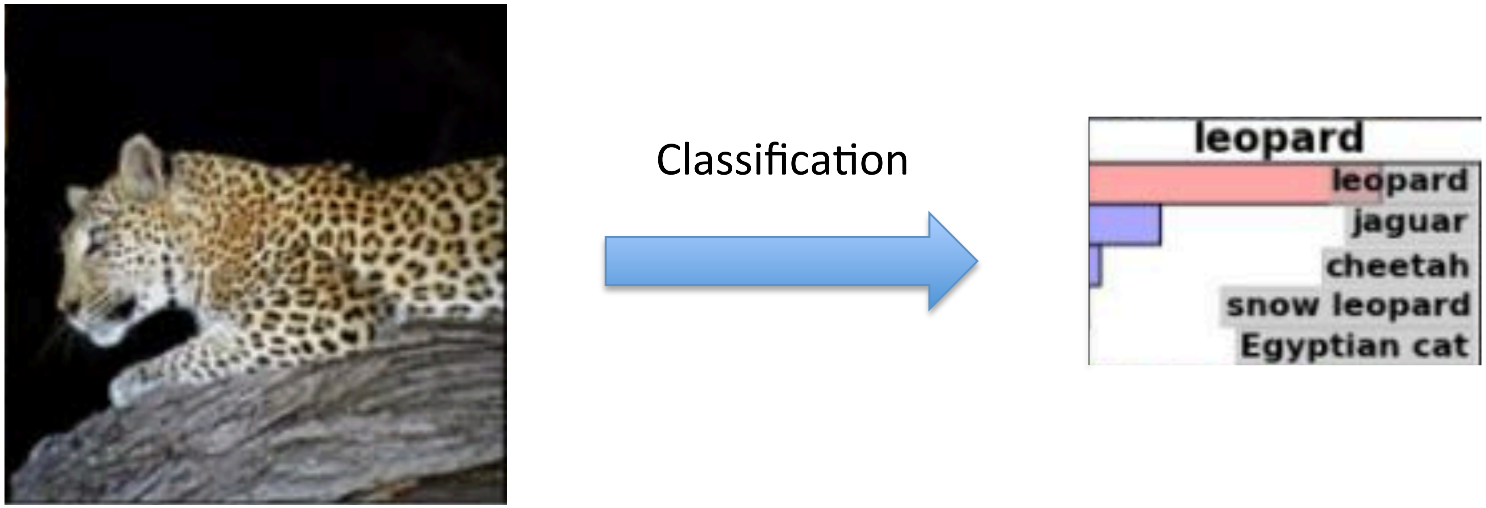


The ImageNet competition

- Over 15M labeled high resolution images
- Roughly 22K categories
- Collected from the Web and labeled by Amazon Mechanical Turk

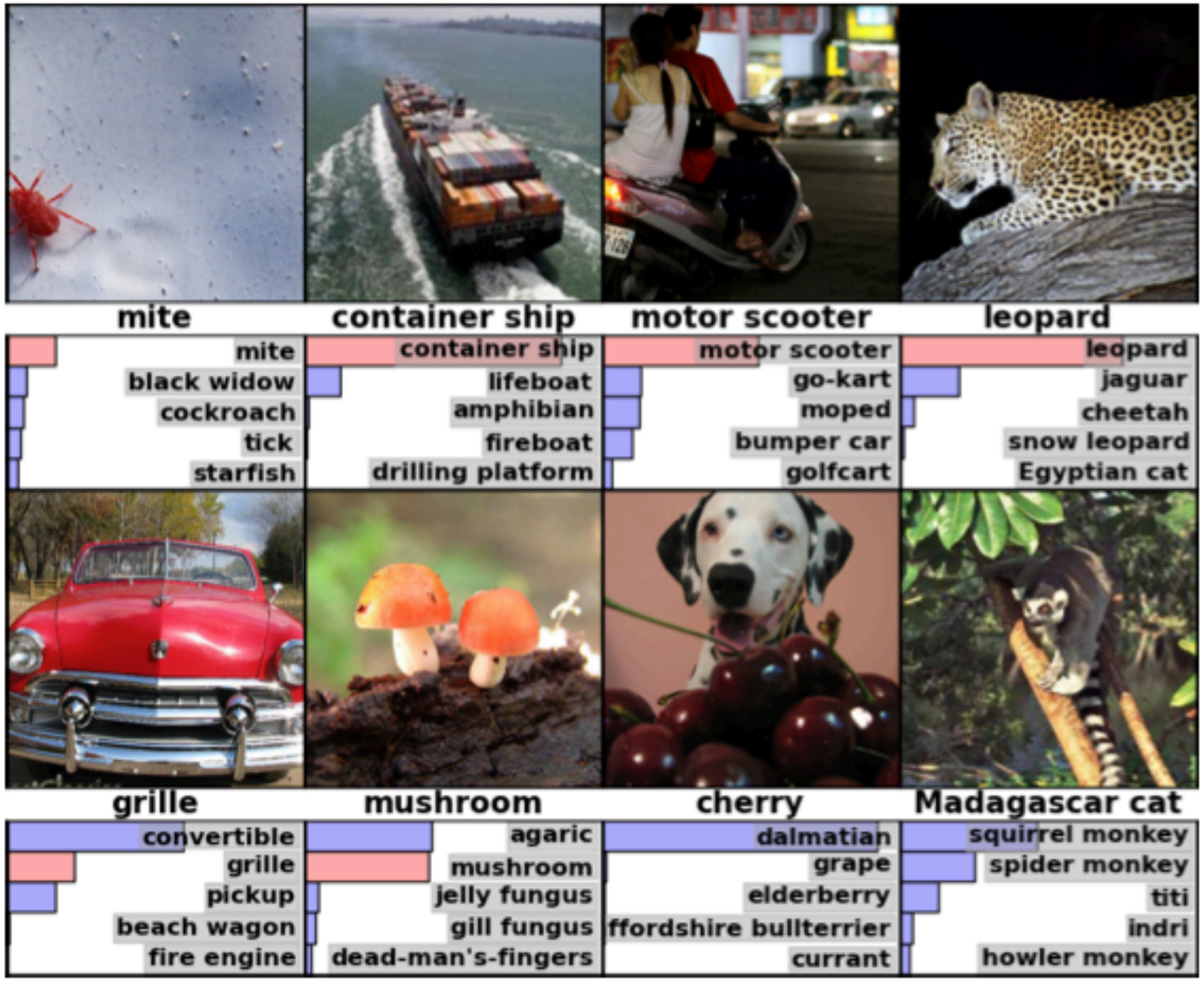
Goal

- Image classification



Results: 8 ILSVRC-2010 test images

- Results



Semantic Image Segmentation

Model trained with a maximum range of 40m and EFS.

