Introduction to AI: Understanding the Technology

Winter Academy on Artificial Intelligence and International Law

Asser Institute – 11 February 2021

Giovanni Sileno
g.sileno@uva.nl

What is Artificial Intelligence?

What is Artificial ~~Intelligence~~?

What is Artificial ~~Intelligence~~?

- What is made by humans?

What is Artificial ~~Intelligence~~?

- What is made by humans?
- What is induced by humans?

What is Artificial ~~Intelligence~~?

- What is made by humans?
- What is induced by humans?
- What is simulated, not true?

"It's a sunset, Billy. It's not selling anything."

What is ~~Artificial~~ Intelligence?

What is ~~Artificial~~ Intelligence?

- Problem-solving ability?

What is ~~Artificial~~ Intelligence?

- Problem-solving ability?
- Capacity of abstraction?

What is ~~Artificial~~ Intelligence?

- Problem-solving ability?
- Capacity of abstraction?
- Capacity of organization?

What is ~~Artificial~~ Intelligence?

- Problem-solving ability?
- Capacity of abstraction?
- Capacity of organization?
- Creativity?

What is ~~Artificial~~ Intelligence?

- Problem-solving ability?
- Capacity of abstraction?
- Capacity of organization?
- Creativity?
- Self-awareness?

What is ~~Artificial~~ Intelligence?

- Problem-solving ability?
- Capacity of abstraction?
- Capacity of organization?
- Creativity?
- Self-awareness?
- Manipulation ability?

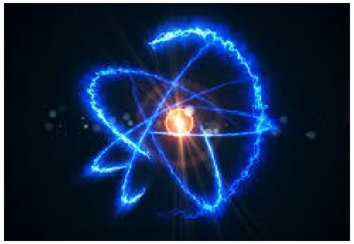
AI as a discipline

- Most disciplines emerge around specific domains of knowledge, settling upon methods deemed adequate to that domain.

Biology

Life and living organisms

Physics



Laws of the universe

Law

Legal systems and justice

Computer science

Computational systems

but Artificial Intelligence?

AI as a discipline

- As a discipline, AI is not primarily connected to a knowledge domain, but to a **purpose**:

conceiving artificial systems that are intelligent

AI as a discipline

- As a discipline, AI is not primarily connected to a knowledge domain, but to a **purpose**:

conceiving artificial systems that are intelligent

- All other disciplines (and their methods, or refinements of their methods) become for AI ***instrumental*** to that purpose (or sub-goals derived from that purpose).
- **But what is meant by this purpose?**

AI as a discipline

- As a discipline, AI is not primarily connected to a knowledge domain, but to a **purpose**:

conceiving artificial systems that are intelligent

- All other disciplines (and their methods, or refinements of their methods) become for AI ***instrumental*** to that purpose (or sub-goals derived from that purpose).
- **But what is meant by this purpose?**

it depends on what we mean by "intelligence"...

Categories of AIs

systems that

think like humans	think rationally
act like humans	act rationally

Russell and Norvig, "Artificial Intelligence: a Modern Approach", chapter 1
available at https://people.eecs.berkeley.edu/~russell/aima1e/chapter01.pdf

Categories of AIs

systems that

MENTAL dimension	think like humans	think rationally
BEHAVIOURAL dimension	act like humans	act rationally

Categories of AIs

systems that

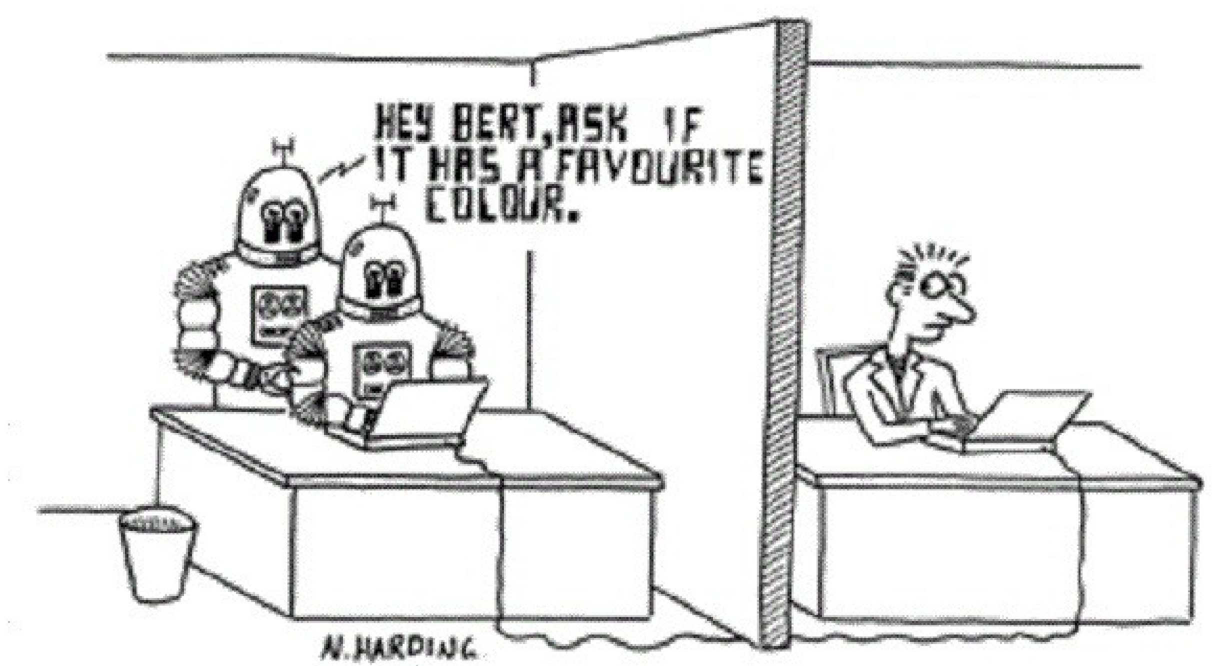
MENTAL dimension	think like humans	think rationally
BEHAVIOURAL dimension	act like humans	act rationally
	DESCRIPTIVE dimension	**PRESCRIPTIVE dimension**
	standards set by **actual** human behaviour	standards set by **ideal** (human) behaviour

systems that

think like humans	think rationally
act like humans	act rationally

Turing test approach

artificial and natural not distinguishable behind a neutral interface



systems that

think like humans	think rationally
act like humans	act rationally

Cognitive modeling approach

AI reproducing cognitive functions observed by humans

NATURA ARTIS MAGISTRA argument

If these cognitive functions are required for our intelligence → they might be required to achieve artificial intelligence

EXPLAINABILITY argument

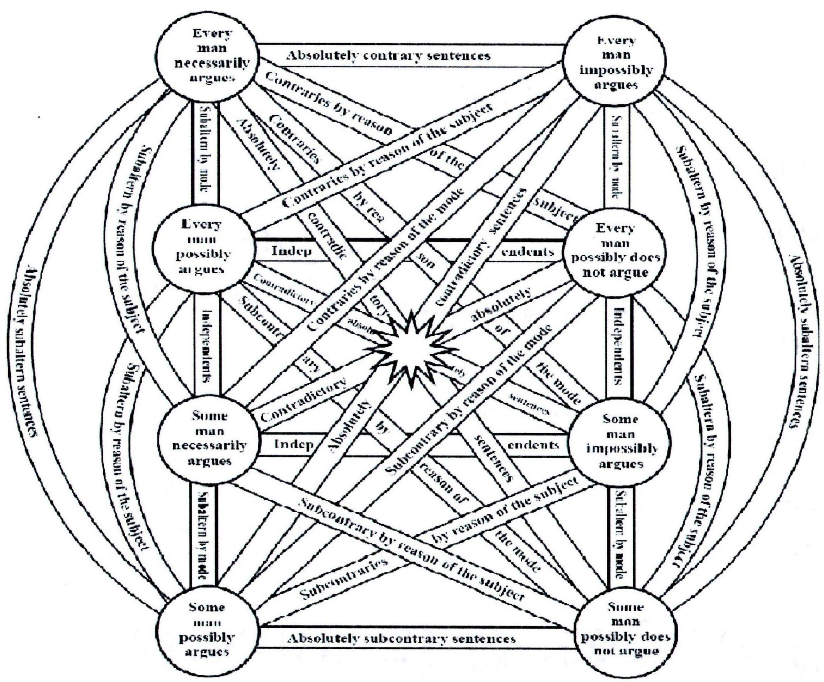
If they explain our internal working → they can help to interpret AI functioning

systems that

think like humans	think rationally
act like humans	act rationally

The "Laws of Thought" approach

AI producing logically valid inferences

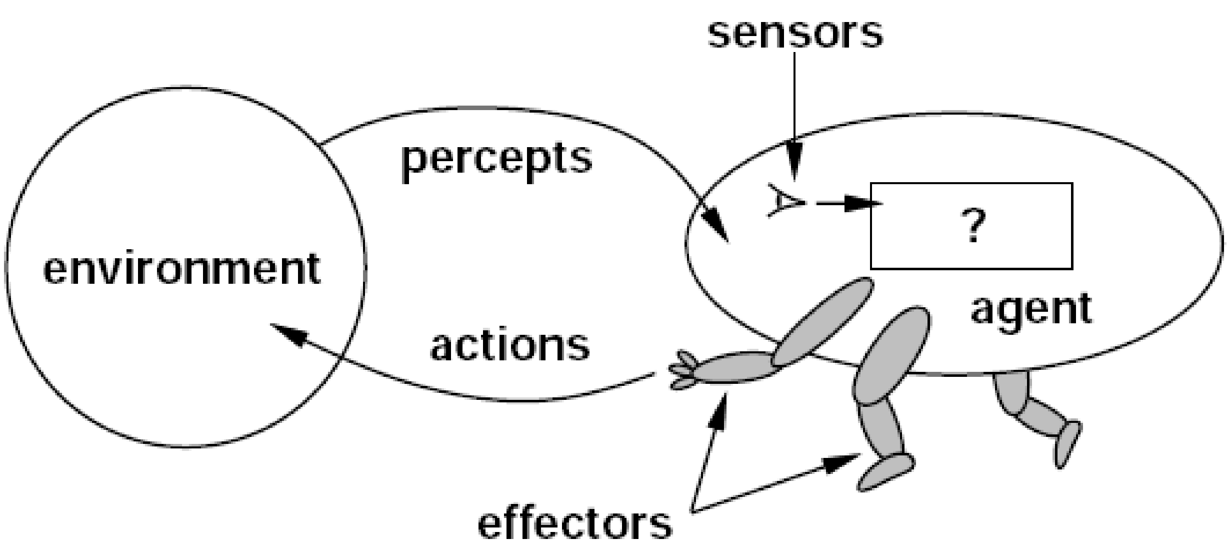


systems that

think like humans	think rationally
act like humans	act rationally

The "Rational Agent" approach

AI decision-making following standards of rationality



- the agent selects the best choice
- to achieve its **goals**
- given its **beliefs**

autonomous entity

Superhuman performances?

- In specific tasks, performance can be easily measured (quantified).
  → *systems can adapt to perform better than humans.*

systems that

think like humans	think rationally
perform like humans	act rationally

outperform humans

in

narrow (specific)	general

contexts

Of the many AI waves

- This variety of topics has been developed through a cycles of ***springs*** (and ***winters***) centered around different topics.

Of the many AI waves

- This variety of topics has been developed through a cycles of ***springs*** (and ***winters***) centered around different topics.
- Some of the peaks:
  - *ad-hoc* systems with handcrafted knowledge (**60s**/**70s**)
  - expert systems/problem solving methods (**80s**)
  - robotics, computer vision, speech recognition (**80s**)
  - evolutionary computing (**90s**)
  - agent-based modeling and multi-agent systems (**90s/00s**)
  - semantic web (**00s**)
  - **deep learning (10s)** ← **NOW**

Key message

- Despite the contemporary (ab)use of the term, AI covers much more than the subfields that have brought the most recent advances:

AI ≠ ML or DL

machine learning deep learning

Key message

- Despite the contemporary (ab)use of the term, AI covers much more than the subfields that have brought the most recent advances:

AI ≠ ML or DL

machine learning deep learning

- This misappropriation is not new (even the term ML was coined by researchers to distinguish themselves from logic-based AI).

logicist

reasoning and decision-making

AI AS ENGINEERING OF THE “MIND”

induction of functions from data

empiricist

monolithical systems

logic

logicist

reasoning and decision-making

AI AS ENGINEERING OF THE “MIND”

heterogeneous systems

induction of functions from data

homogeneous systems

artificial neural networks (ANNs)

monolithical systems

empiricist

probability

logicist

The main problem here is collecting the <u>relevant</u> knowledge

→ **EXPLICITATION**

reasoning and decision-making

AI AS ENGINEERING OF THE "MIND"

induction of functions from data

The problem here is inducing the <u>tacit</u> behavioural model, not applying it!

→ **INTERNALIZATION**

empiricist

Working principles of Symbolic AI

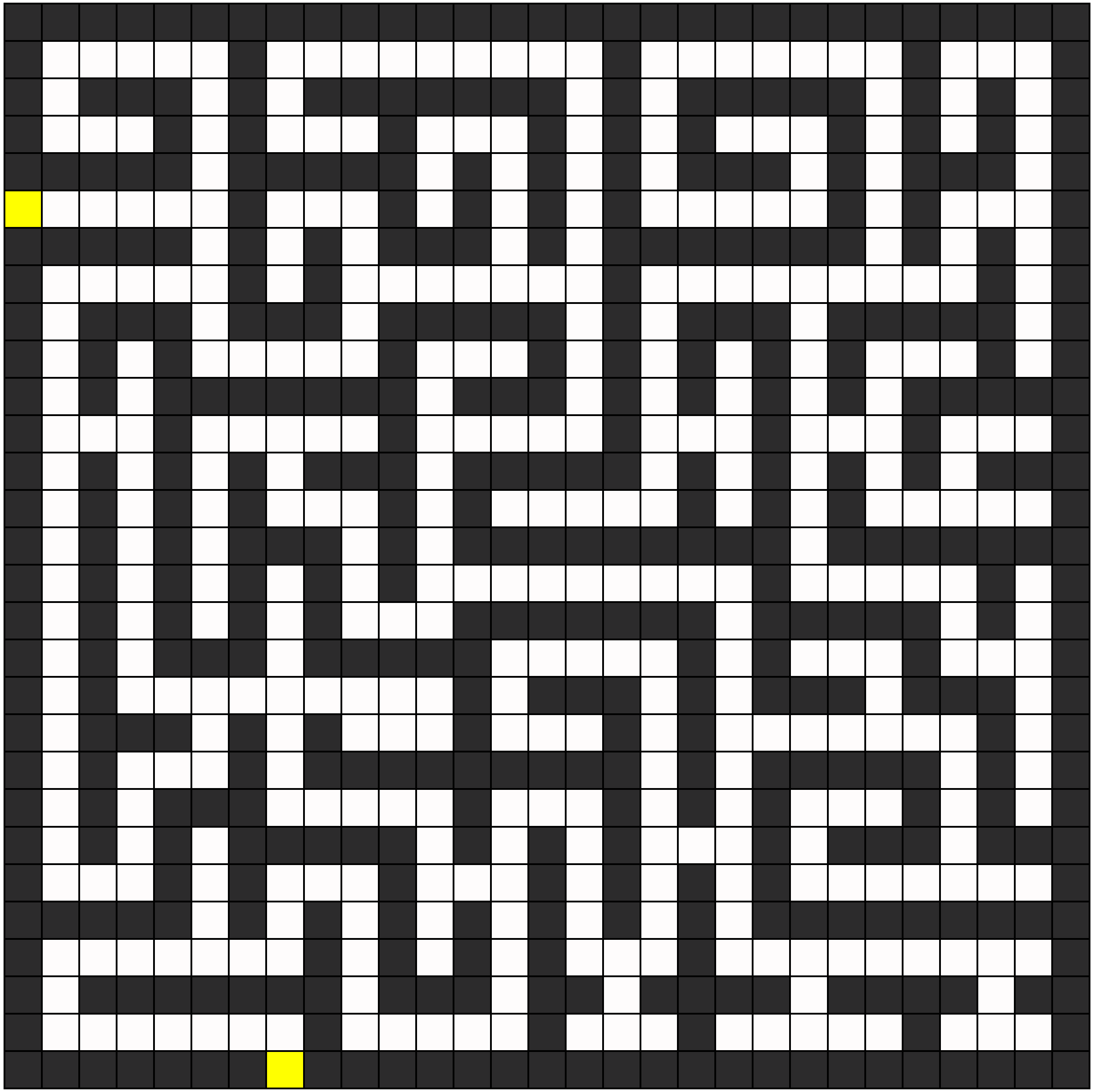
Algorithm = Logic + Control

"An algorithm can be regarded as consisting of

- a **logic component**, which specifies the *knowledge* to be used in solving problems, and
- a **control component**, which determines the *problem-solving strategies* by means of which that knowledge is used.