- Autonomous vehicles

And YOU?

- Machine **translation**

- Change of **paradigm**

- A set of **new tools**
  - **Data analysis** (e.g. neural networks)
  - **Simulation** (e.g. Multi-Agent systems)
  - New **goals** (e.g. recommendation)

1. **Old** paradigm

- **Construct a hypothesis** (e.g. such and such treatment should have such and such an effect)
- Build an experimental design to **test the validity of the hypothesis**
- The experimental design and the data collected **serve only to test the given hypothesis**

1. **Old** paradigm
   - **Construct a hypothesis** (e.g. such and such treatment should have such and such an effect)
   - Build an experimental design to **test the validity of the hypothesis**
   - The experimental design and the data collected **serve only to test the given hypothesis**

2. **New** paradigm
   - Be “open” minded: we are ready to look for (unexpected) patterns in the mass of available data
   - Infinite re-use of data is possible (even though they were not collected for this specific purpose)

This is « **data mining** »

(Almost) all field are concerned

- Environment
  - Follow the **dynamics** of urban areas, of coastal erosion, of the Artic ocean, ... From satellite images
  - The **climate change**
    - The Harvard forest (Long Term Ecological Project). 1600ha, of which 20 are equipped with electric heaters.
  - Understand the **interplay between species** in an ecosystem
  - What are the **genes** that participate in the resistance to hydraulic stress
- Nutrition
  - What are the **determinant** of our preferences for animal proteins
- Sociology
  - How **rumors** are born and spread

Outline

1. A brief history of AI
2. AI now: the triumph of deep neural networks
3. AI in the near future
4. There are limits
5. The case of XAI
6. Conclusion

Where one speaks of "data flood"

- Our data is captured in abundance whenever we go **on Internet**
  - Which sites are visited
  - Which time, for how long, the clicks, what has been bought, ...
- **Smartphones**
  - **Location** even when you did not agree
  - A lot of apps **full of curiosity**
- Connected **Bracelets**
- Means of **payment** (bank)
- Sensors in **vehicles** (insurance)
- Linky meters

« data flood » in the field

- Agriculture
  - Sensors in the **field**
  - Sensors in the **soil**
  - Sensors on **animals**, in the farm
  - **Drones**
  - Data on the **local markets** (e.g. in India)
  - Data on the **stock markets**
  - Metereological data
  - Data on the **social networks**: producers and consumers
  - **Cold chains** and **distribution**

The world is **yours**

AI + Internet of Things

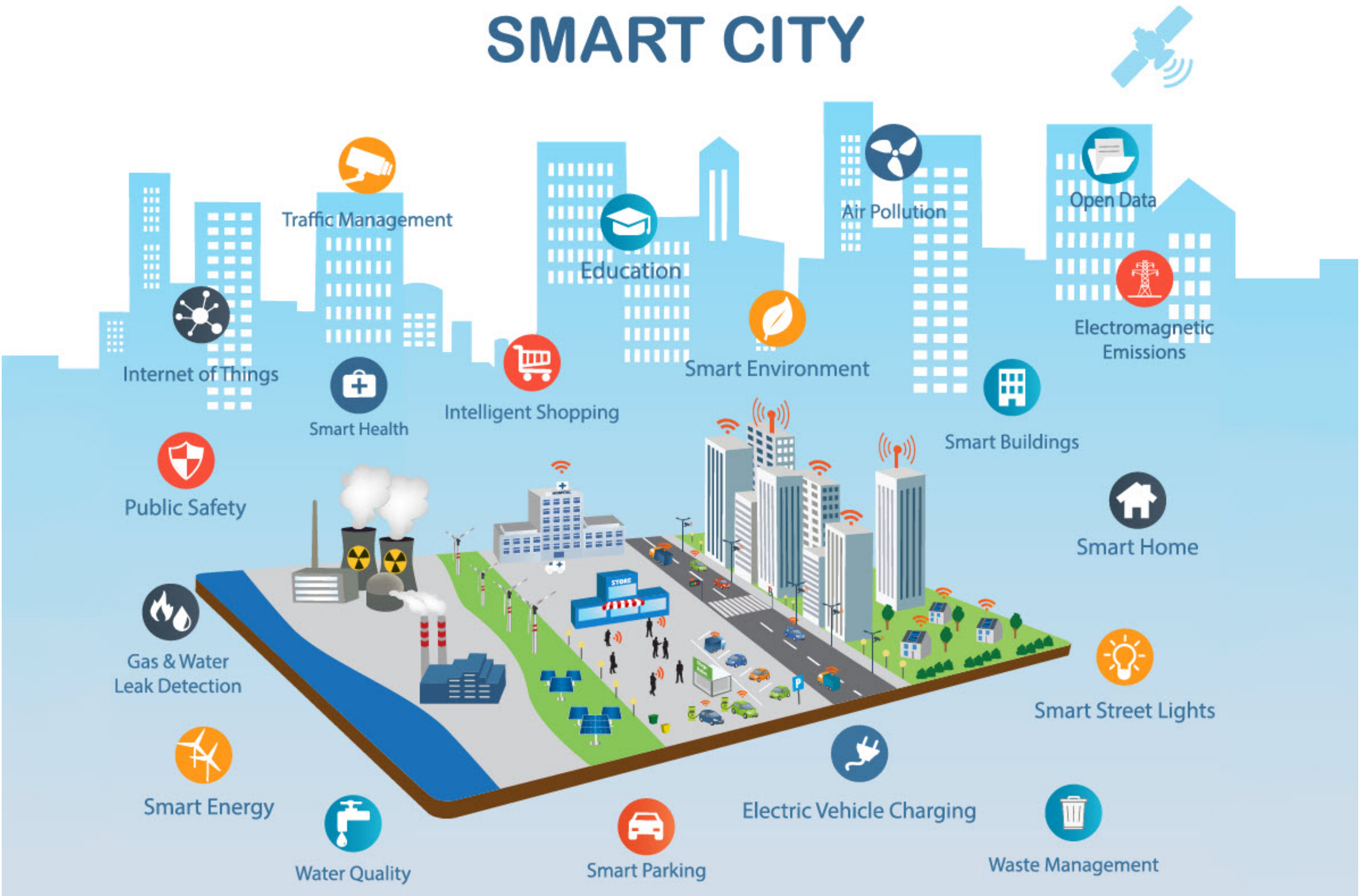
- It cares for you ... or so it seems
  - Sensors, cameras, smartphones, car, ... EVERYWHERE and ALL THE TIME

- Scenario
  - You enter your local supermarket. You are recognized by the camera or thanks to your smartphone. An automatic concierge greets you:
  - "Hi, Mr. Smith, I understand that you're wife's birthday is coming up. We know she loves Napa wines. We've just got a shipment of some fantastic Napa wines, ..."
  - "We can also recommend you some travel place for your next vacation ..."

Completely **fluid** and **targeted to you**

The world is yours

...



Other AI goodies

- Personal **assistant**
  - Help you **plan** your next holiday vacation
  - Help you **optimize** your revenue declaration
  - Help you **choose** the best meal
- Personal assistant **for scientists**
  - A **super Mathematica**
  - **Alpha fold**: discovering the 3D conformation of proteins
- **Specialized** devices
  - "be my eyes" for the blind and vision impaired people

Outline

1. A brief history of AI
2. AI now: the triumph of deep neural networks
3. AI in the near future
4. There are limits
5. The case of XAI
6. Conclusion

One example can says a lot

- Examples are described by:

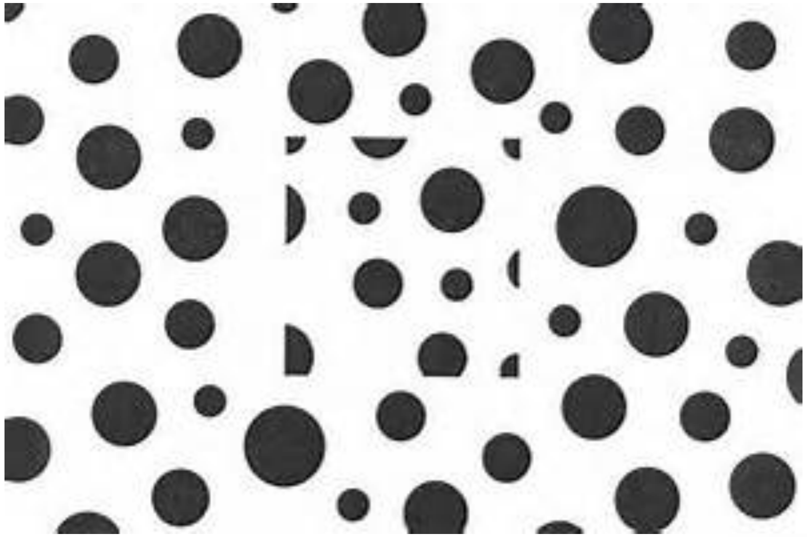
  Number (1 or 2); ***size*** (small or large); ***shape*** (circle or square); ***color*** (red or green)

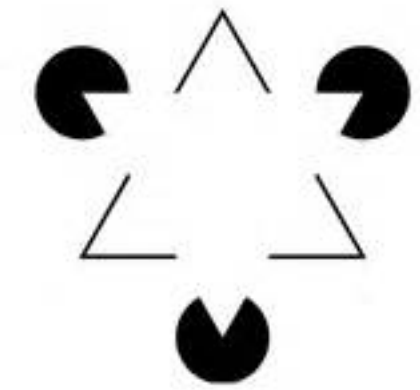
Description	**Your** prediction	**True** class
1 large red square		-
1 large green square		+
2 small red squares		+
2 large red circles		-
1 large green circle		+
1 small red circle		+

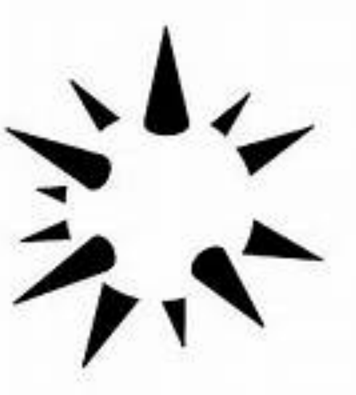
How many possible functions altogether from X to Y ? $2^{2^4} = 2^{16} = 65{,}536$

How many functions do remain after 6 training examples? $2^{10} = 1024$

Learning is induction







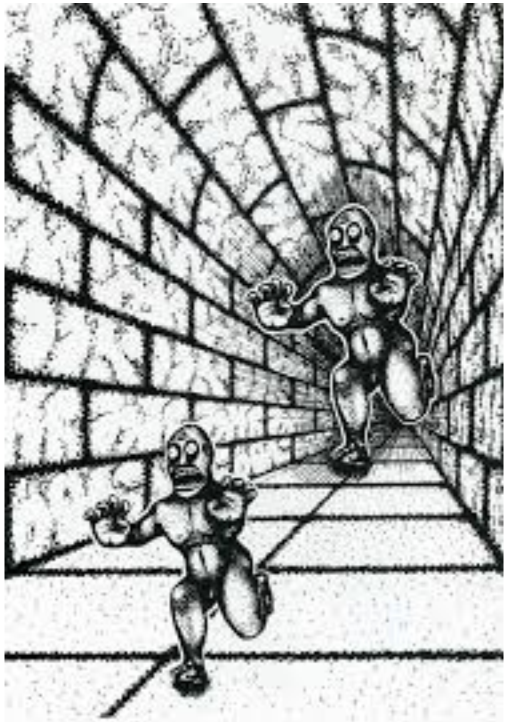
...