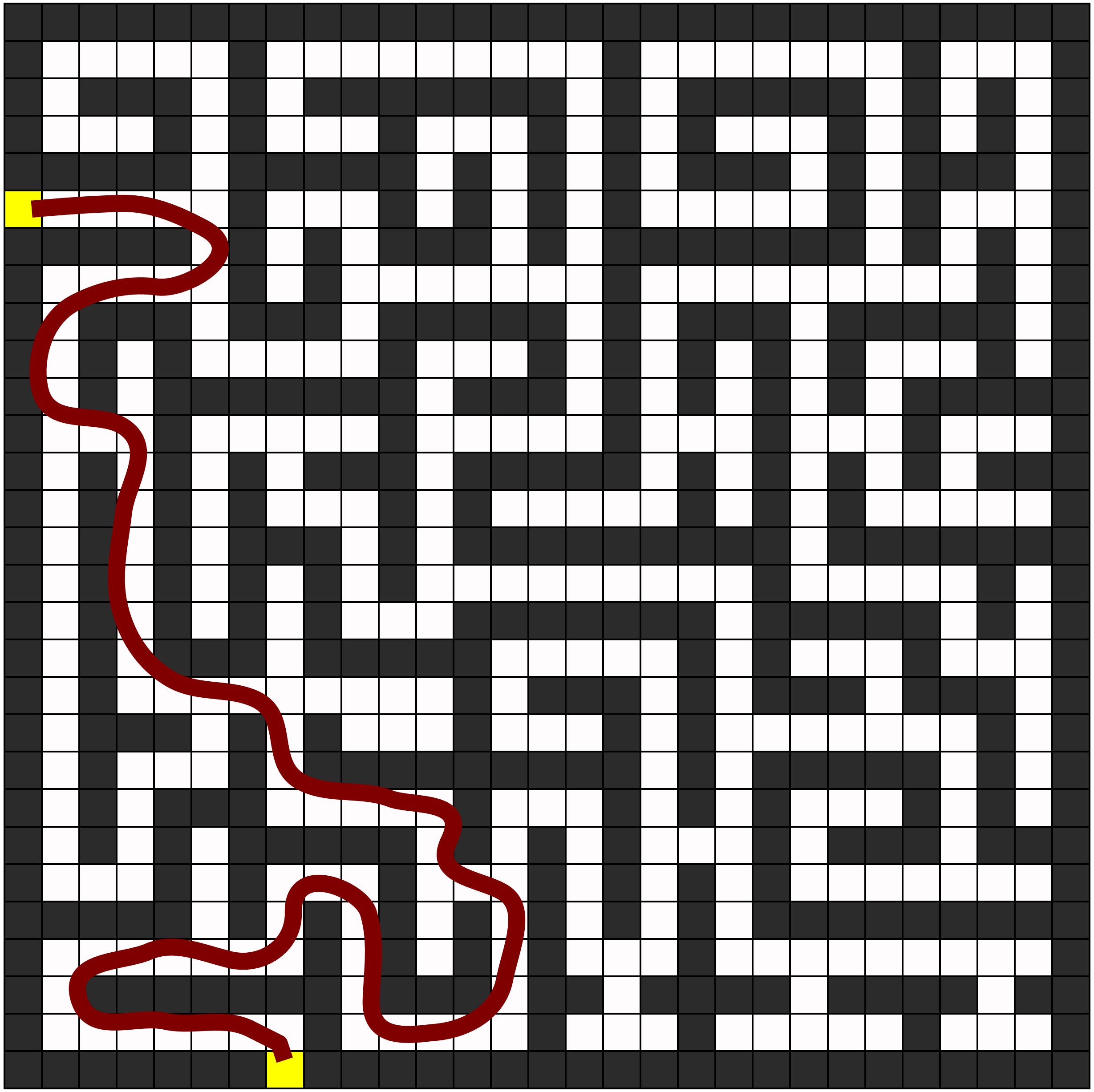
The logic component determines the meaning of the algorithm whereas the control component only affects its efficency."

Kowalski, R. (1979). Algorithm = Logic + Control. Communications of the ACM, 22(7), 424–436.



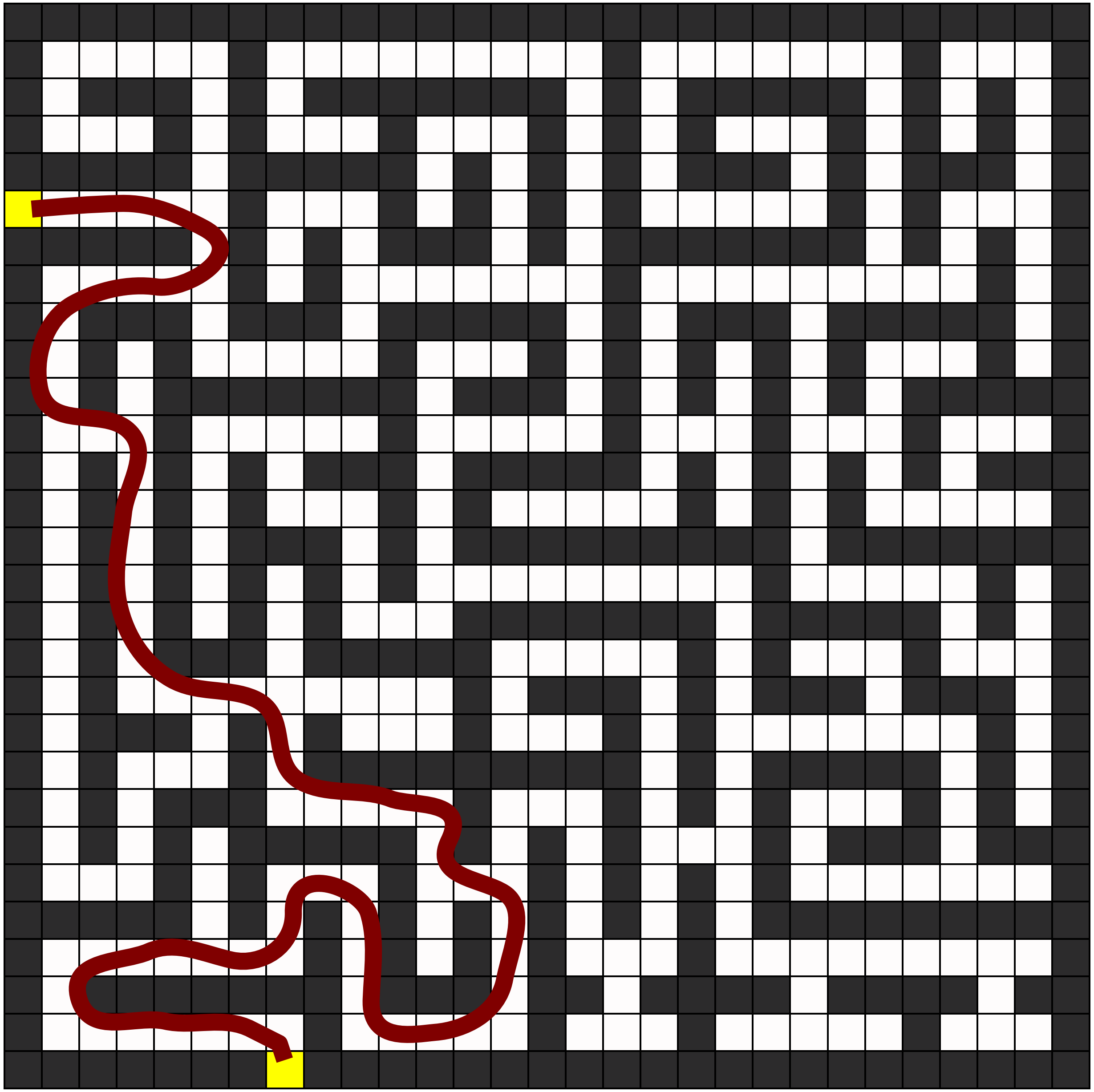
Imperative approach

you command the directions



Imperative approach

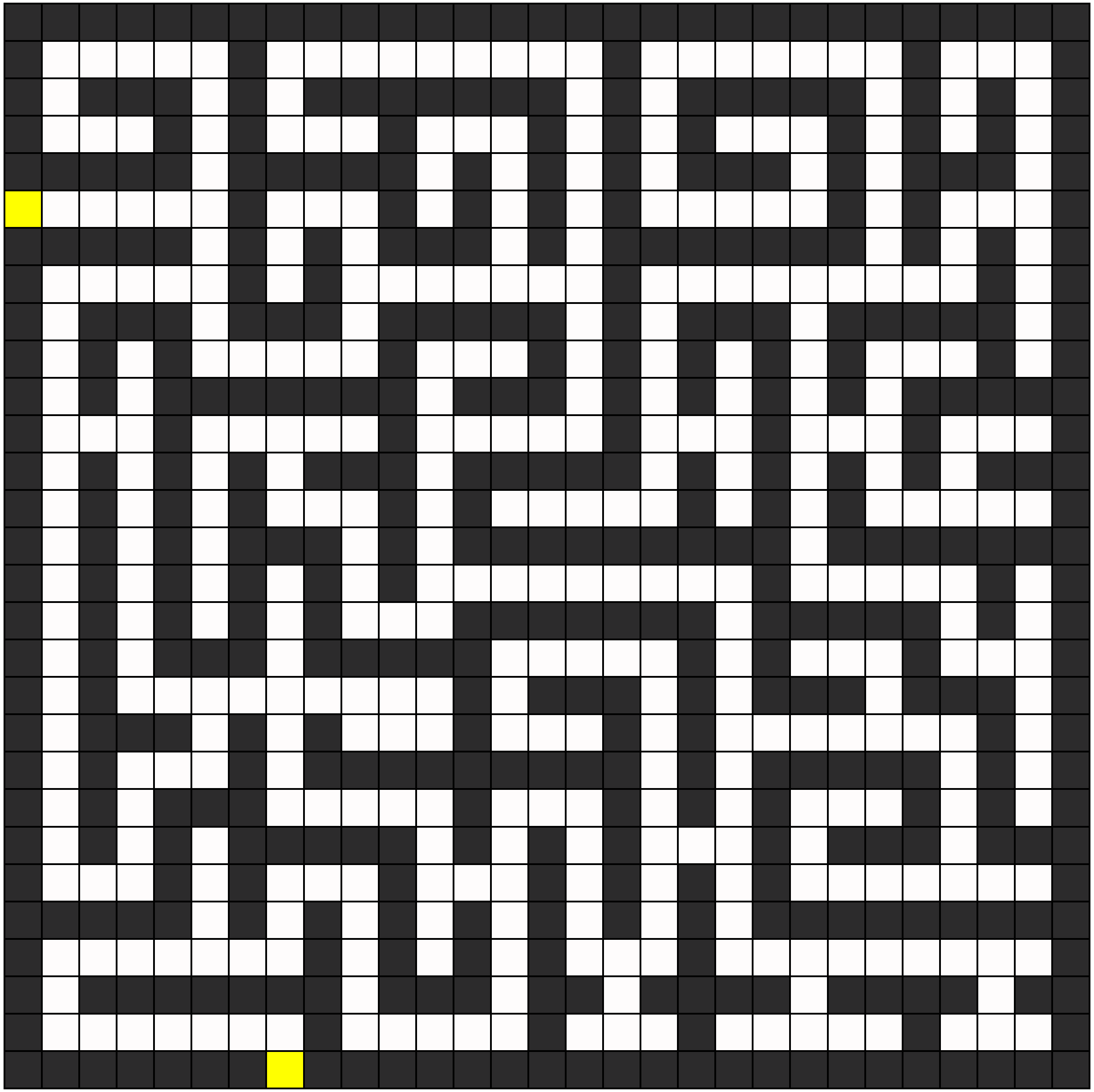
you command the directions



Imperative approach

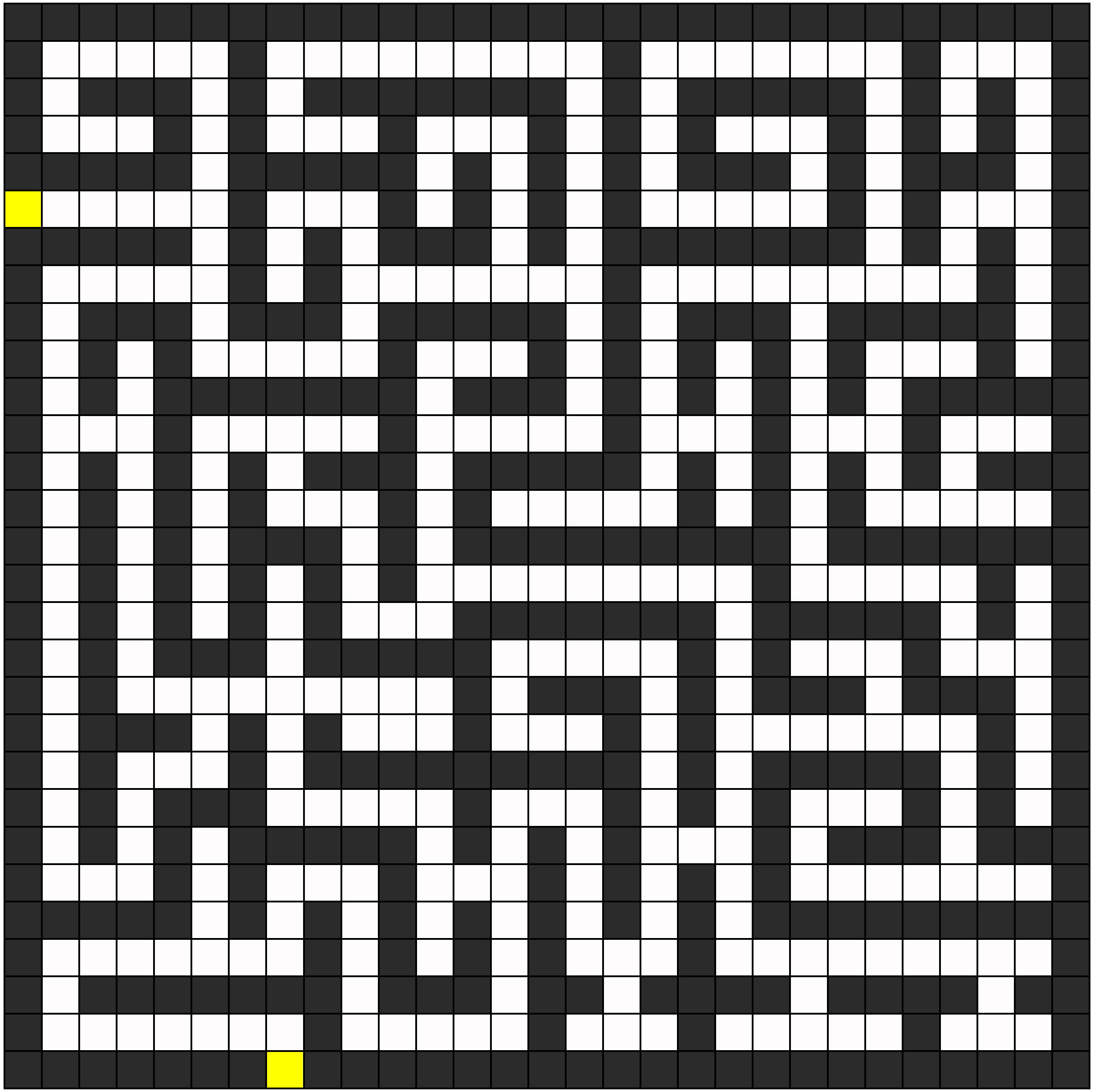
you command the directions

- What if the labyrinth changes?



Declarative approach
you give just the labyrinth.

the computer finds the way.

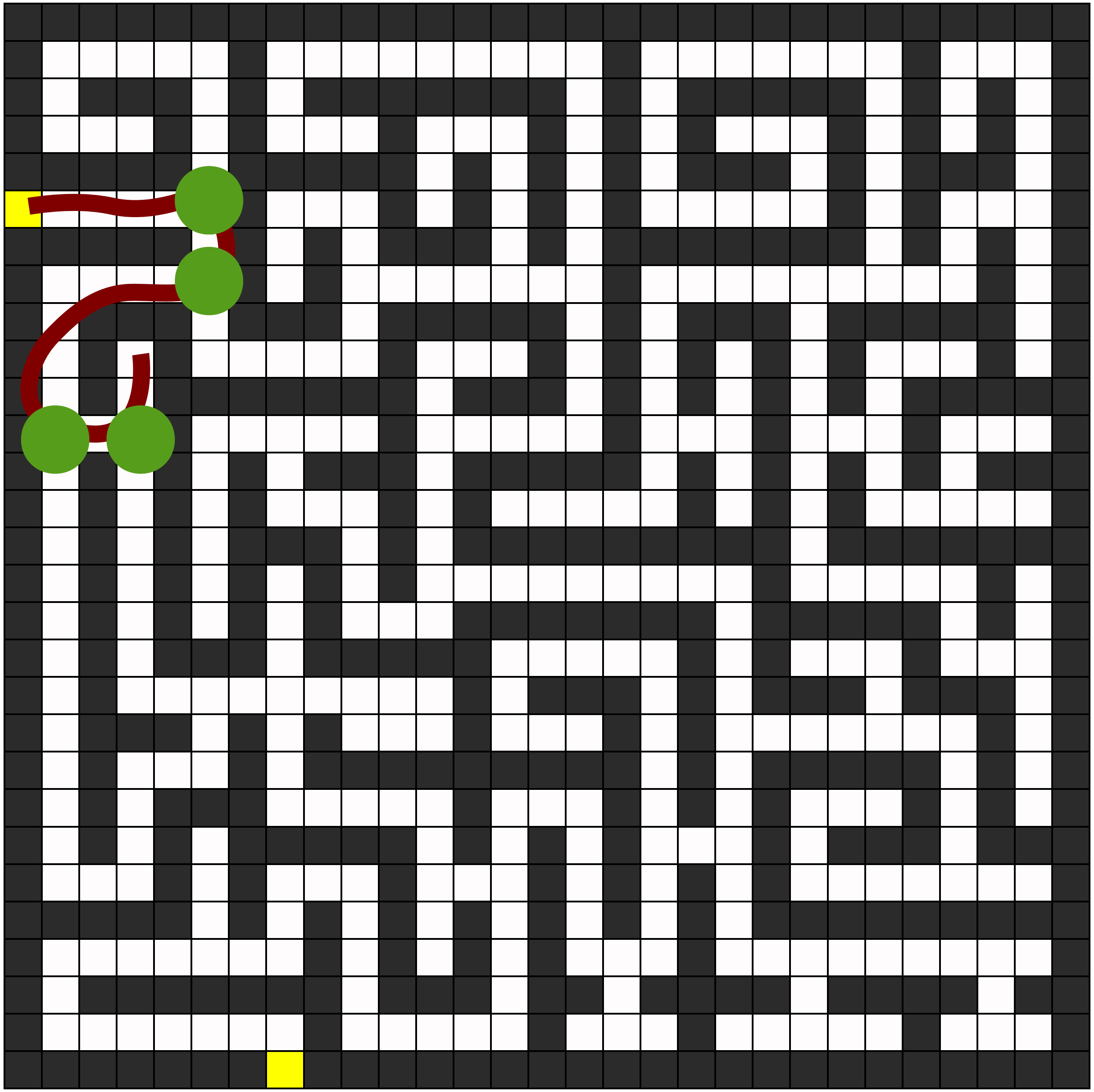


Declarative approach

you give just the labyrinth.

the computer finds the way.

- For instance, via ***trial***, ***error*** and ***backtracking***.