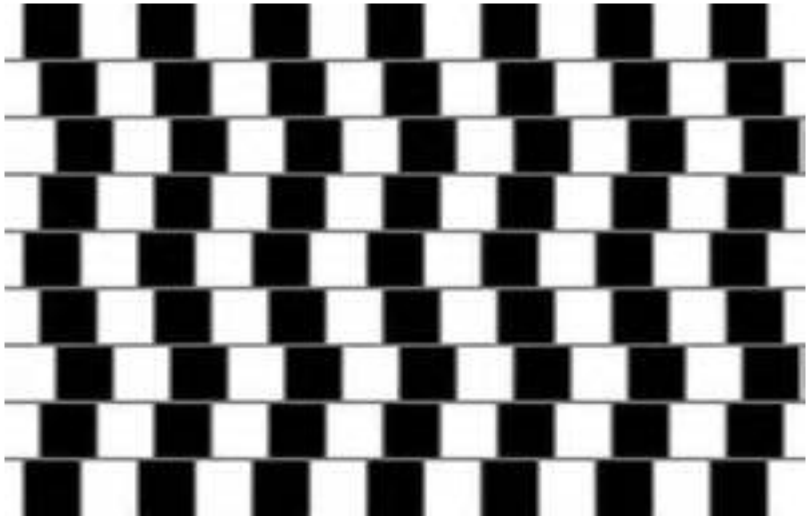
Learning is induction

- There are ambiguities

Learning is induction

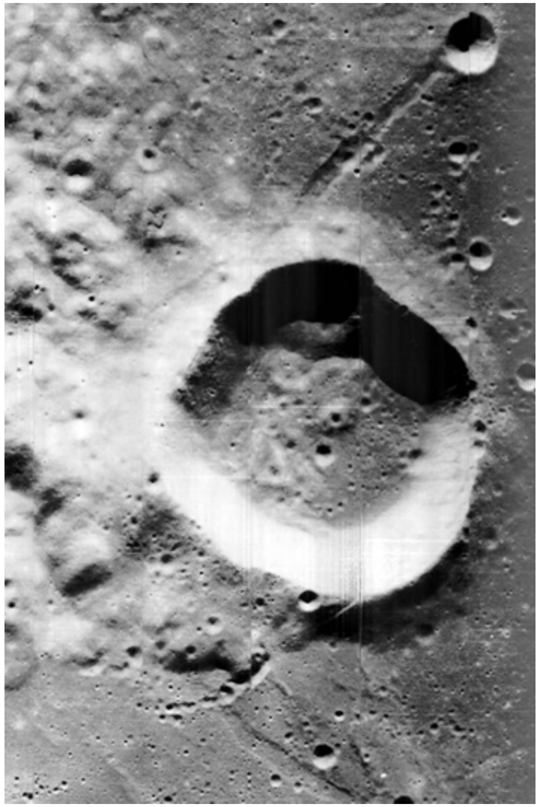
- ... therefore fallible

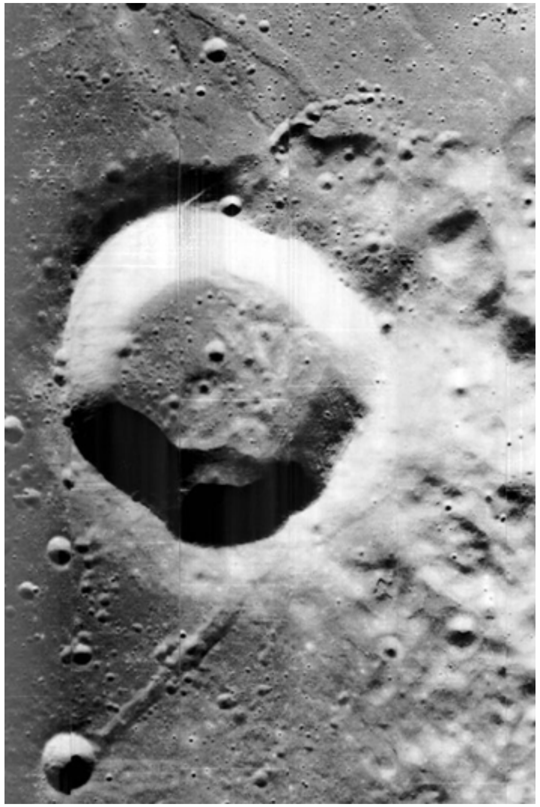




Learning is induction

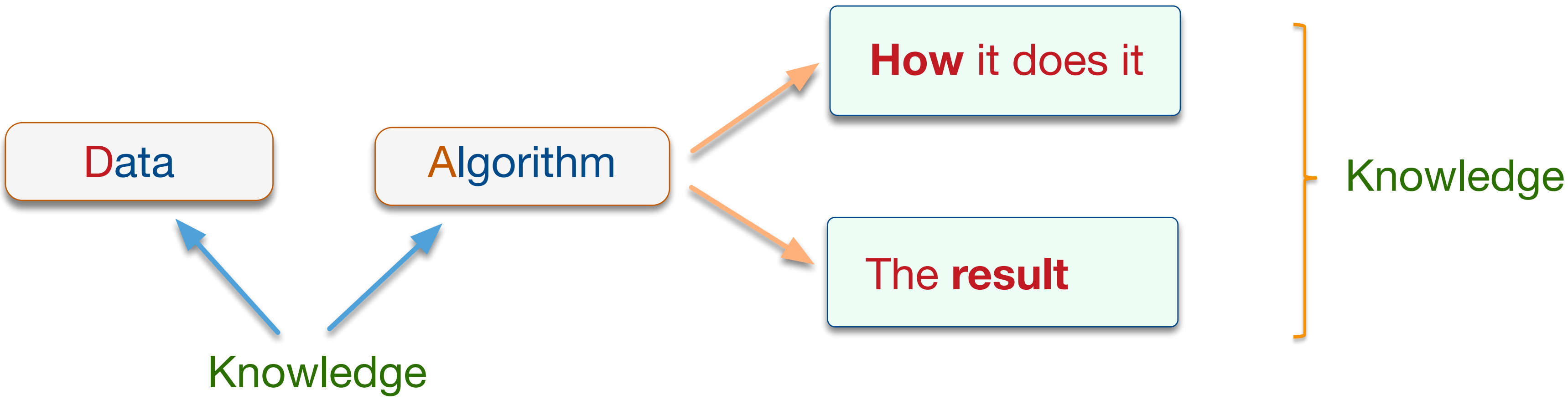
- There are uncertainties





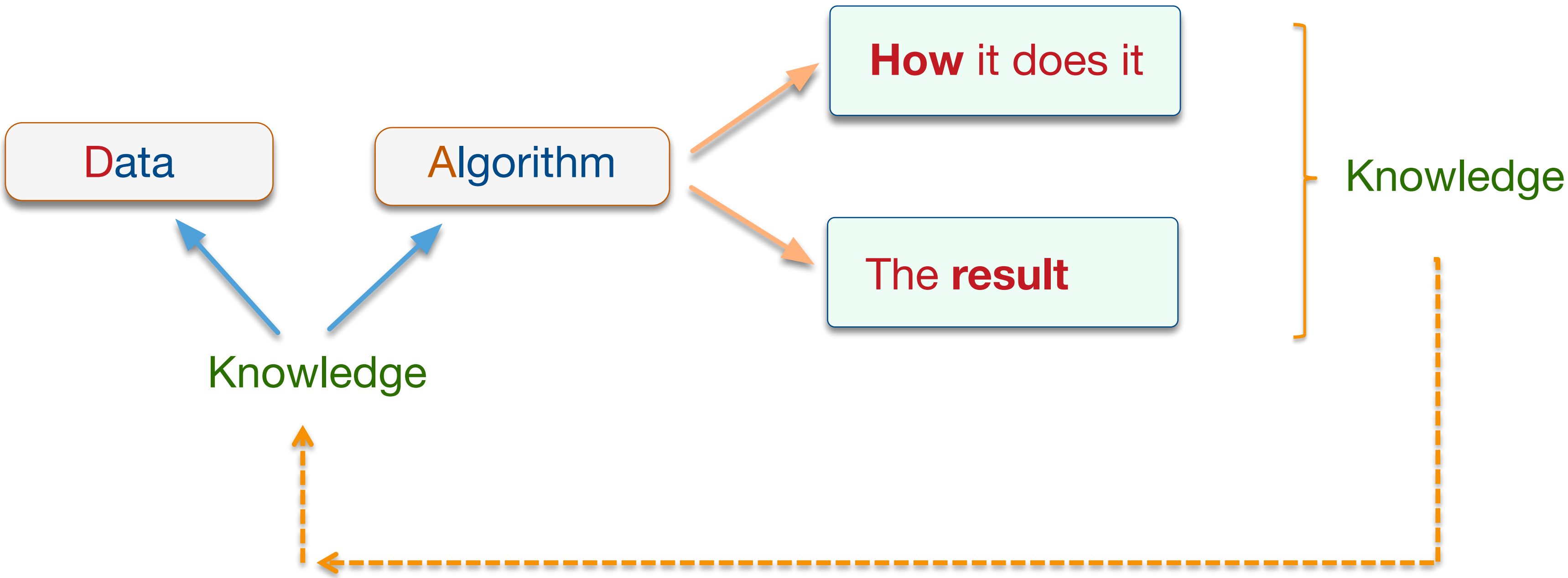
Crater *or* hill?

Inductive learning: what **it does**



...

Inductive learning: what **it does**



...

Induction is a **risky** business

1. You have to **invest a lot**

2. And **be very careful** about the yield

Machine Learning DOES NOT produce absolute truths

Do not give up your **critical sense** at every step!

Machine translation

- Very **impressive** and **useful** (see DeepL)

- But

Le drone volait à une altitude de 30m au-dessus du sol	The drone was flying at an altitude of 30 m $ above the ground

Machine translation

- Very **impressive** and **useful** (see DeepL)

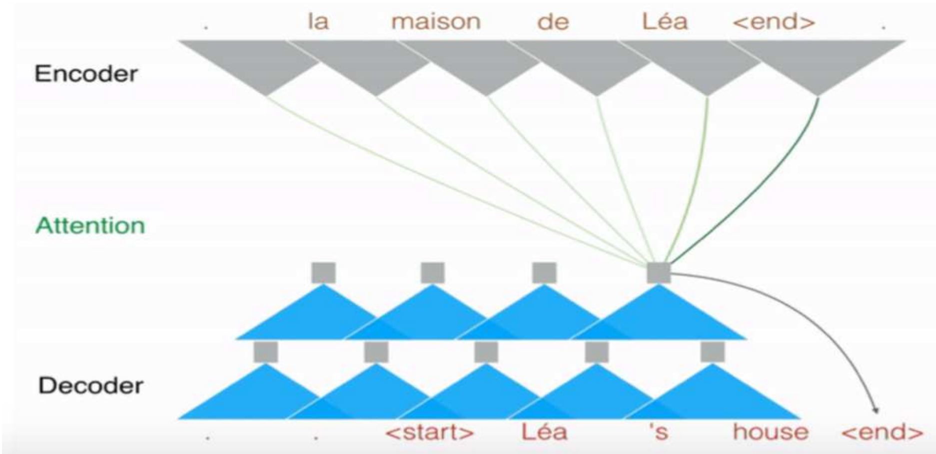
- But

???