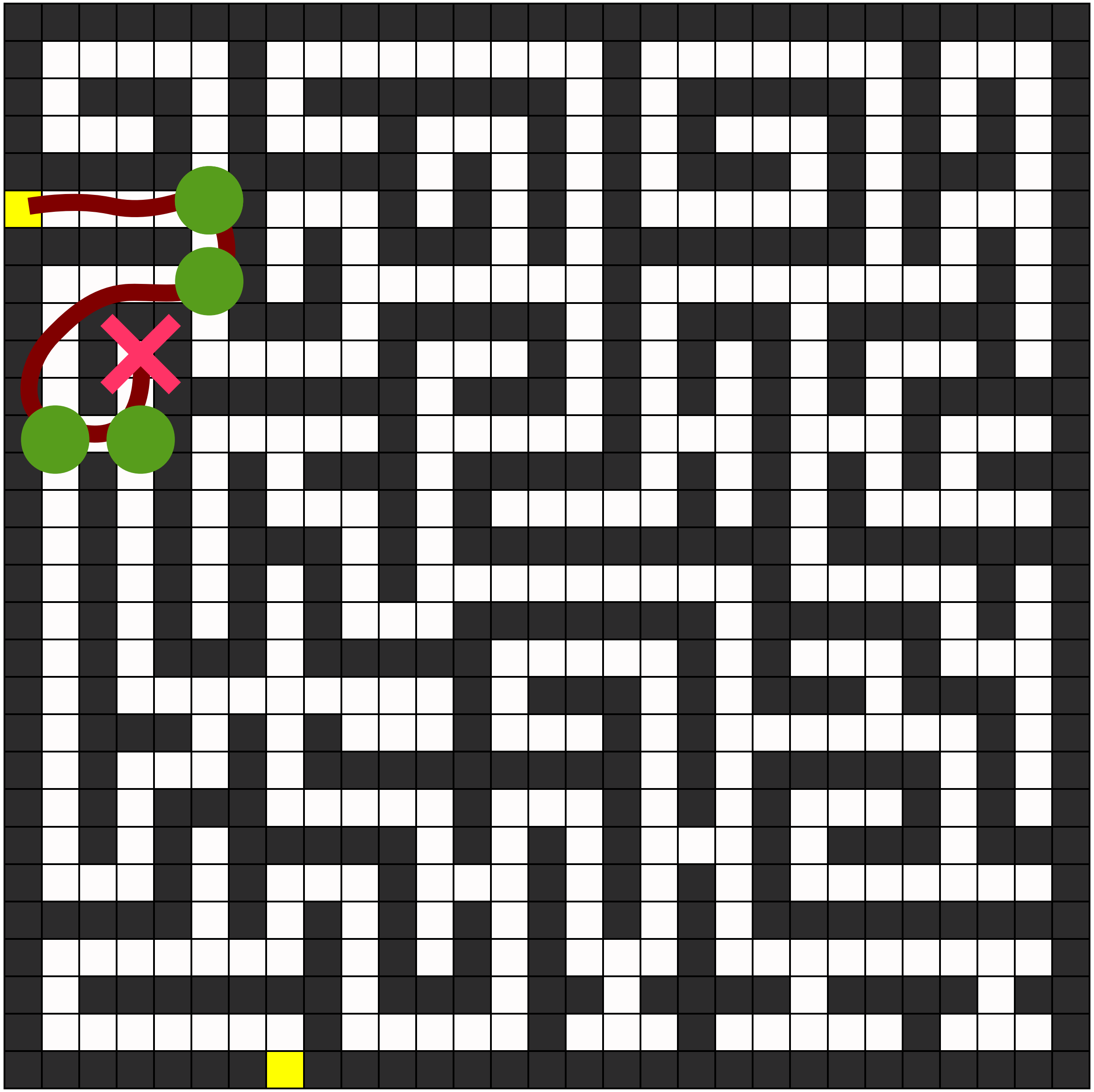


Declarative approach

you give just the labyrinth.

the computer finds the way.

- For instance, via ***trial***, ***error*** and ***backtracking***.

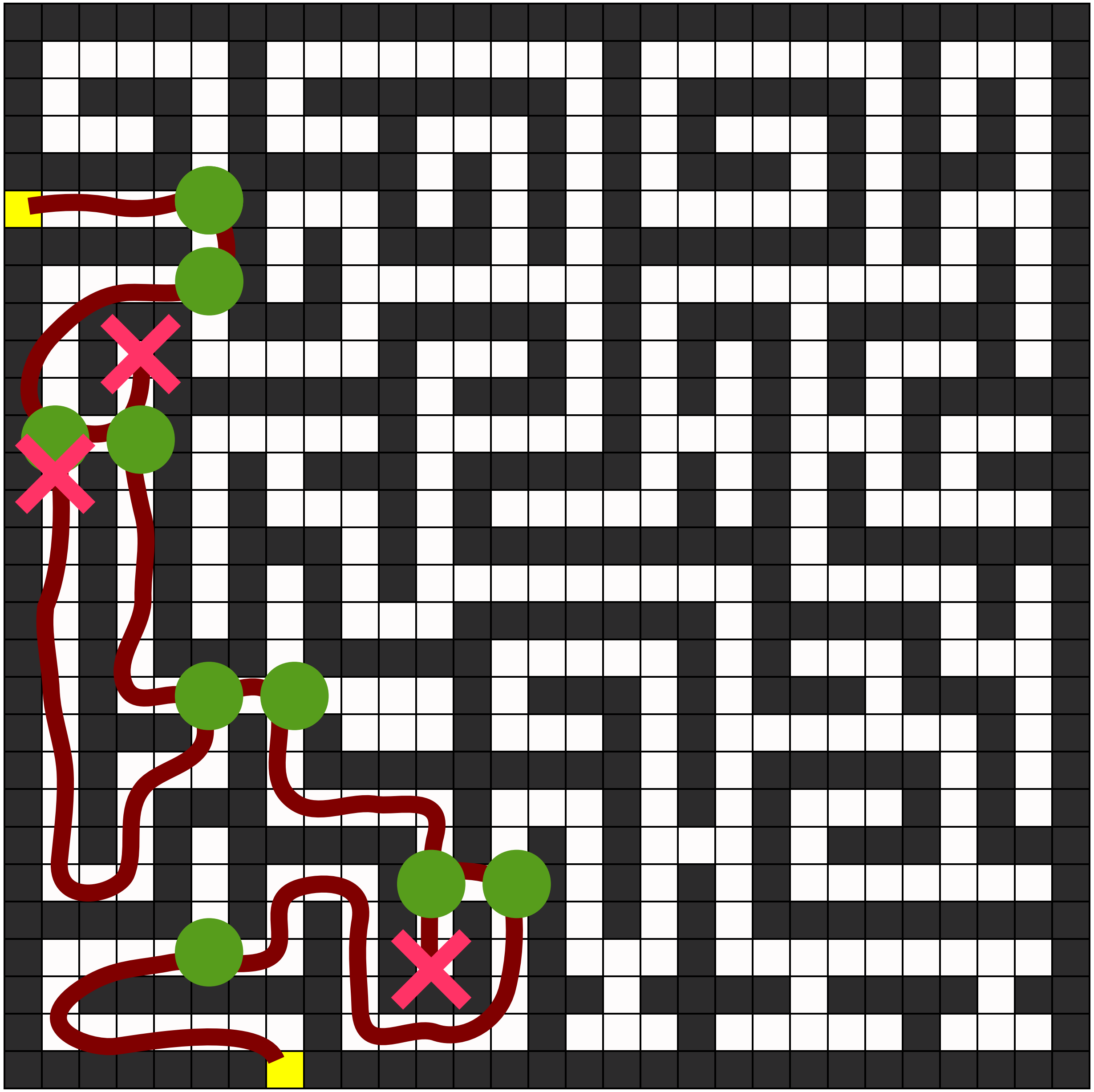


Declarative approach

you give just the labyrinth.

the computer finds the way.

- For instance, via ***trial***, ***error*** and ***backtracking***.



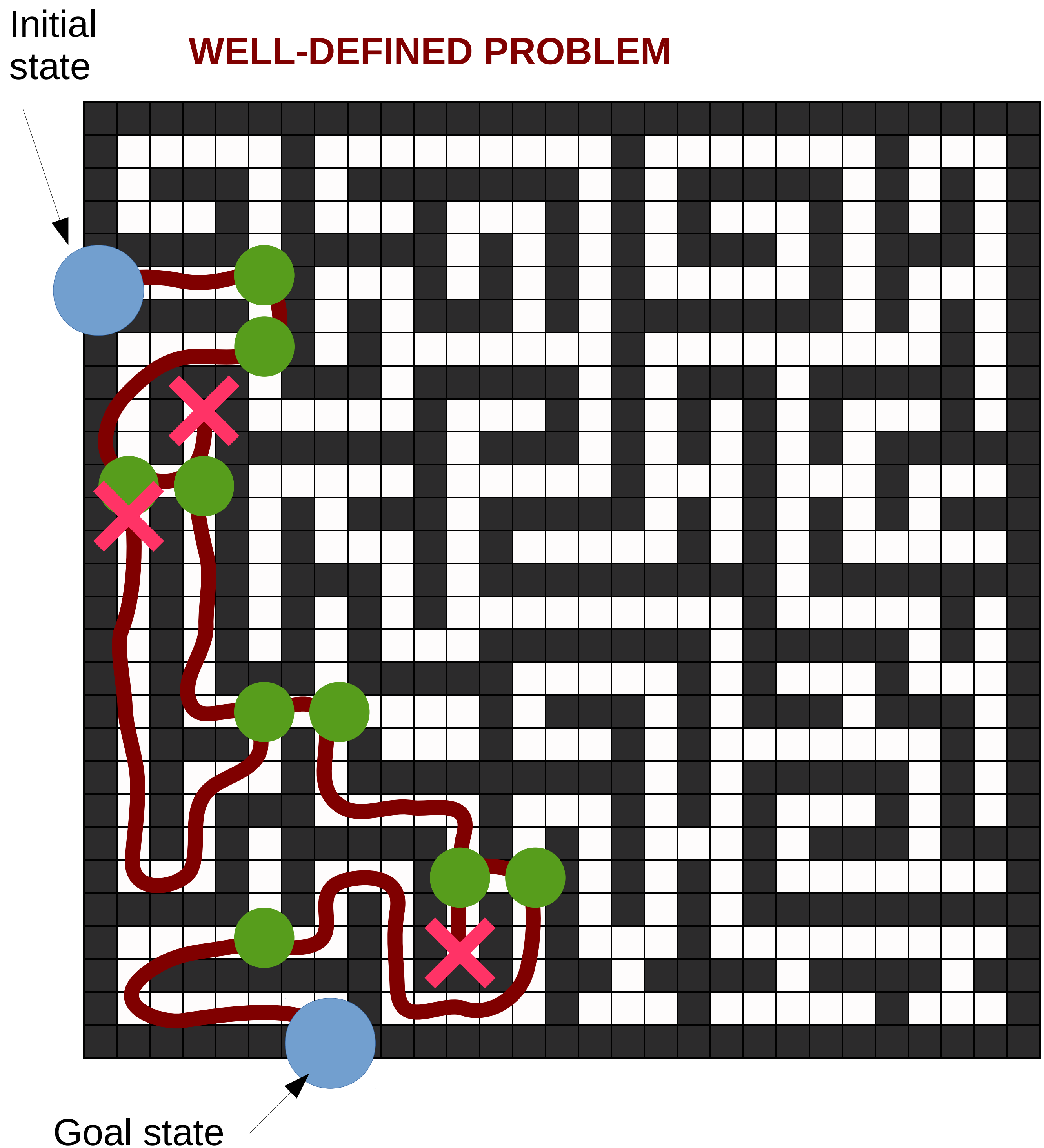
Declarative approach

you give just the labyrinth.

the computer finds the way.

- For instance, via ***trial***, ***error*** and ***backtracking***.

WELL-DEFINED PROBLEM



KNOWLEDGE

Declarative approach

you give just the labyrinth.

the computer finds the way.

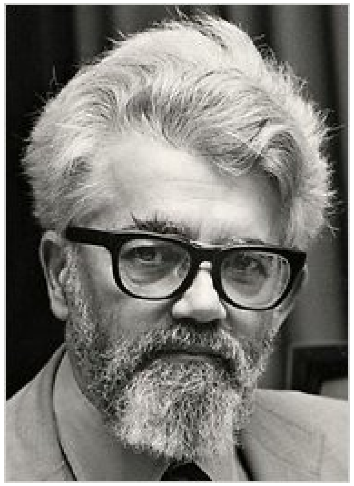
- For instance, via ***trial***, ***error*** and ***backtracking***.

PROBLEM-SOLVING METHOD

Well-defined problems & problem spaces

Problems are *well-defined* when there is a simple test to conclude whether a solution is a solution.

J. McCarthy (1956) The inversion of functions defined by Turing machines. Automata Studies, Annals of Mathematical Studies, 34:177 – 181.



People solve problems by *searching* through a problem space, consisting of the *initial state*, the *goal state*, and *all possible states in between*.

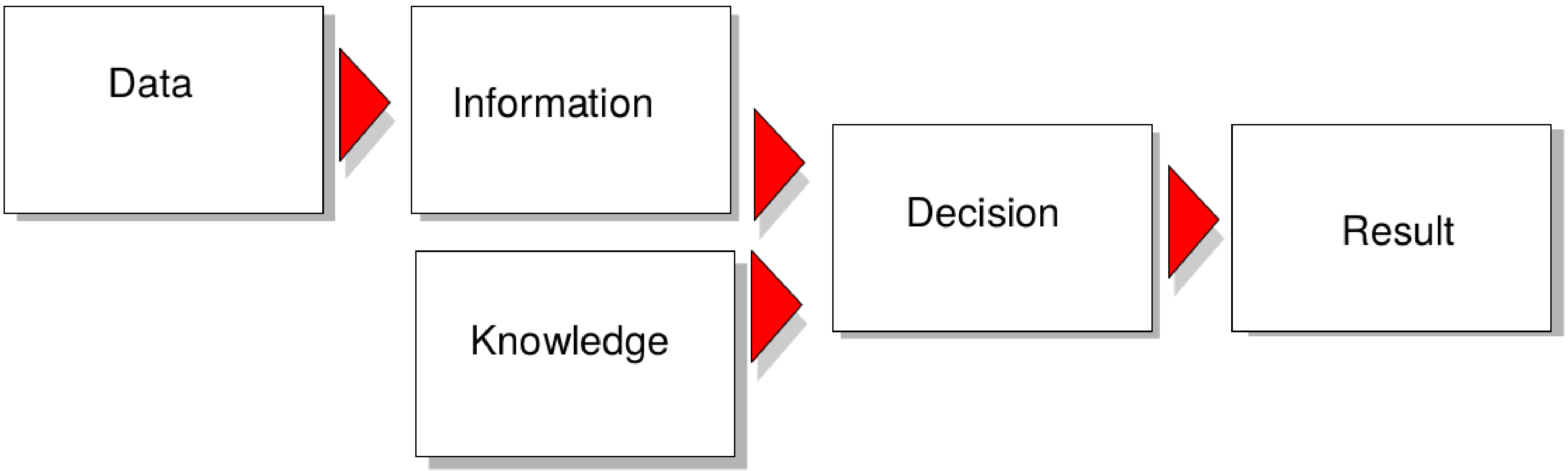
Newell, A., & Simon, H. A. (1972). Human problem solving.

Intelligence as search: Garry Kasparov vs IBM's Deep Blue (1997)

200 million positions per second
alpha-beta min-max search
static evaluation heuristics based on matches played by masters

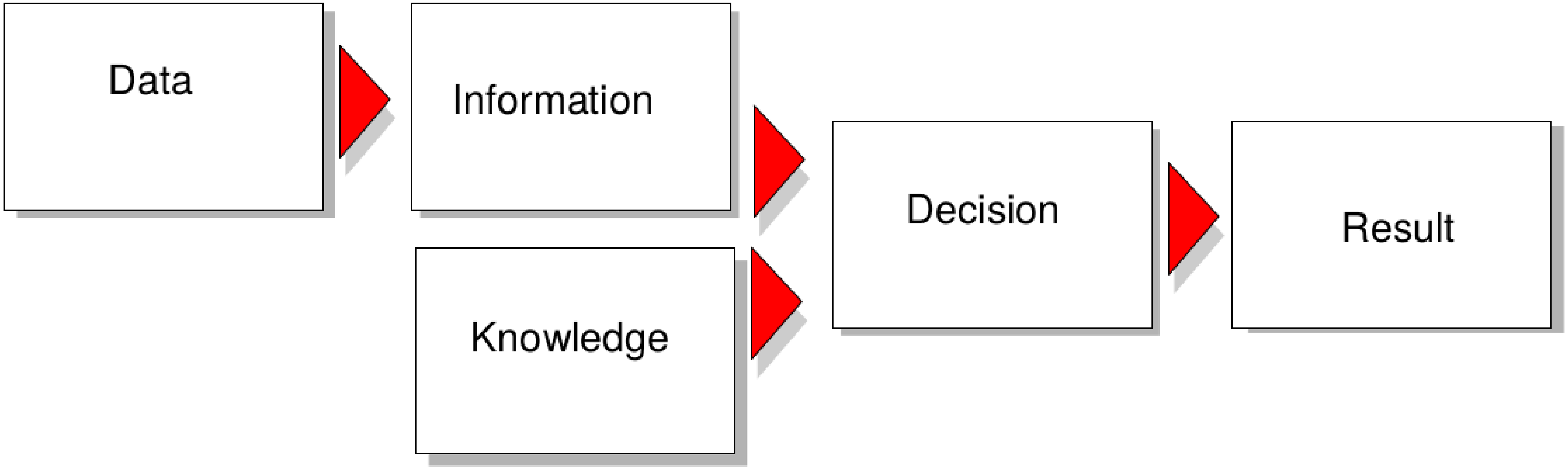
Data, Information, Knowledge

- **Data**: uninterpreted signals or symbols, e.g. 14



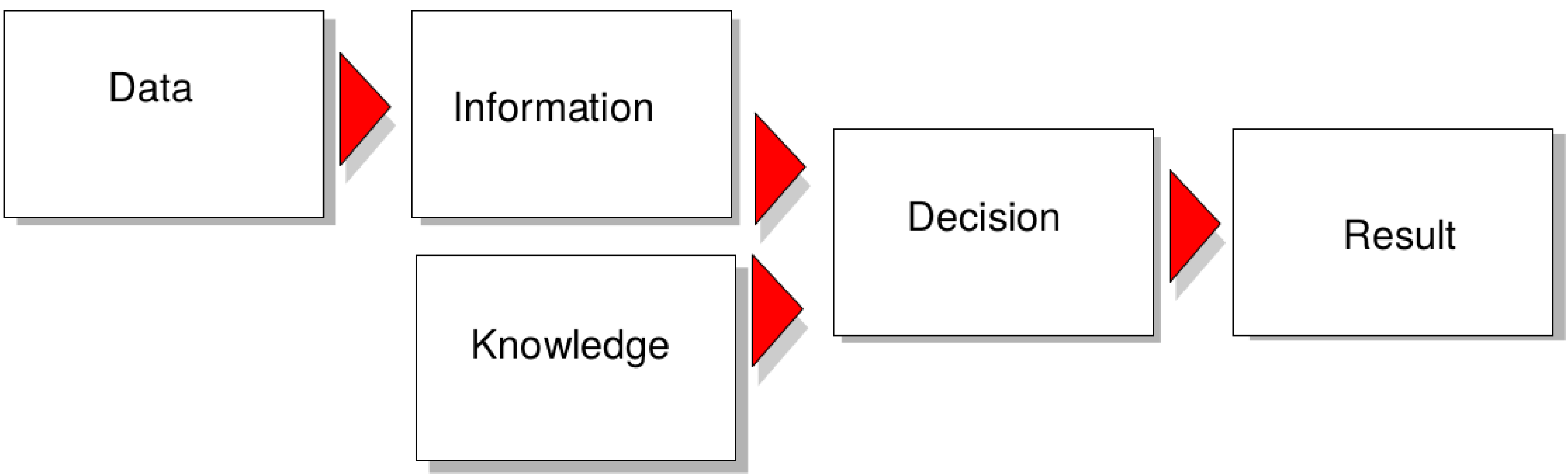
Data, Information, Knowledge

- **Data**: uninterpreted signals or symbols, e.g. 14
- **Information**: data with added meaning, e.g. 14°



Data, Information, Knowledge

- **Data**: uninterpreted signals or symbols, e.g. 14
- **Information**: data with added meaning, e.g. 14°
- **Knowledge:** all data and information that people use to *act*, *accomplish tasks* and to *create new information* (e.g. know-how, -why, -who, -where and -when), e.g. 14° is the temperature now, in this room, measured by …



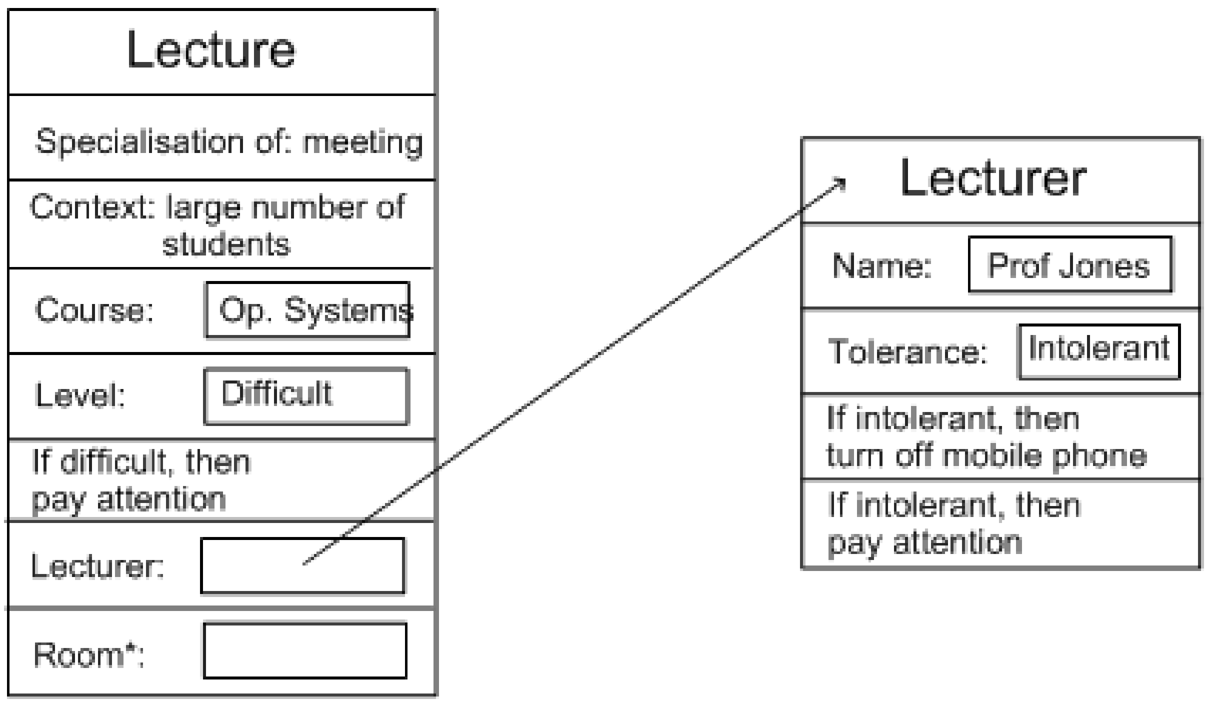
Expert system (rule base)

if flower **and** seed **then** phanerogam
if phanerogam **and** bare-seed **then** fir
if phanerogam **and** 1-cotyledon **then** monocotyledonous
if phanerogam **and** 2-cotyledon **then** dicotyledonous
if monocotyledon **and** rhizome **then** thrush
if dicotyledon **then** anemone
if monocotyledon **and** ¬rhizome **then** lilac
if leaf **and** flower **then** cryptogamous
if cryptogamous **and** ¬root **then** foam
if cryptogamous **and** root **then** fern
if ¬leaf **and** plant **then** thallophyte
if thallophyte **and** chlorophyll **then** algae
if thallophyte **and** ¬ chlorophyll **then** fungus
if ¬leaf **and** ¬flower **and** ¬plant **then** colibacille

rhizome + flower + seed + 1-cotyledon ?

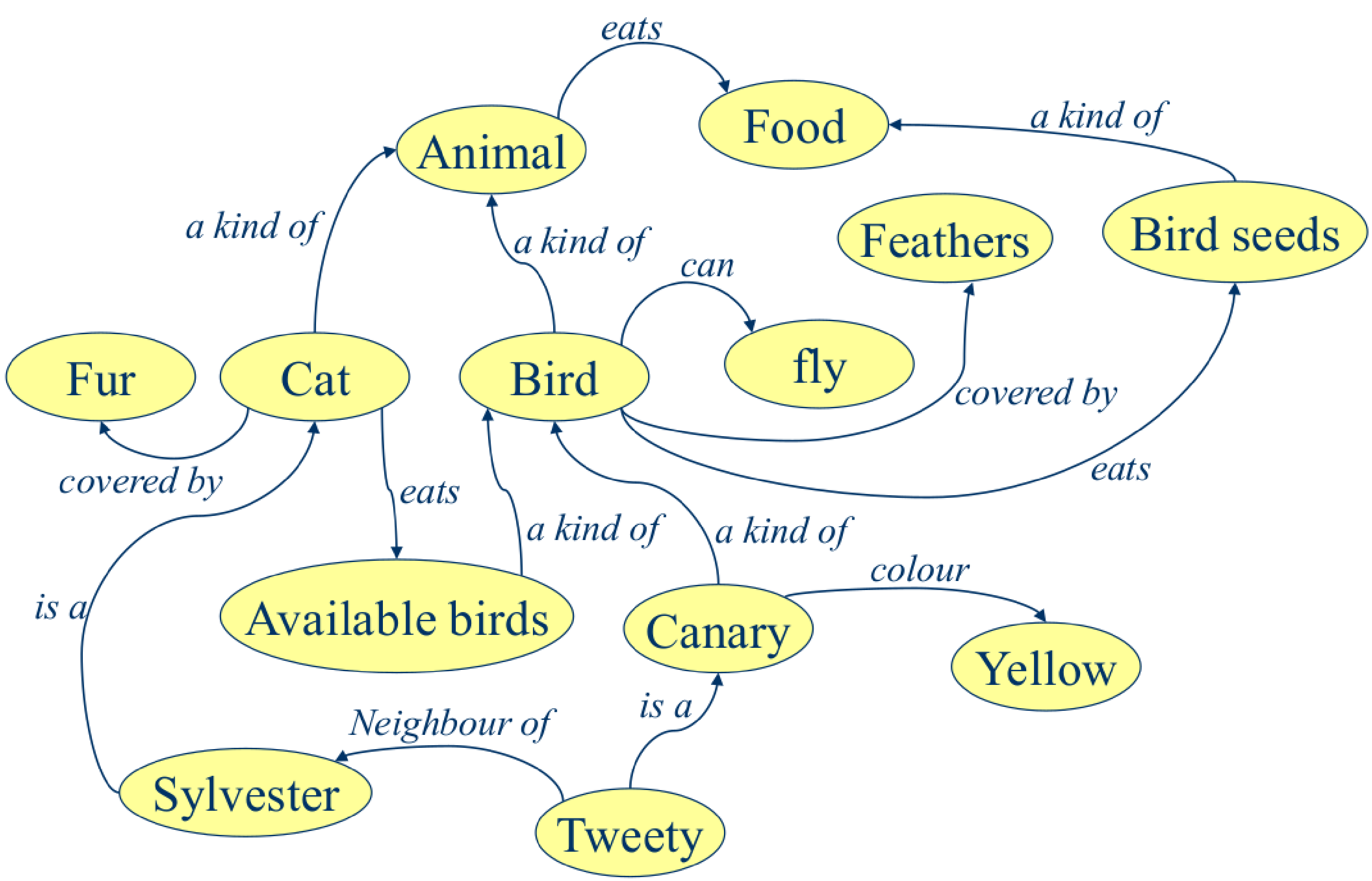
Frames

- Frames are "stereotyped" knowledge units representing situations, objects or events or (classes) sets of such entities.



(base for the **Obiect-Oriented Programming** paradigm**)**

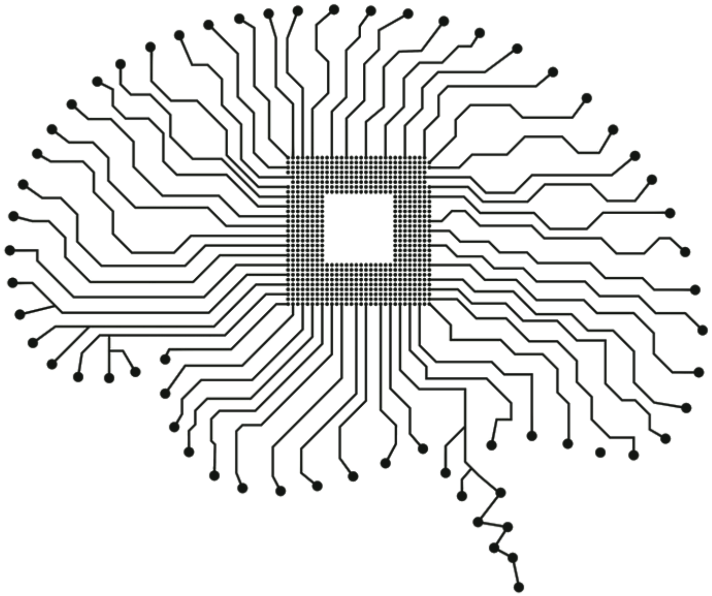
Semantic Networks



(used in contemporary **Semantic Web** technologies**)**

In sum...

- Symbolic AI presents ***transparent*** techniques to effectively model and solve problems that can be described in symbolic terms *(where expertise can be verbalized)*.
- All IT systems of organizations today rely on some of the technologies introduced or emerged during the first AI wave.
- ***But these results are much inferior than what promised..*** *(even more in the 70s).*



A physical symbol system has the *necessary* and *sufficient* means for general intelligent action

Allen Newell and Herbert A. Simon

Computer Science as Empirical Inquiry: Symbols and Search (1976)

Acknowledged limitations

- **knowledge acquisition bottleneck**
- **scaling or modularity**
- **tractability** (e.g. *ramification problem*)
- **symbol grounding**

Acknowledged limitations

- **knowledge acquisition bottleneck**
- **scaling or modularity**
- **tractability** (e.g. *ramification problem*)
- **symbol grounding**
- natural language
- sensory-motor tasks
  - *computer vision,*
  - *speech recognition,*
  - *actuator control*

Acknowledged limitations

- **knowledge acquisition bottleneck**
- **scaling or modularity**
- **tractability** (e.g. *ramification problem*)
- **symbol grounding**

- natural language
- sensory-motor tasks
  - *computer vision,*
  - *speech recognition,*
  - *actuator control*

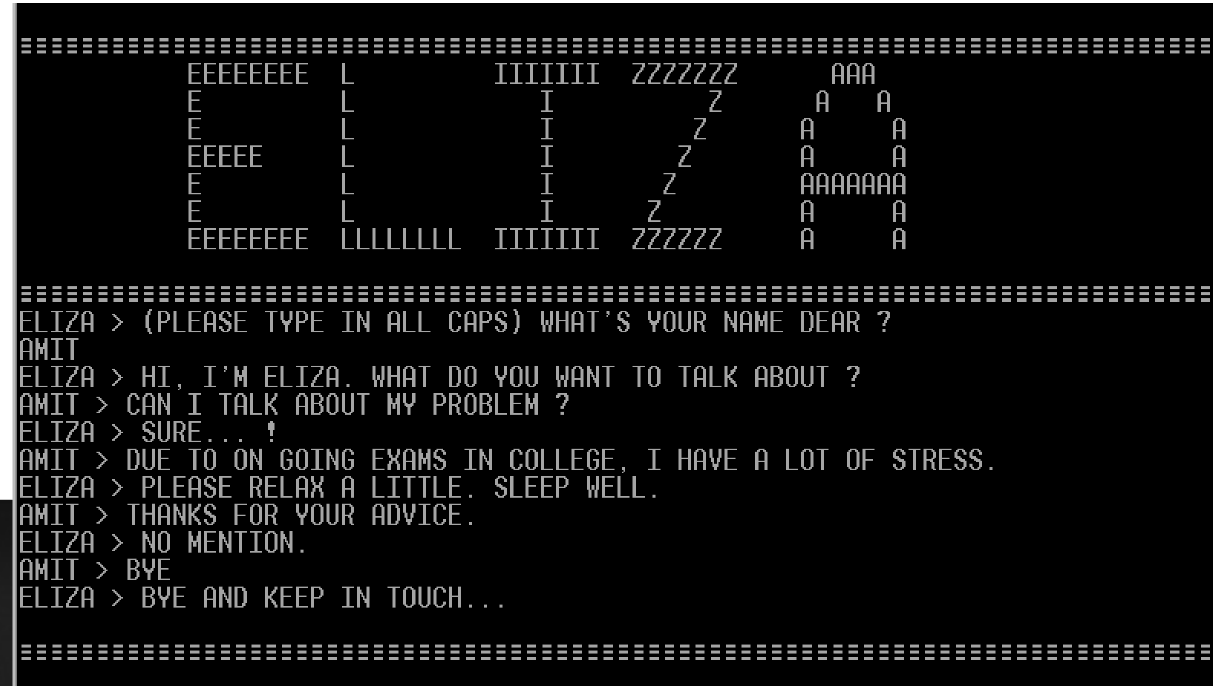
Hacking solutions

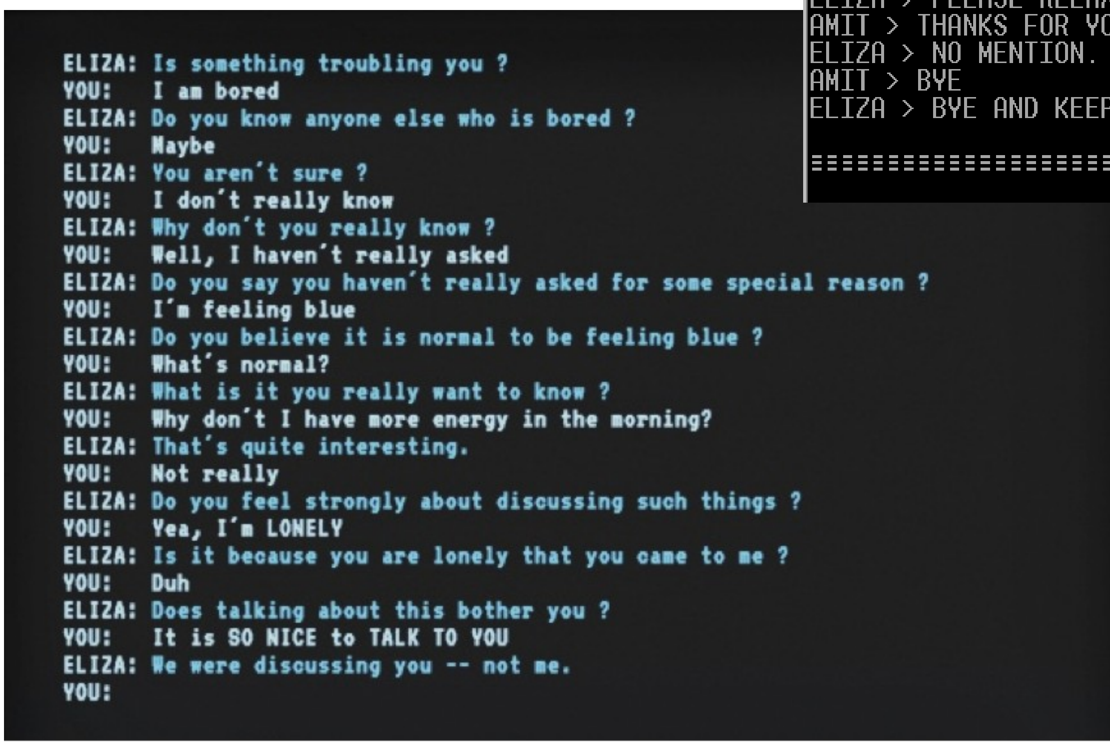
- Scruffies never believed the mind was a monolithical system, so they tinkered with heuristics, *ad-hoc* methods, and opportunistically with logic ("neat shells for scruffy approaches").