Le drone volait à une altitude de 30m au-dessus du sol	The drone was flying at an altitude of 30 m $ above the ground

Machine translation

- Still far from perfect, but …



From Hofstädter (2018)

66

Explanations and deep neural networks

Optical illusions: how to explain them?

Boxer: 0.40 Tiger Cat: 0.18

(a) Original image

Airliner: 0.9999

(b) Adversarial image

!!??

[Selvaraju et al. (2017) « *Grad-CAM: Visual explanations from deep networks via gradient-based localization* »]

Annotation d'images

Figure 2.11: "A group of young people playing a game of frisbee"—that caption was written by a computer with no understanding of people, games or frisbees.

Exemple en médecine

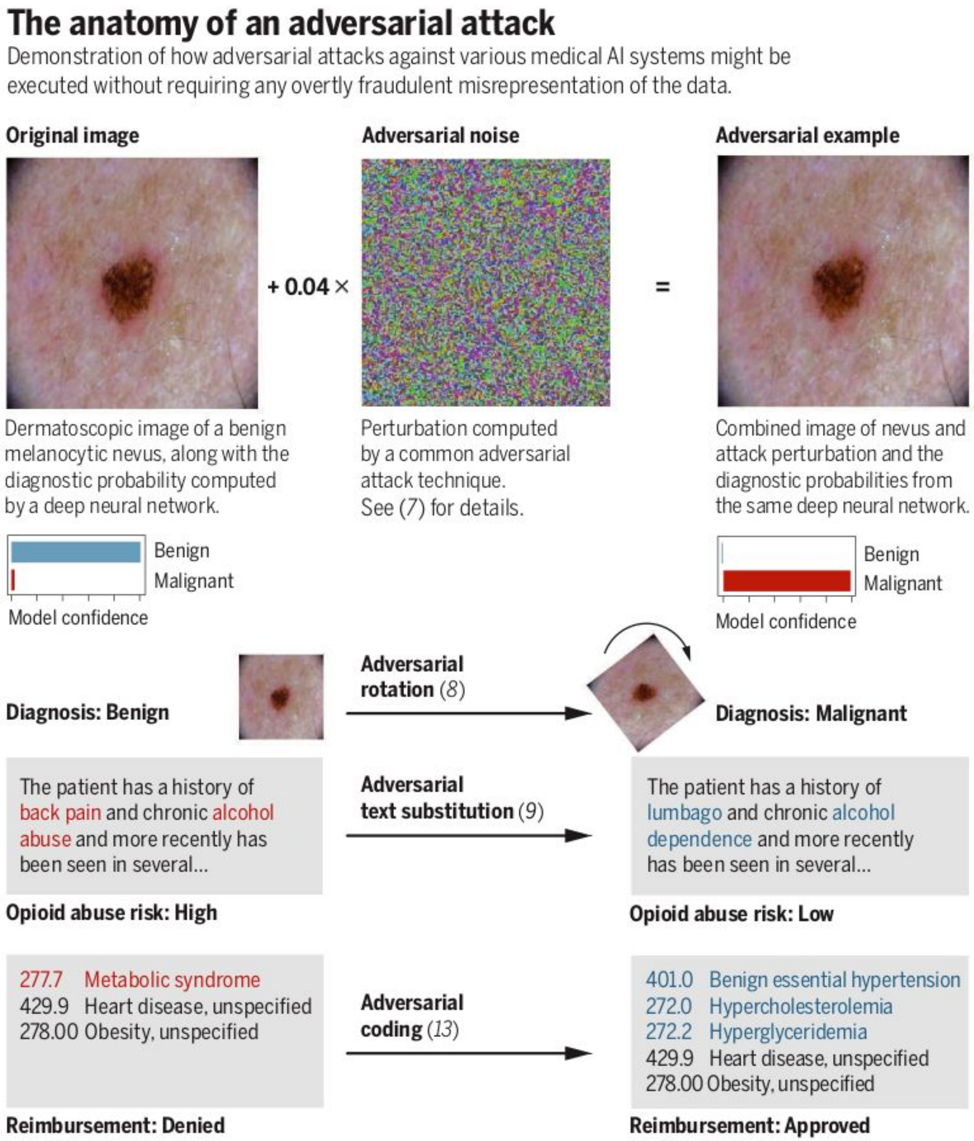
MACHINE LEARNING

Science

Adversarial attacks on medical machine learning

Emerging vulnerabilities demand new conversations

22 March 2019



Car in a swimming pool

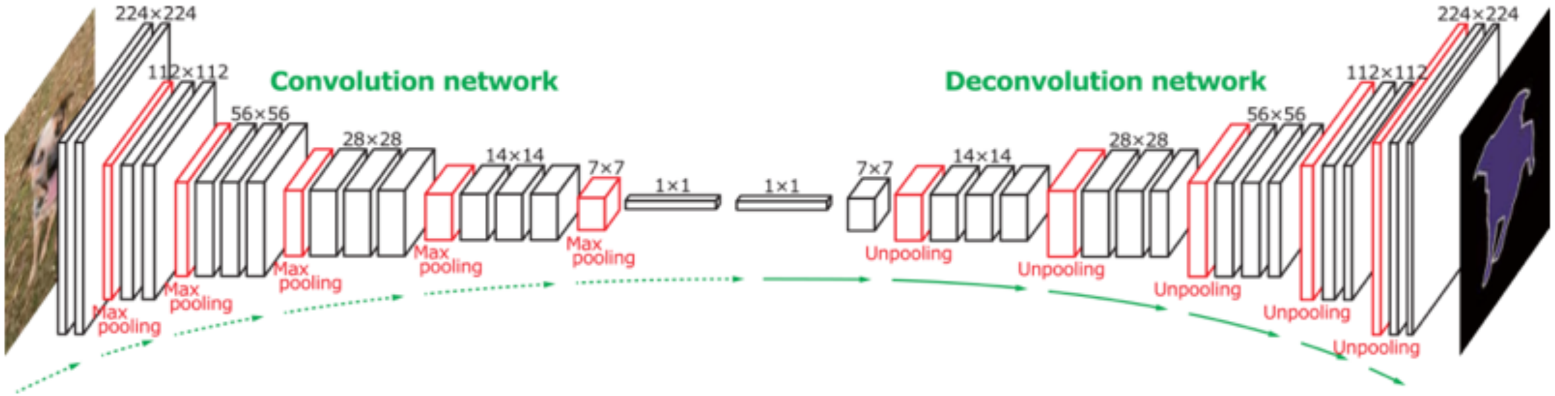
- … or not … ?