(the first chatbot)

ELIZA

Weizenbaum ~1965

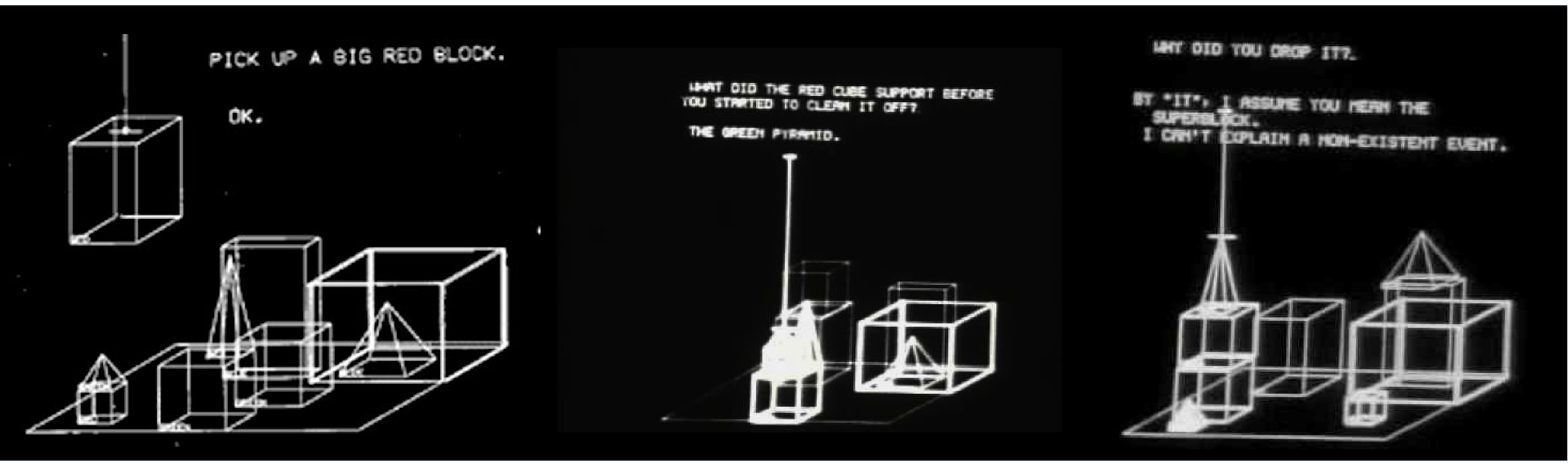




Still running e.g. on: https://www.masswerk.at/elizabot/eliza.html

SHRDLU

Winograd ~1969



- Deeper linguistic understanding
- but limited to simple *blocks* worlds

https://hci.stanford.edu/~winograd/shrdlu/

Acknowledged limitations

- **knowledge acquisition bottleneck**
- **scaling or modularity**
- **tractability** (e.g. *ramification problem*)
- **symbol grounding**
- natural language
- sensory-motor tasks
  - *computer vision,*
  - *speech recognition,*
  - *actuator control*

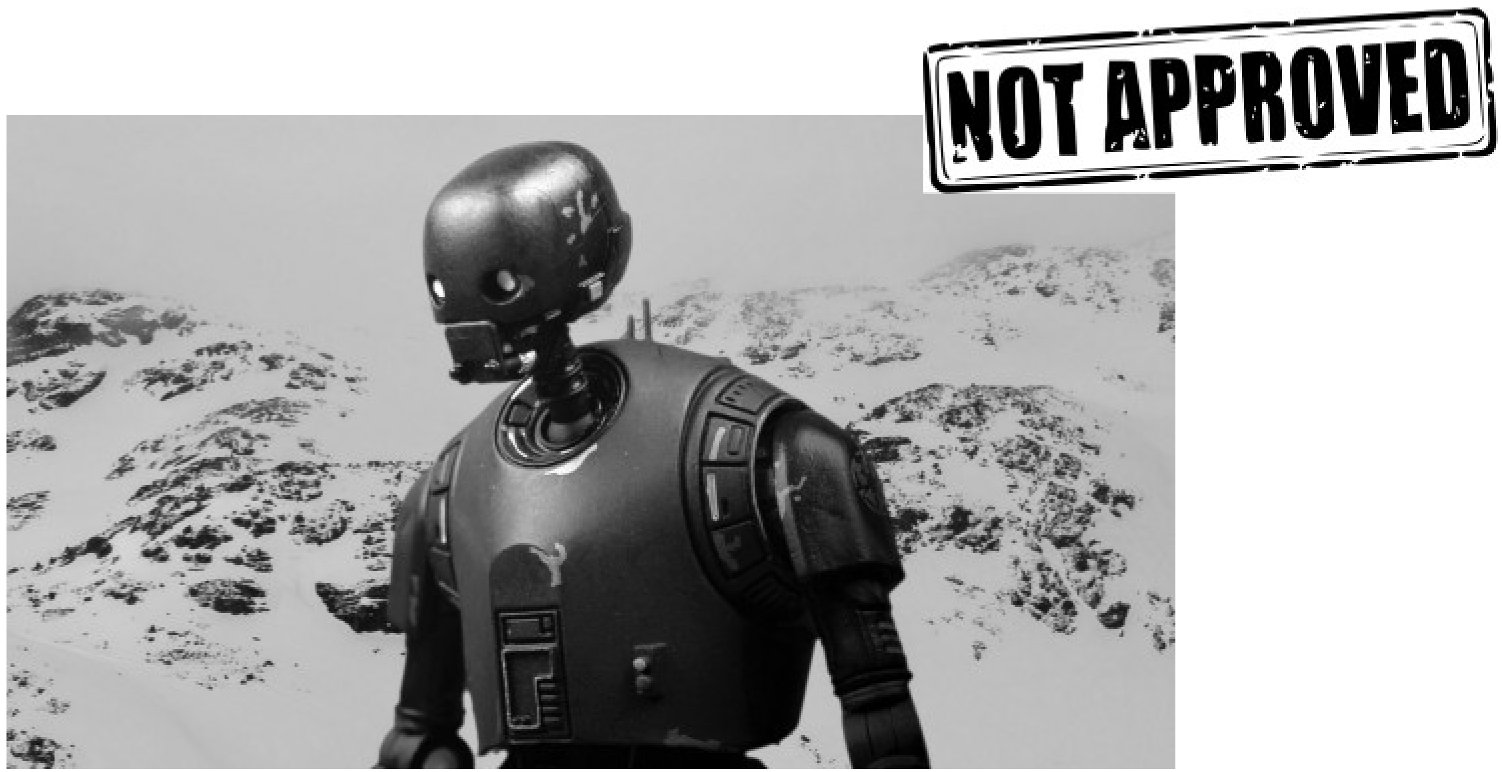
Hacking solutions

Scruffies never believed the mind was a monolithical system, so they tinkered with heuristics, *ad-hoc* methods, and opportunistically with logic ("neat shells for scruffy approaches").

- **but these successes were impossible to generalize.**

AI Winter (early 70s/80s)

- After a series of critical reports, funding to AI projects reduced massively. Researchers started to seek other names for their own research fields.



- Facing overwhelming difficulties to go beyond from toy problems, radically different paradigms started to be (re)considered, **renouncing** to symbolic representations.
- As Rodney Brooks famously put it:

“Elephants don't play chess”

The revenge of machine learning

Machine learning

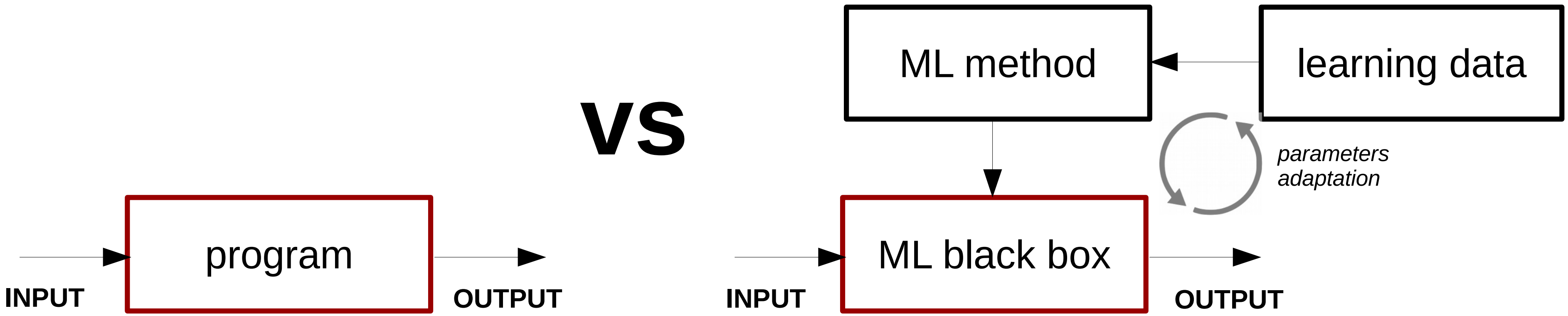
Machine learning is a process that enables artificial systems to improve from experience.

according to well-defined criteria

Machine learning

Machine learning is a process that enables artificial systems to improve from experience.

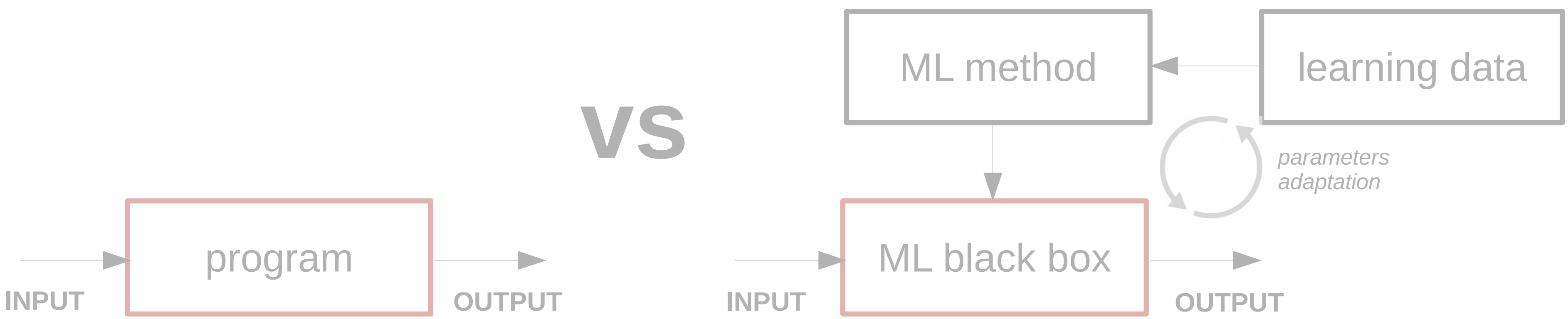
- Rather then writing a program, here the developer has to collect adequate training data and decide a ML method.



function approximation

How to adapt to training data is often not straightforward!

- Rather then writing a program, here the developer has to collect adequate training data and decide a ML method.

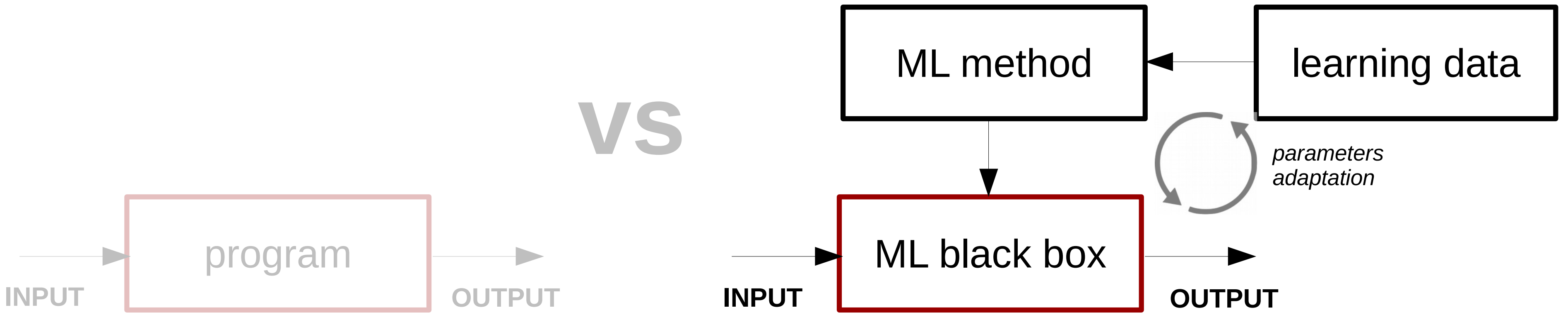


function approximation

Machine learning

Machine learning is a process that enables artificial systems to improve from experience.

- Rather then writing a program, here the developer has to collect adequate training data and decide a ML method.



- Unfortunately, an **adequate parameter adaptation can be highly data-demanding**, especially for rich inputs.

Machine learning

Machine learning is a process that enables artificial systems to improve from experience.

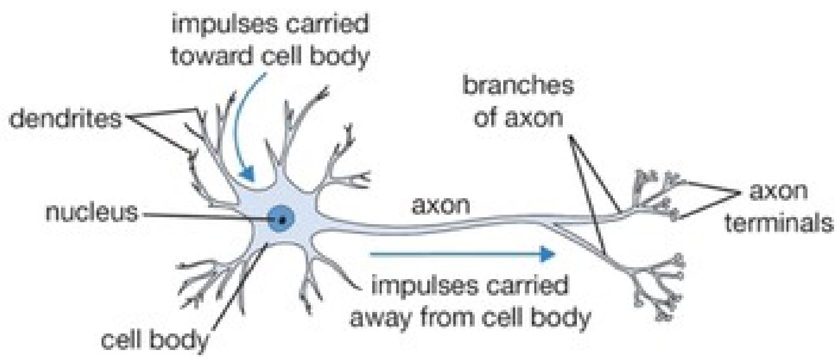
- Many learning methods are available, but studied and used by different communities!
- **Neural networks** are only one among many.

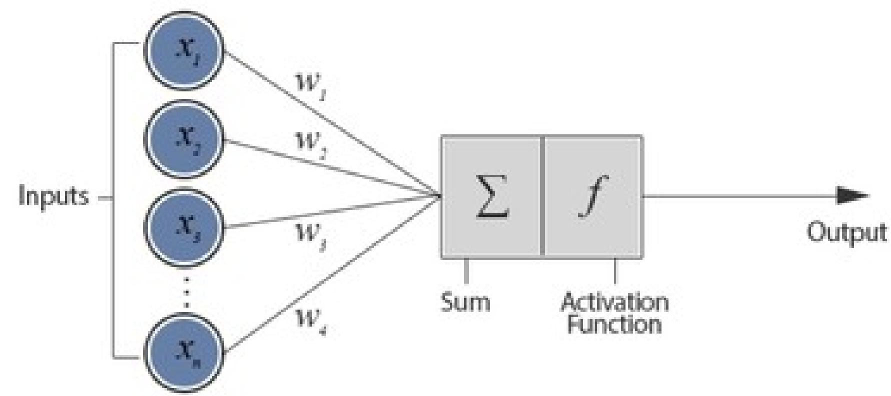
  (e.g. ***evolutionary algorithms*** can also be of use)

Nice video applying evolutionary algorithms: https://www.youtube.com/watch?v=pgaEE27nsQw

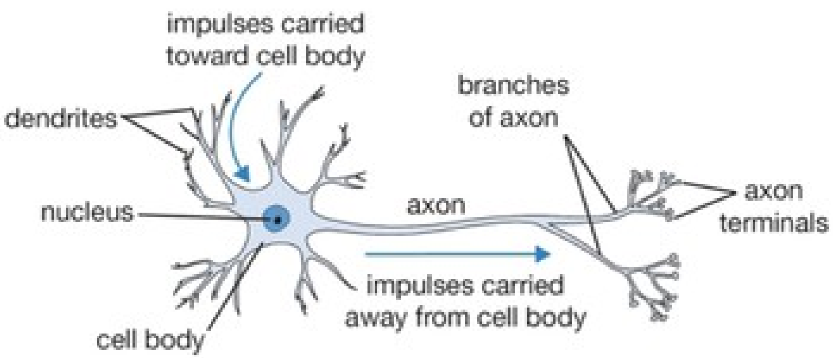
From T. Geijtenbeek, M. van de Panne, F. van der Stappen, Flexible Muscle-Based Locomotion for Bipedal Creatures. In ACM Transactions on Graphics, Vol. 32, Nr. 6 (Proc. of SIGGRAPH Asia 2013)

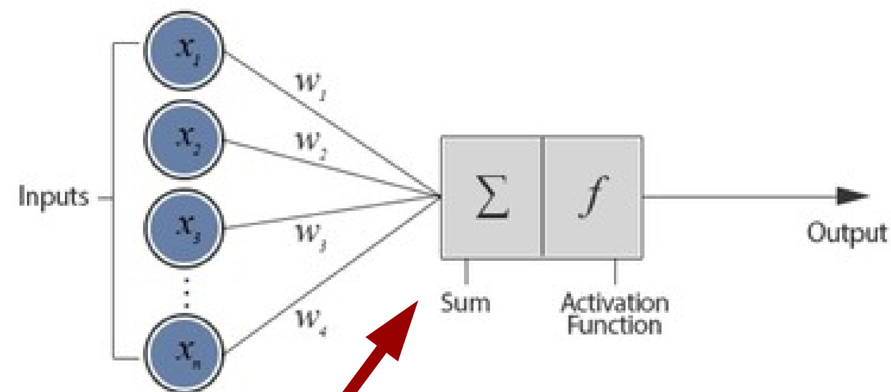
Biological neurons vs ANN nodes





Biological neurons vs ANN nodes

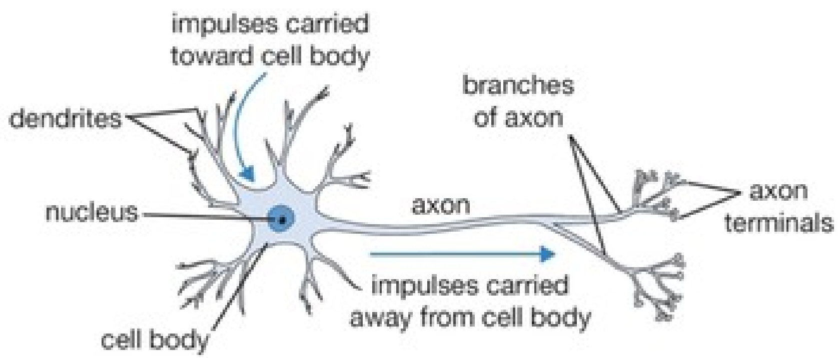


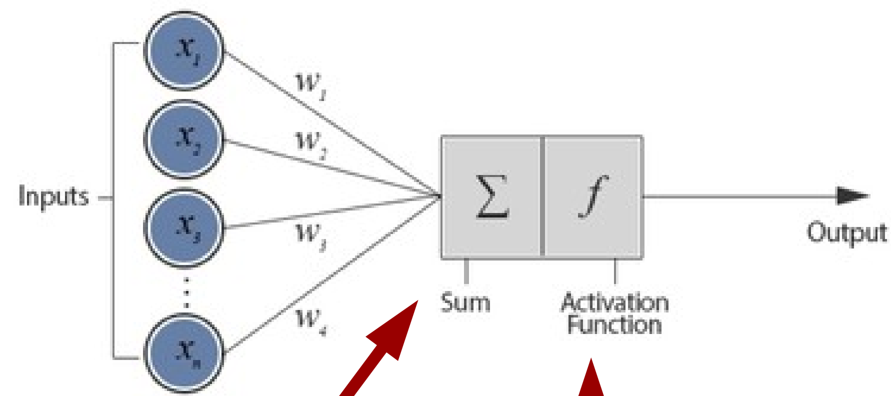


weighted
accumulation

artificial neural networks

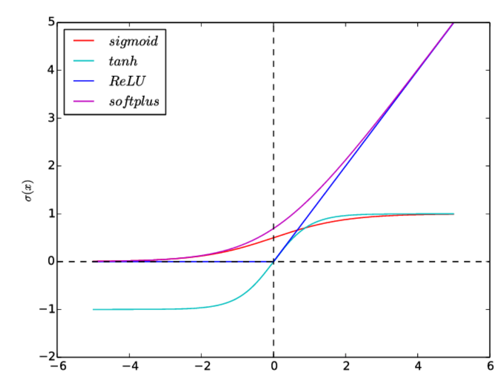
Biological neurons vs ANN nodes



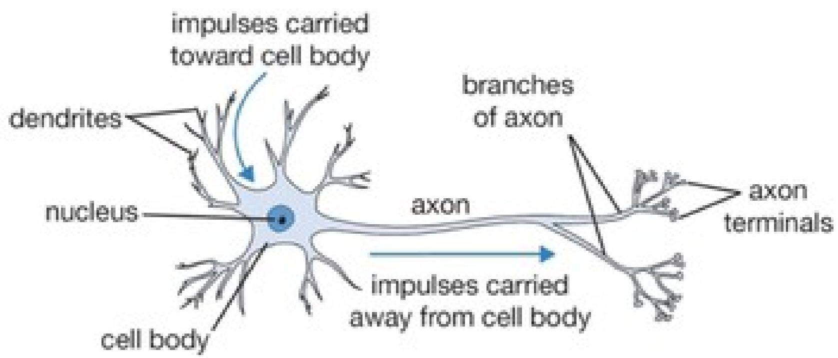


weighted accumulation

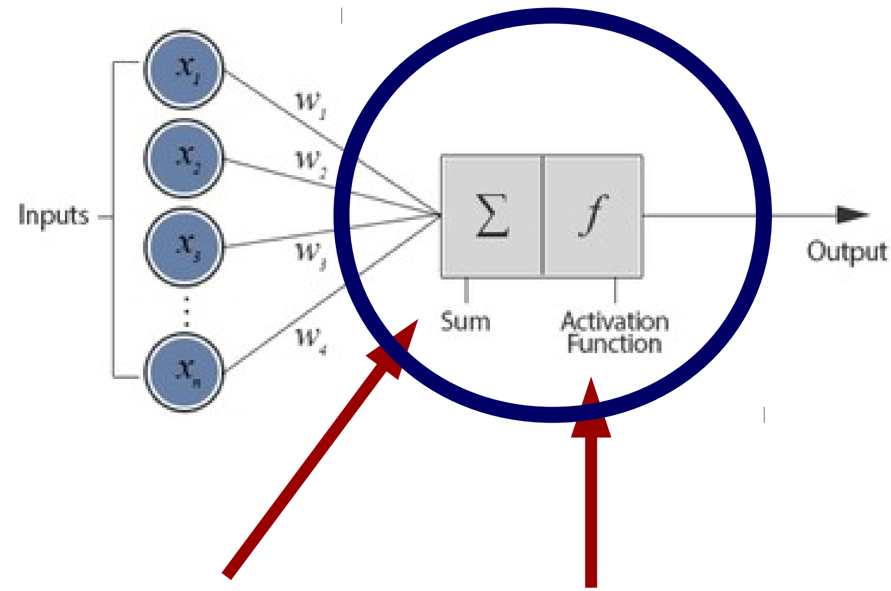
non-linearization



Biological neurons vs ANN nodes

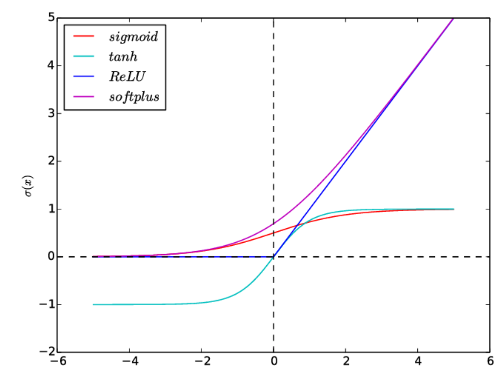


a sort of informational filter



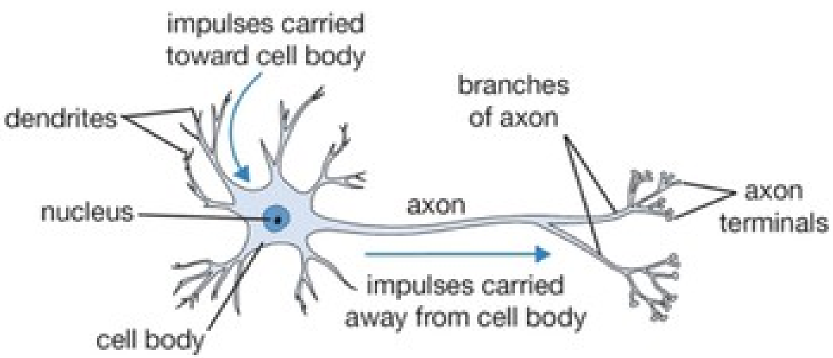
weighted accumulation

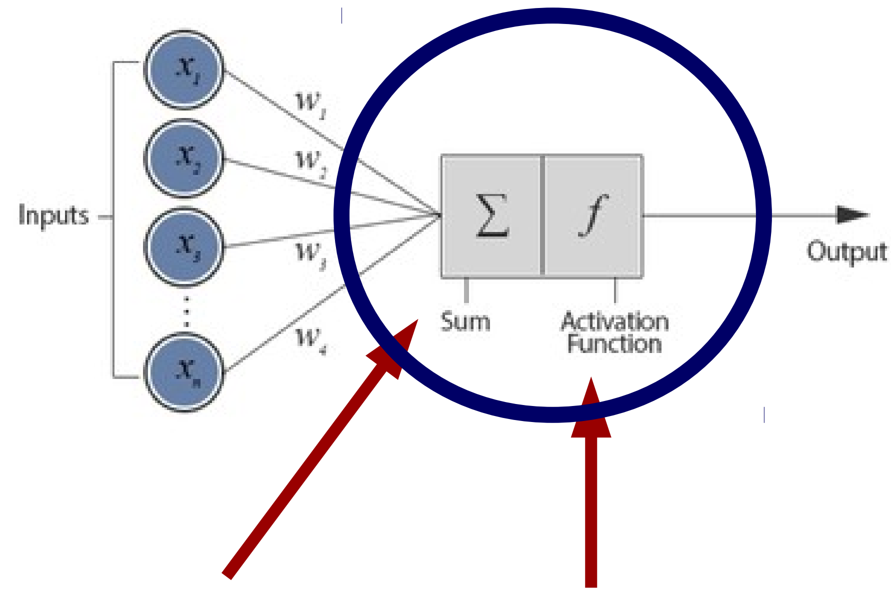
non-linearization



Biological neurons vs ANN nodes

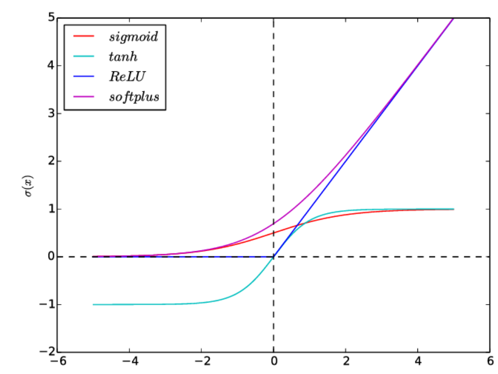
a sort of informational filter





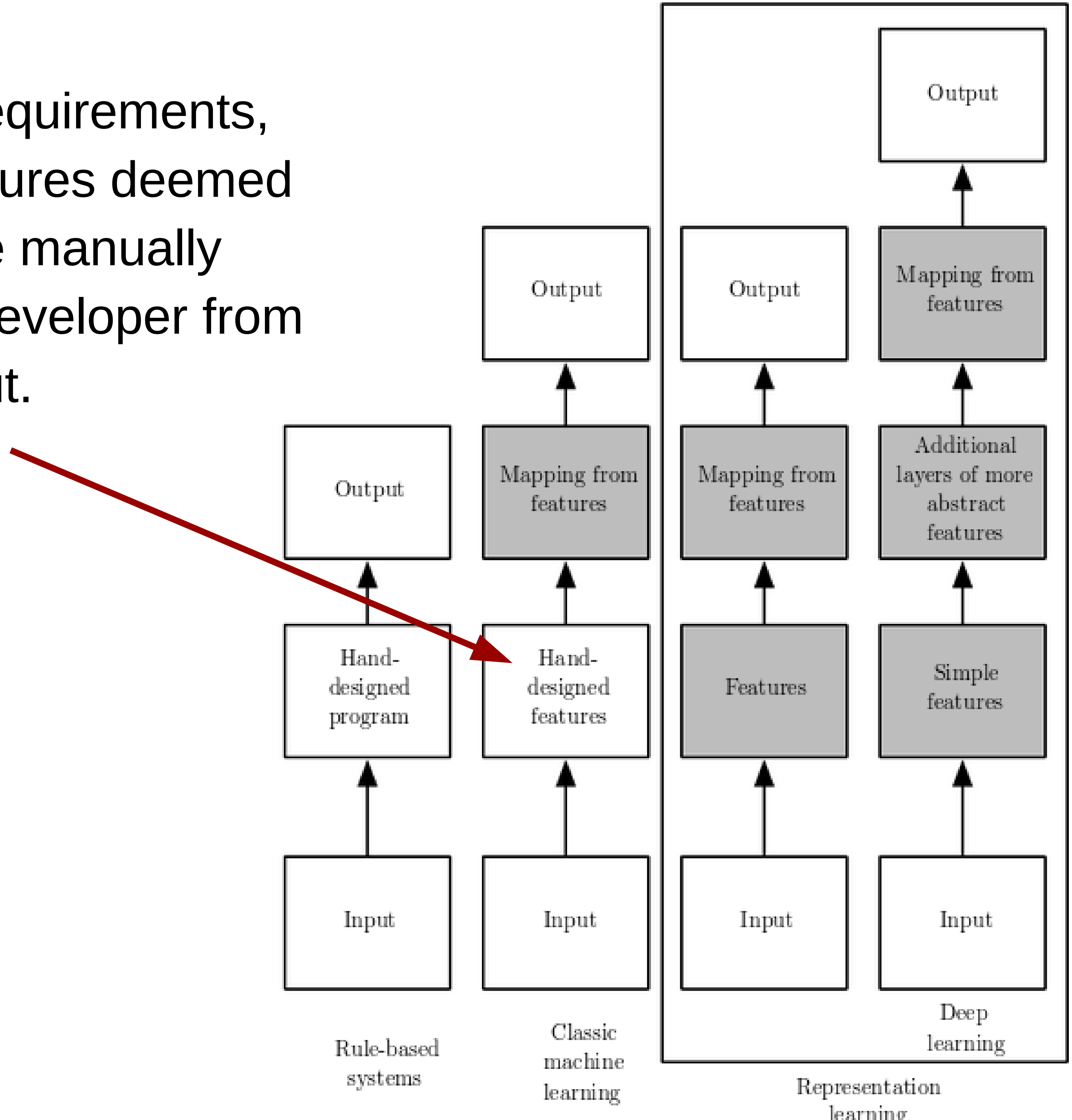
weighted accumulation

non-linearization



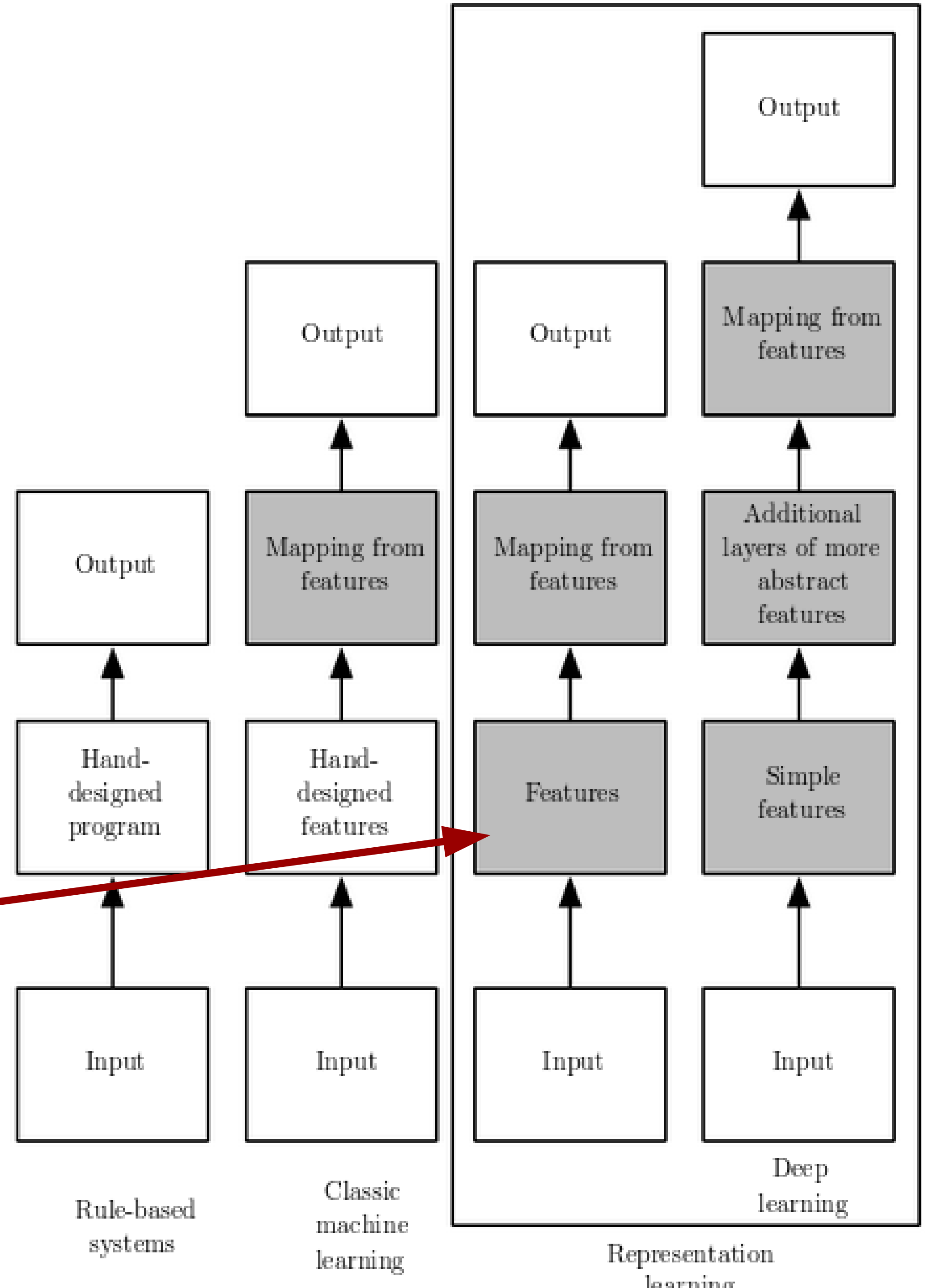
- A multi-layered artificial neural network is similar to a ***cascade* of filters**, that can be used to extract what is relevant and transform it adequately.

- To reduce data requirements, in classic ML features deemed to be relevant are manually selected by the developer from the available input.

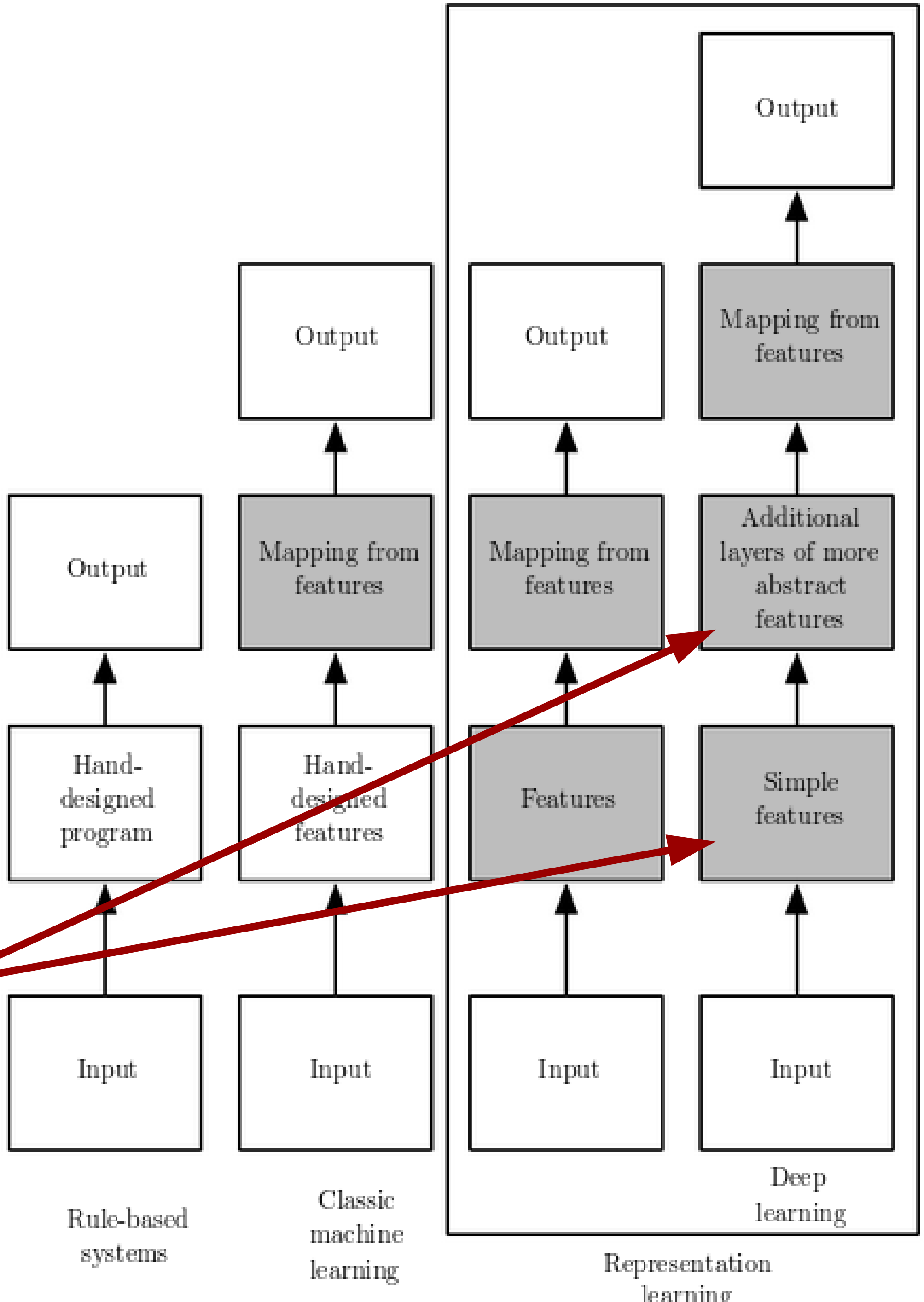


Goodfellow, Benjo and Courville, "Deep Learning" (2016)

- When this is not possible, features have to be extracted as well, through some ***representation learning***.