Is this less of a car
because the context is wrong?

[Léon Bottou (ICML-2015, invited talk) « *Two big challenges in Machine Learning* »]

Neural networks are “black boxes”

- A very large number of numbers (10^7 – 10^9 parameters)



Outline

1. A brief history of AI
2. AI now: the triumph of deep neural networks
3. AI in the near future
4. There are limits
5. The case of XAI
6. Conclusion

The case AlphaGo

- Plays like an “alien”
- An amazing game
- Revolutionizes the way we play
- Effervescence in go schools

Autonomous vehicle

- The National Highway Traffic Safety Administration (NHTSA) is currently investigating 23 accidents related to Tesla's Autopilot system

- Questions
  - Who is **responsible**?
    - The driver?
    - Tesla (the programmer)?
    - The other person?
  - What is **the reason** for the accident?
    - So as to correct the autopilot system (and systems around)

Problem

- So far efficient predictors are often black boxes

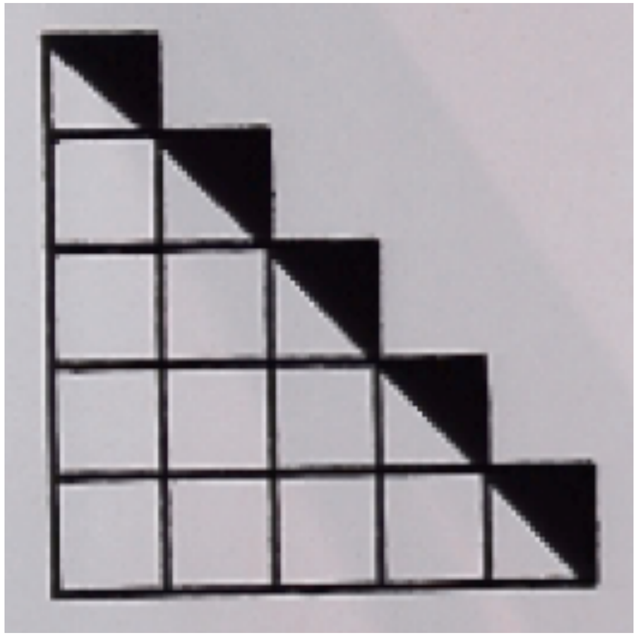
- This is an issue for a number of applications (e.g. in medicine)
  - We want to be able to be **confident** in the system
  - It can justify its **decisions**
  - It can justify its **reasoning**

The ability of providing explanations is **required in Europe** since May 2018 (GDRP, Recital 71)

XAI: Explainable Artificial Intelligence

Lots of types of "explanations"

$$1 + 2 + 3 + \ldots + n \overset{?}{=} \frac{n^2}{2} + \frac{n}{2}$$

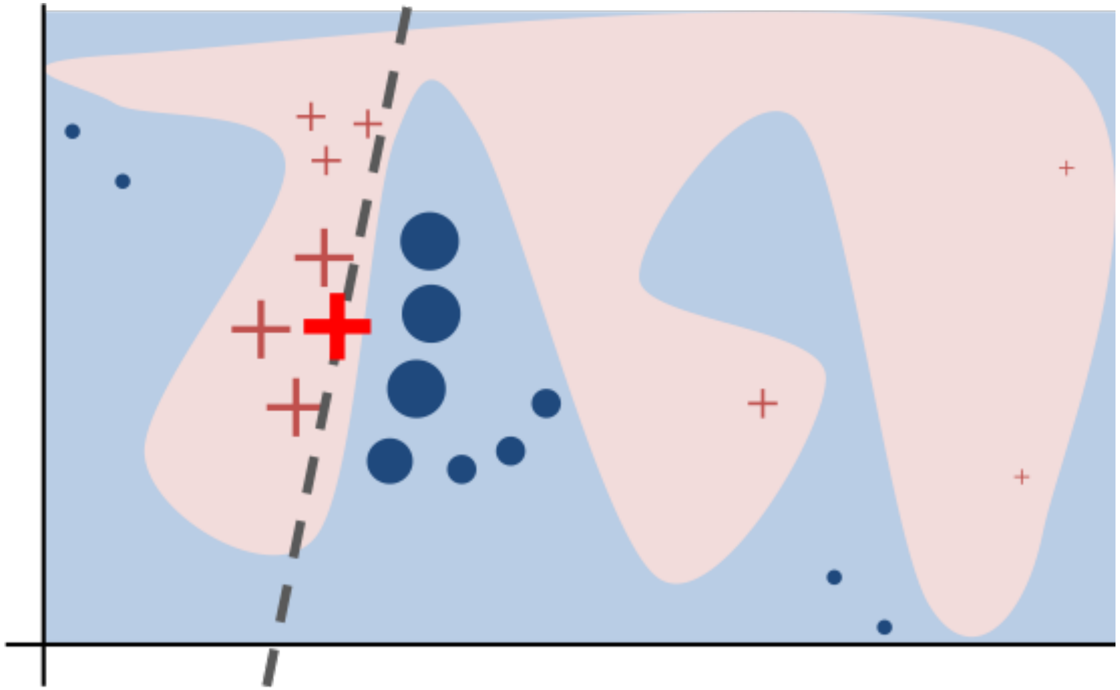


...