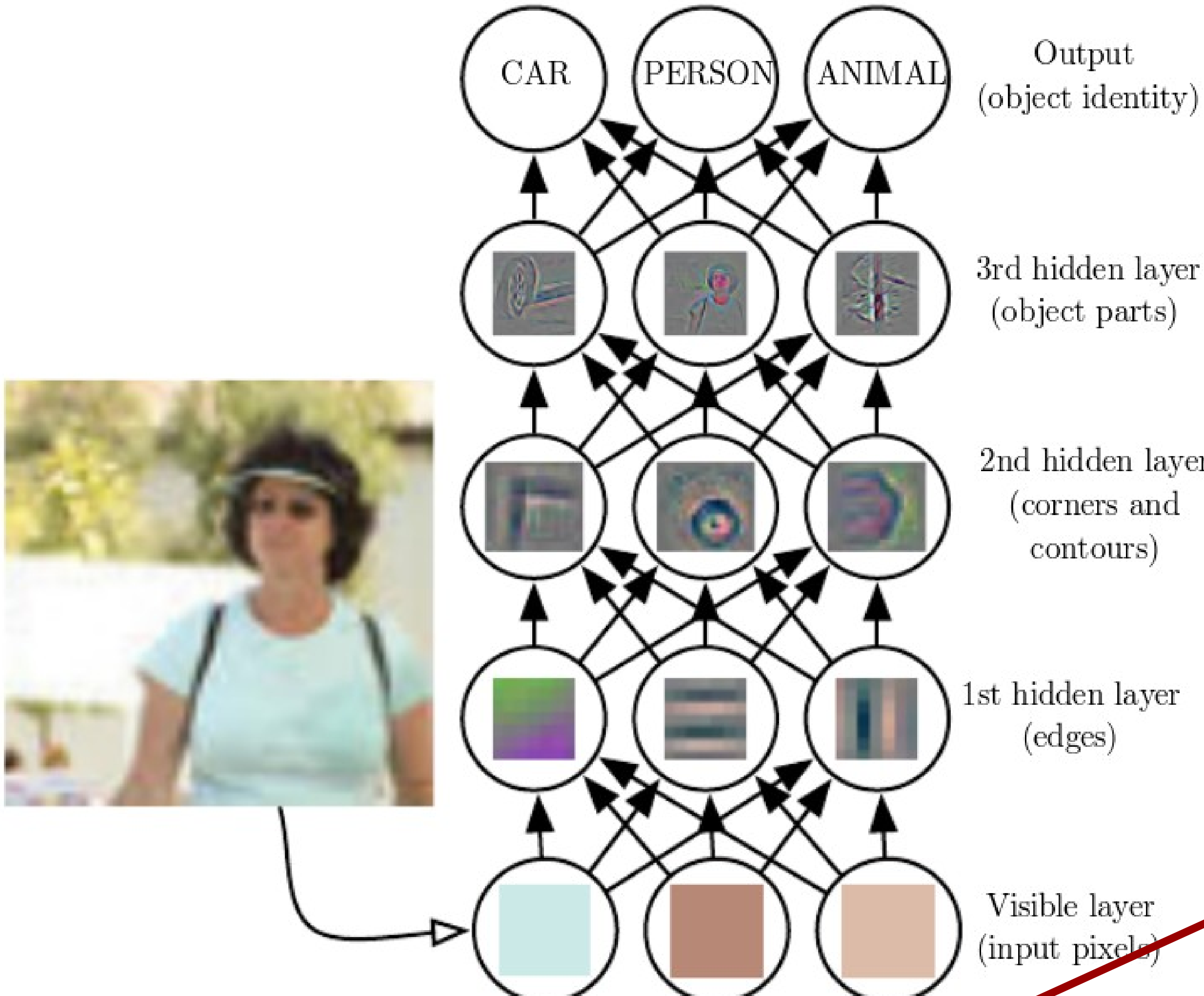


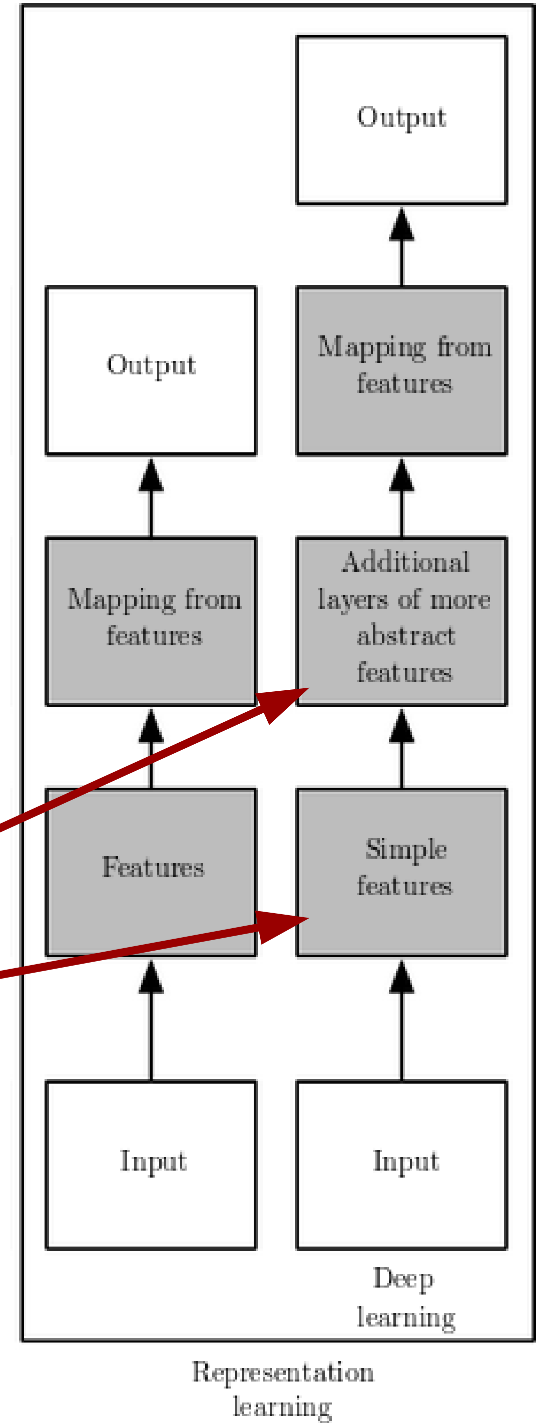
Goodfellow, Benjo and Courville, "Deep Learning" (2016)



- ***Deep learning*** relies on a **hierarchy** of representation learning, producing different level of abstractions

Goodfellow, Benjo and Courville, "Deep Learning" (2016)





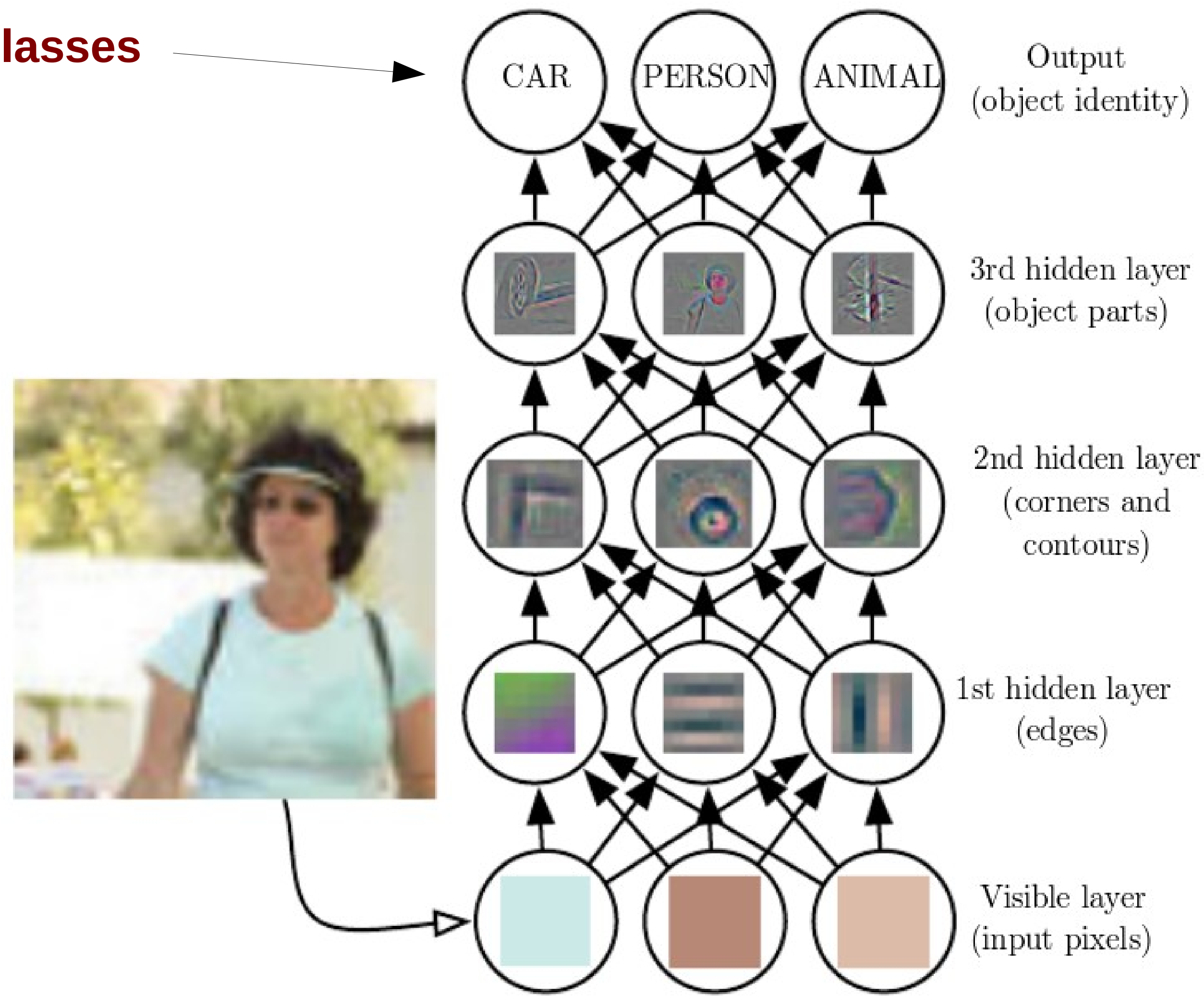
- ***Deep learning*** relies on a **hierarchy** of representation learning, producing different level of abstractions

Goodfellow, Benjo and Courville, "Deep Learning" (2016)

Example 1: Classifiers

Classifiers associate symbols to (typically) non-symbolic inputs

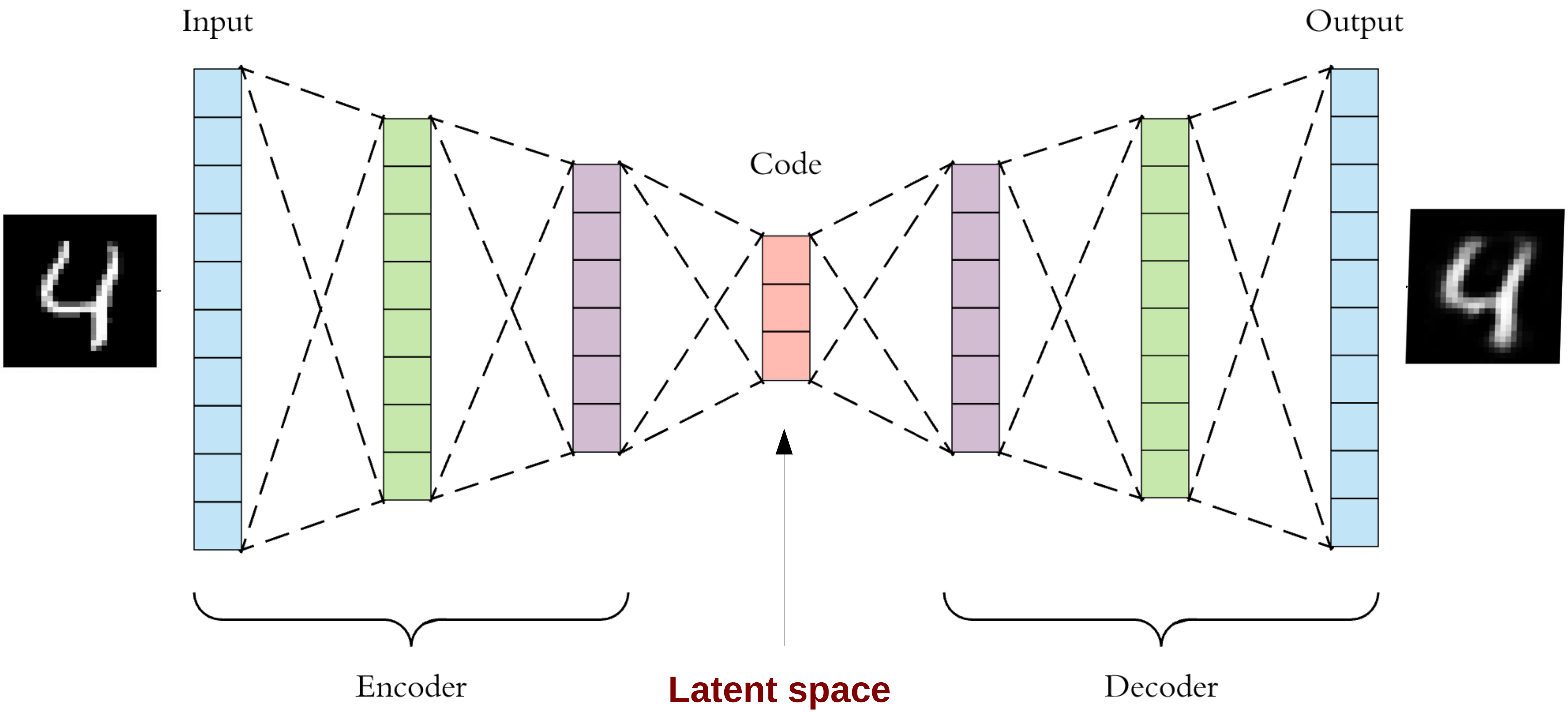
Target classes



Supervised method: you need the labels for the training set.

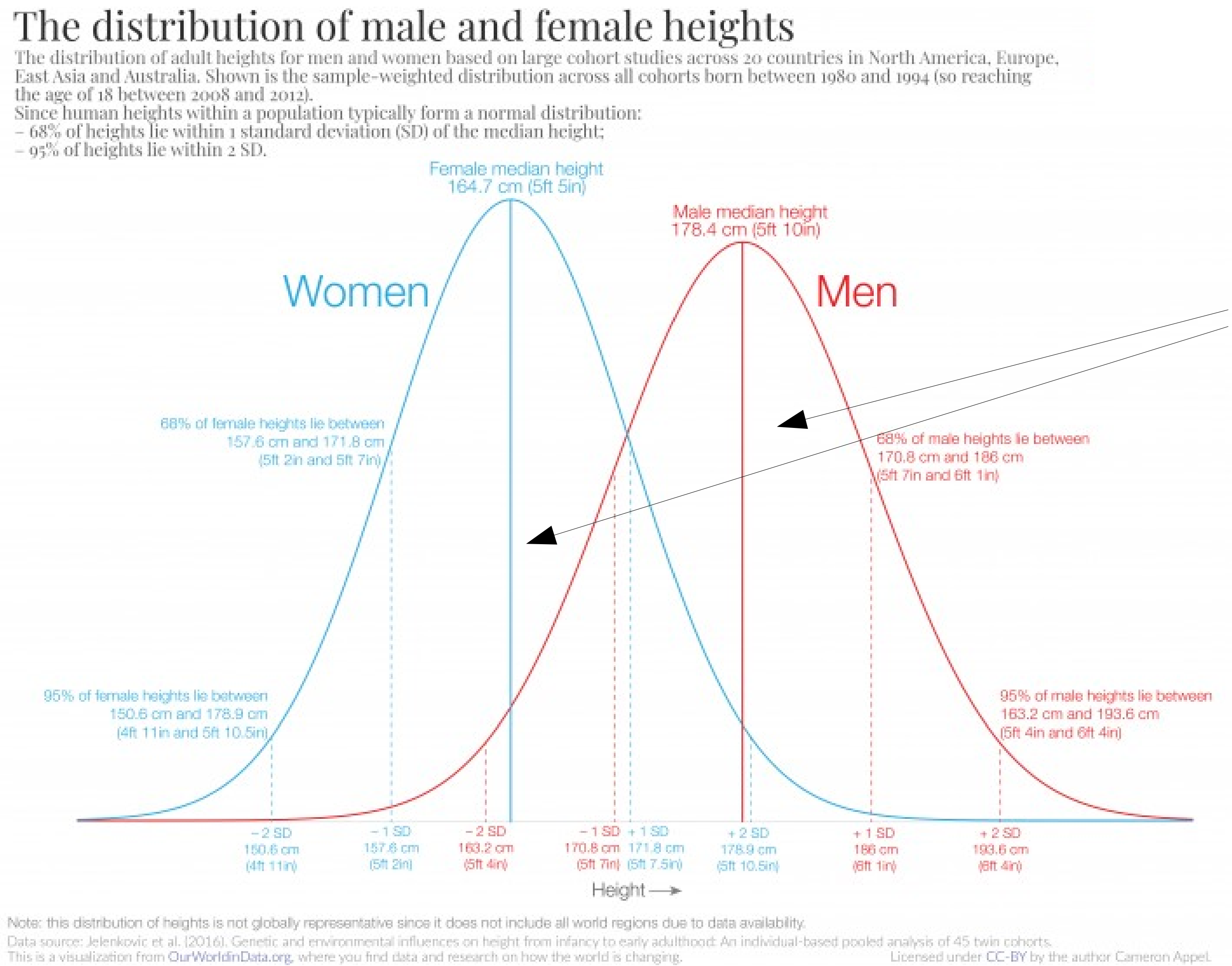
Example 2: Autoencoders

Autoencoders function as “compressors”: what are the features that make the output as much similar as possible to the input



Unsupervised method: you just need the data.