Counterfactuals

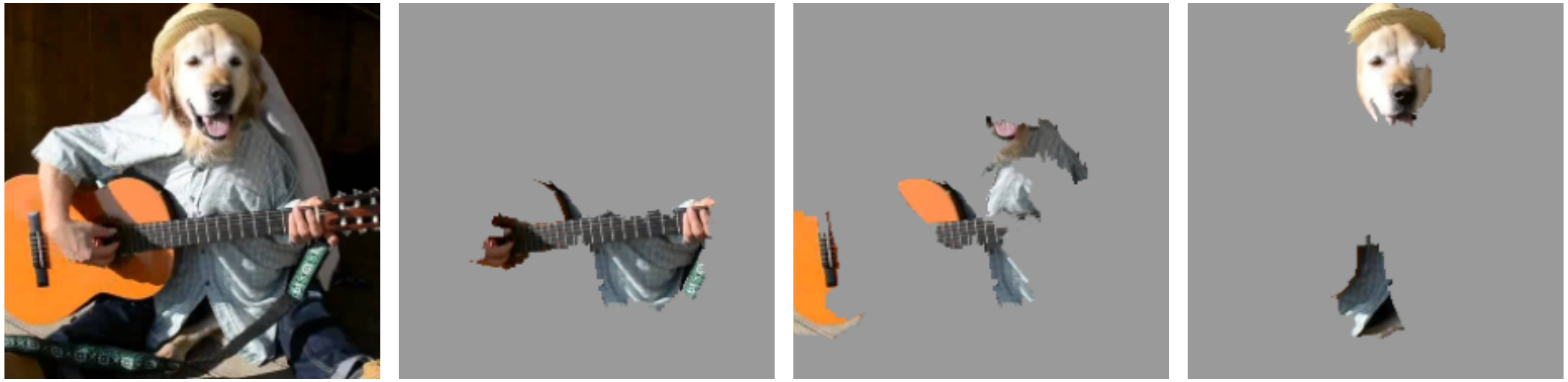
- **If** James Dean had **taken the train** the day of his car accident, he would not have died

Local simplification



- LIME

Sensitivity analysis



- The pixels that best "explain"
  - The recognition of a **electric guitar**
  - The recognition of an **acoustic guitar**
  - The recognition of a **dog**

- Still very rudimentary

What kind of knowledge can we **extract**?

- When interpretability is NOT needed?

What kind of knowledge can we **extract**?

- When interpretability is NOT needed?
  - When **low risk** associated with the decision
    - E.g. *recommendation for a movie*
  - When **good guarantees** on performance exist
    - E.g. *character recognition*

What kind of knowledge can we **extract**?

- When interpretability IS needed?

What kind of knowledge can we **extract**?

- When interpretability IS needed?

  1. With **high risk decisions**
     - *E.g. chirurgical operation*
     - *E.g. shutting down a nuclear plant*
     - *E.g. autonomous vehicle*

What kind of knowledge can we **extract**?

- When interpretability IS needed?

  1. With **high risk decisions**
     - *E.g. chirurgical operation*
     - *E.g. shutting down a nuclear plant*
     - *E.g. autonomous vehicle*

  2. Satisfying **curiosity** (what science is about)
     - *E.g. explain surprising results*
     - *E.g. when no easy explanation exists*
     - *E.g. when the decision function must be included in a larger inference system (a domain theory)*

What kind of knowledge can we **extract**?

- When interpretability IS needed?

  3. **Debugging**

     - *E.g. why is that decision wrong (counterfactual)*
     - *E.g. if a bicycle is recognized because it has two wheels, what if one is hidden behind side bags?*
     - *E.g. why the system seems gender biased?*

What kind of knowledge can we **extract**?

- When interpretability IS needed?

  3. **Debugging**
     - *E.g. why is that decision wrong (counterfactual)*
     - *E.g. if a bicycle is recognized because it has two wheels, what if one is hidden behind side bags?*
     - *E.g. why the system seems gender biased?*

  4. Interpretability **demands higher standard predictive systems**
     - *An interpretable system **can be manipulated***
       - *E.g. if someone knows that a loan is granted if you have more than 2 credit cards*
     - *In order **not to be manipulated**, the predictive system must use **causal factors***

- Why is Machine Learning currently **lacking**?
  - The **exclusive focus on predictive performance** leads to an incomplete learning problem formulation

- Why is Machine Learning currently **lacking**?
  - The **exclusive focus on predictive performance** leads to an incomplete learning problem formulation
  - We want also
    - **Interpretability** of the **results**
    - **Interpretability** of the **process**
    - Gaining a **better understanding of the world** when including the learned decision function in an existing theory

- Why is Machine Learning currently **lacking**?
  - The **exclusive focus on predictive performance** leads to an incomplete learning problem formulation
  - We want also
    - **Interpretability** of the **results**
    - **Interpretability** of the **process**
    - Gaining a **better understanding of the world** when including the learned decision function in an existing theory

Somehow, we have to change the inductive criterion used in Machine Learning

Outline

1. A brief history of AI
2. AI now: the triumph of deep neural networks
3. AI in the near future
4. There are limits
5. The case of XAI
6. Conclusion

- There are reasons to be stunned
  - Enormous **progress** the last few years
  - In combination with **IoT**, a new era is coming
- But also to be cautious
  - These systems **do not understand**
  - They **do not explain**
  - They are essentially **black boxes**
  - And not well understood yet

A lot remains to be done

Some of us still dream the old dream

Towards a General Artificial Intelligence?

- AI far surpasses humans at **narrow tasks** that can be **optimized based on data**
- BUT, it cannot engage in cross-domain thinking on creative tasks or ones requiring complex strategies

- For **future** research:
  - **Multidomain** learning
  - Real **understanding**
  - **Common sense** reasoning
  - Learning from very **few examples**
  - Understanding **humor**
  - Self-awareness?