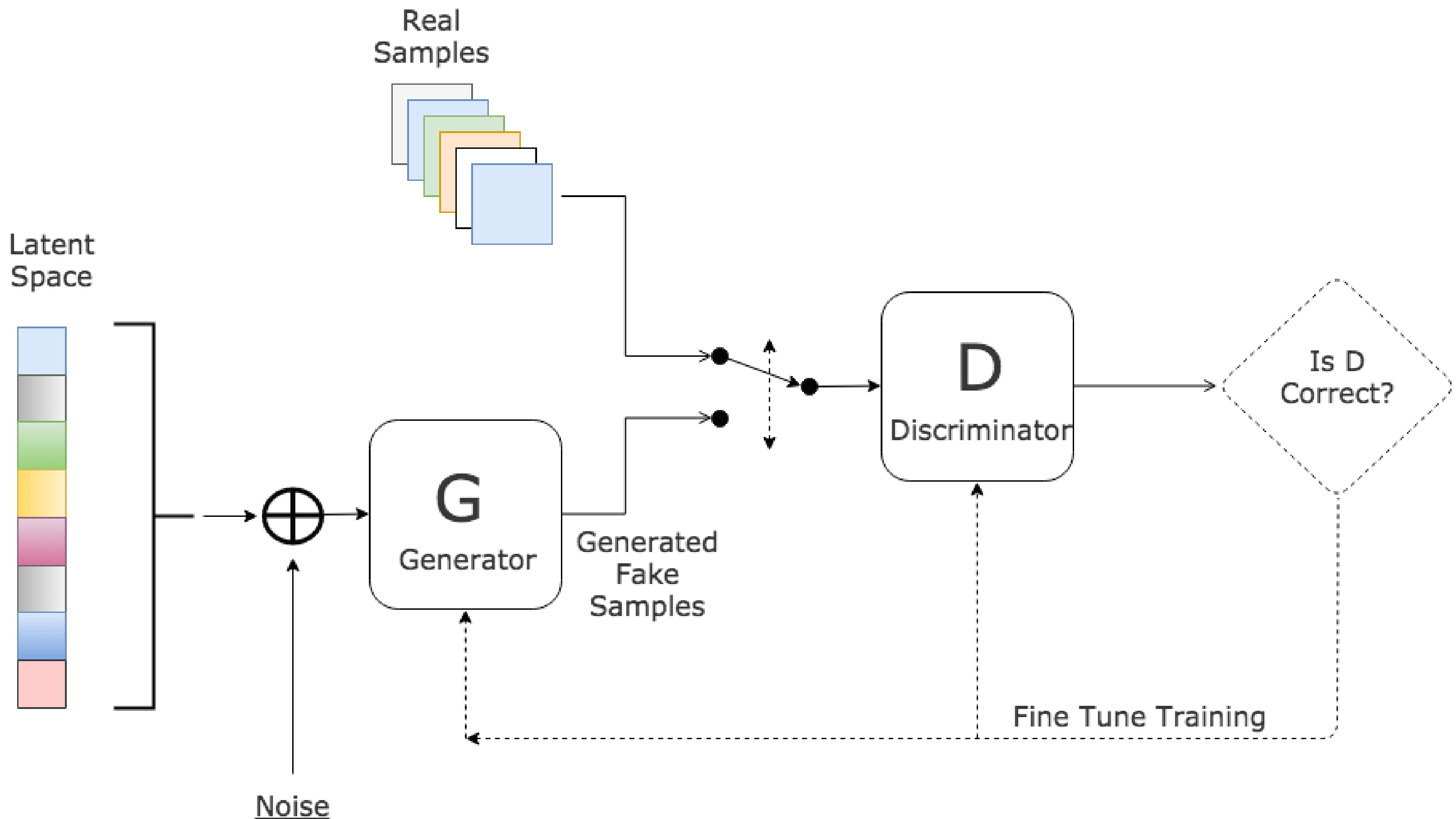
Why it works? Plausible it targets convexity regions in the wild



maintain the protoypes, not the instances

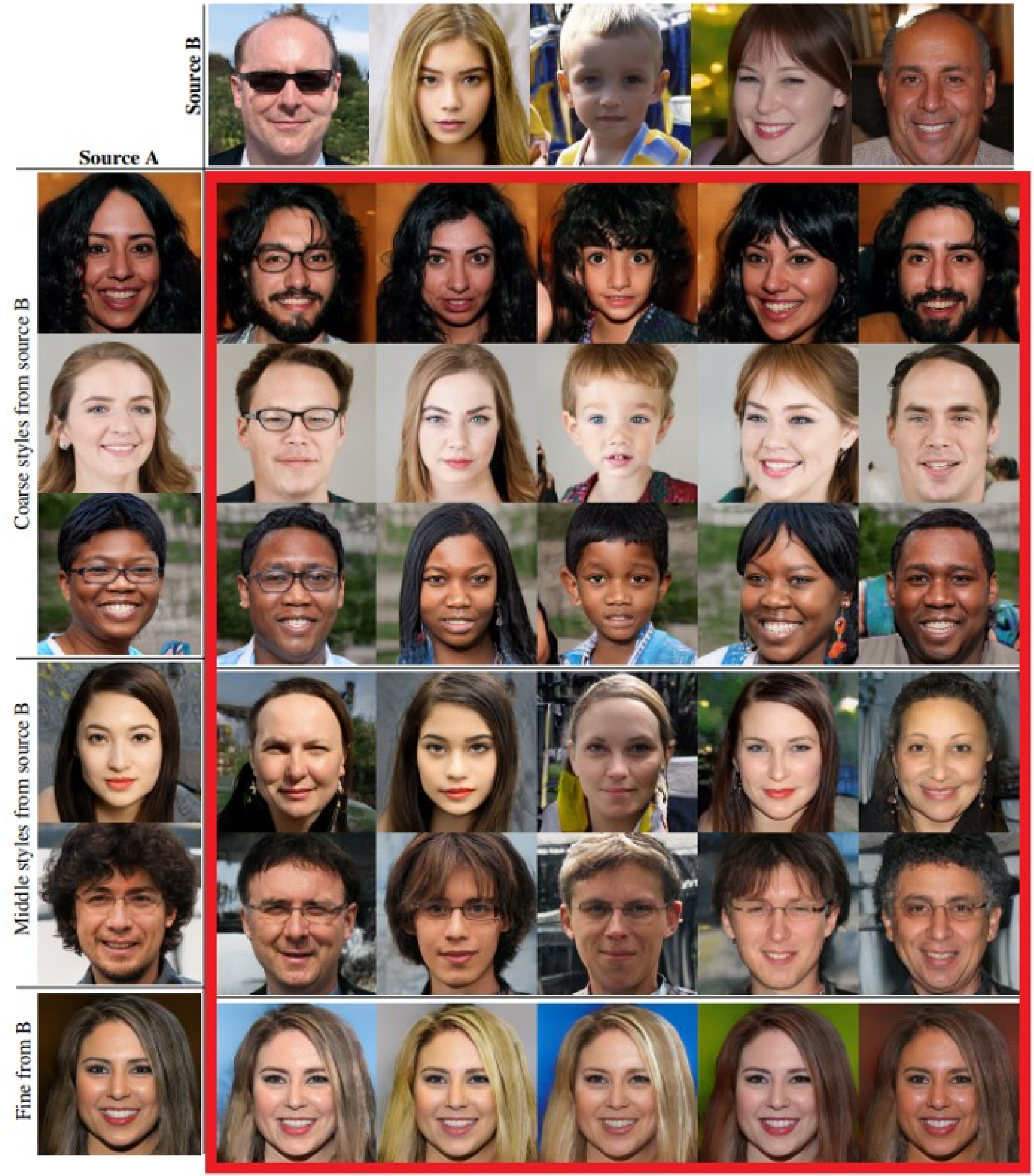
Example 3: Generative Adversarial Networks (GANs)

GANs are based on training both a discriminator (true/fake sample) and a generator. This allows to improve the discrimination by using artificially constructed inputs and to improve the generator by means of the discriminator.

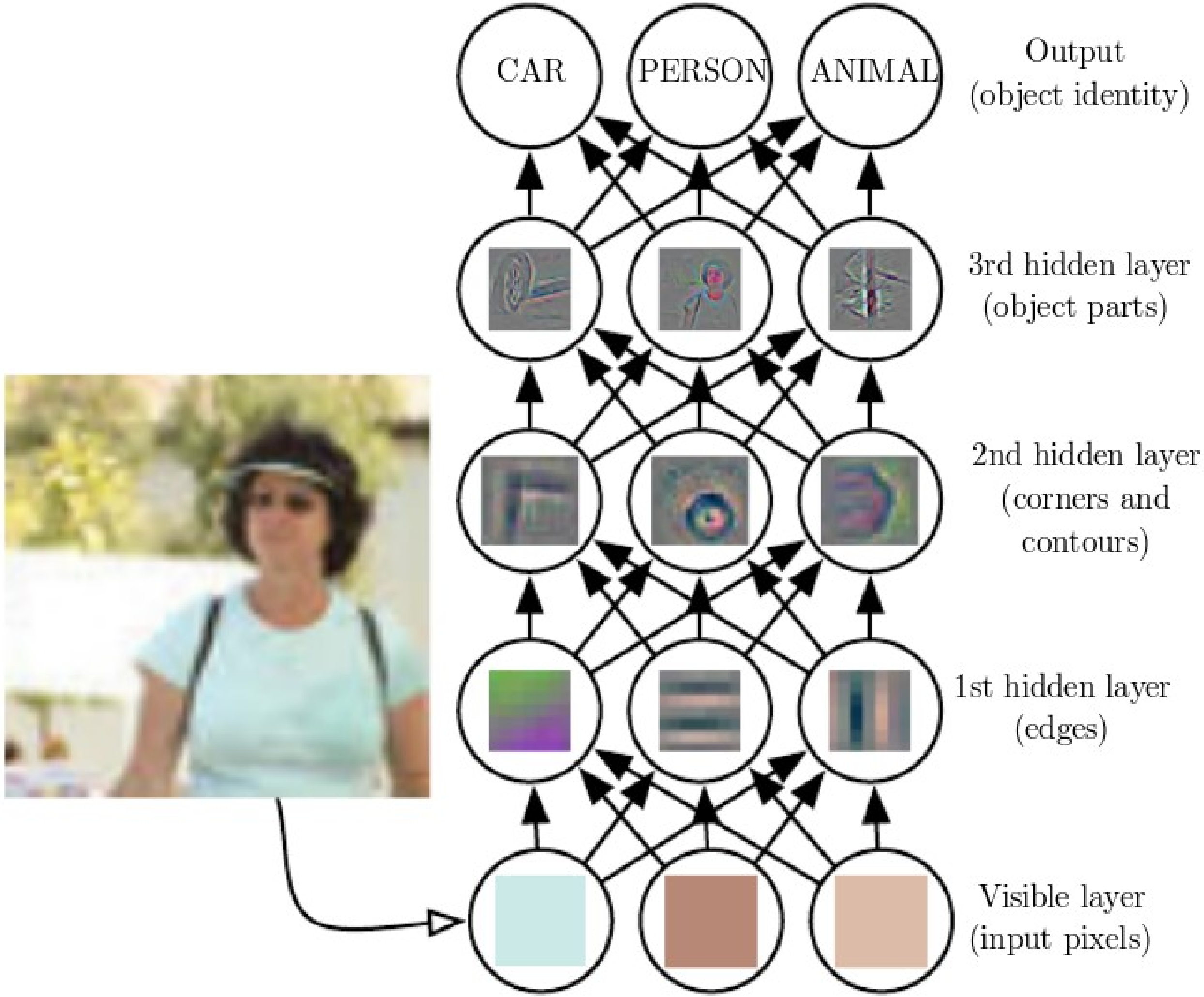


Unsupervised method: you just need the data.

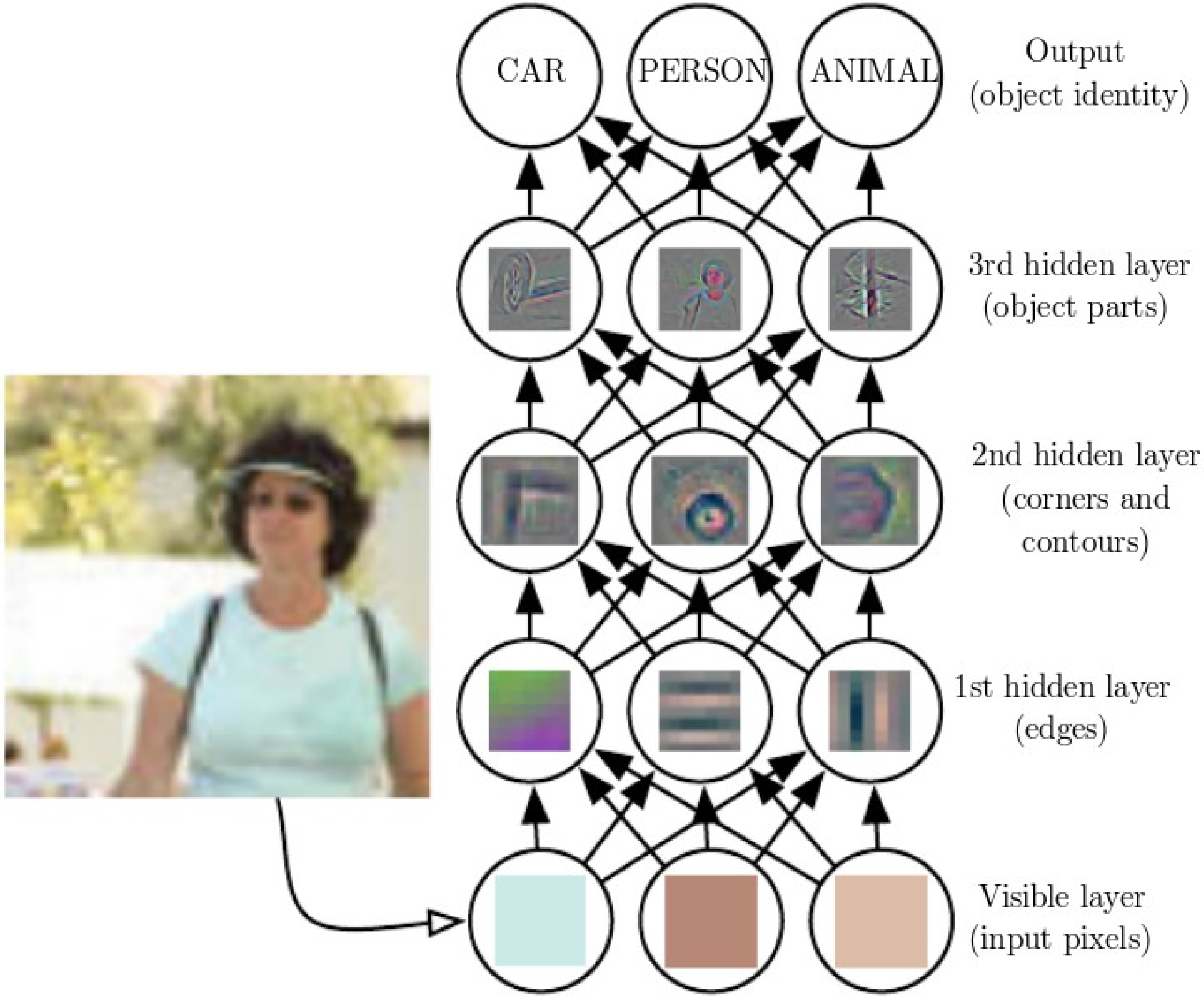
Example of what you can do with a latent space:



Tero Karras, Samuli Laine, Timo Aila, A Style-Based Generator Architecture for Generative Adversarial Networks, CVPR 2019



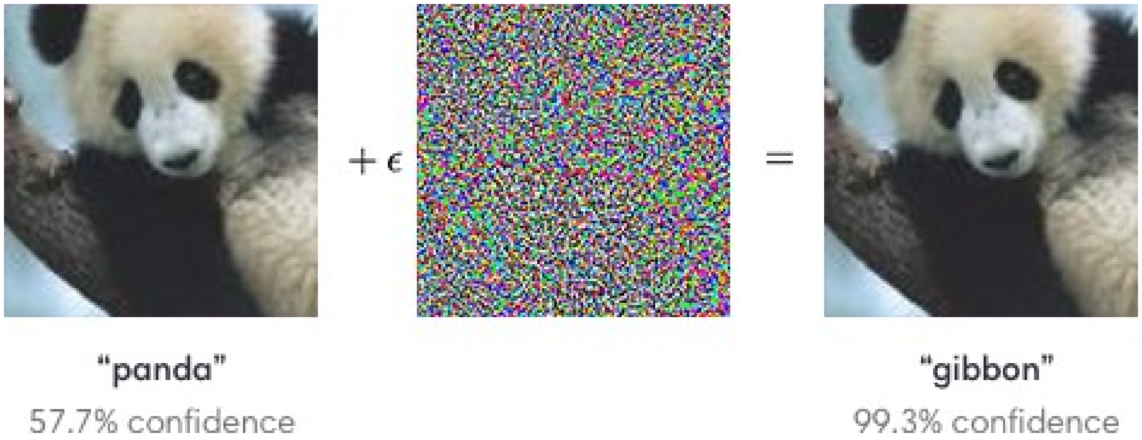
- **Problem**: the developer does not have direct control on which features are considered to be **relevant** to the task.



even qualitative introspection is subjective and arguable

- **Problem**: the developer does not have direct control on which features are considered to be **relevant** to the task.

Adversarial attacks



- Knowing what is deemed of attention by the machine can be exploited by an attacker can produce targeted "optical illusions" for the machine, but not for us.

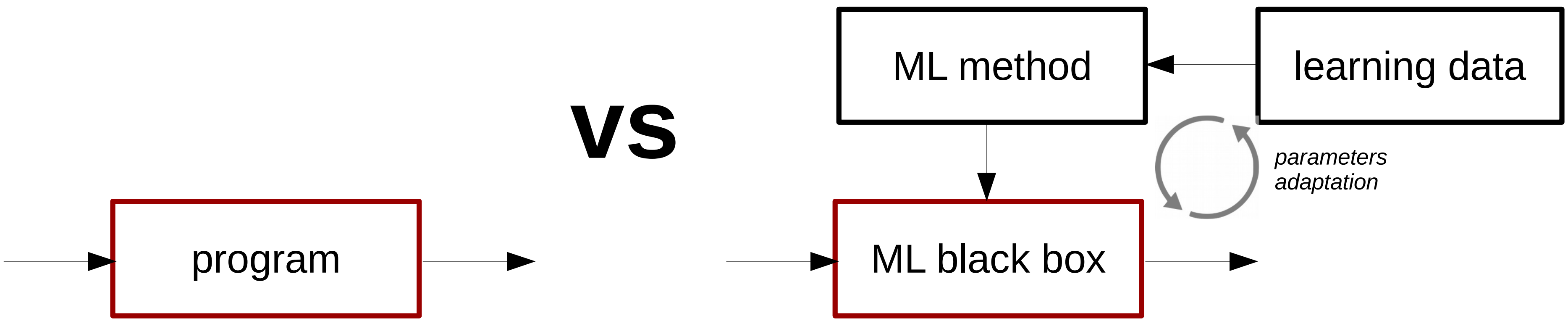
https://blog.openai.com/adversarial-example-research/

Using encoding/decoding abilities of deep learning

- On the other hand, knowing what is relevant to *our vision,* someone can play dirty tricks.

Face to face: https://www.youtube.com/watch?v=ohmajJTcpNk
Voice to lips: https://www.youtube.com/watch?v=9Yq67CjDqvw

From software/knowledge engineering to data engineering



- Clearly, the outcome of applying a ML method critically depends on the training data.

From software/knowledge engineering to data engineering

- Country A’s army demands a classifier to recognize whether a tanks is from country A or country B. It provides the developers with a series of photos of tanks from both countries.

From software/knowledge engineering to data engineering

- Country A’s army demands a classifier to recognize whether a tanks is from country A or country B. It provides the developers with a series of photos of tanks from both countries.
- After the training, the developers investigate by introspection the activation patterns. They discover that “**daylight**” is a major factor supporting a B-tank classification. Returning on the source data, the developers discovered that there was *no photo of B-tanks at night*.

From software/knowledge engineering to data engineering

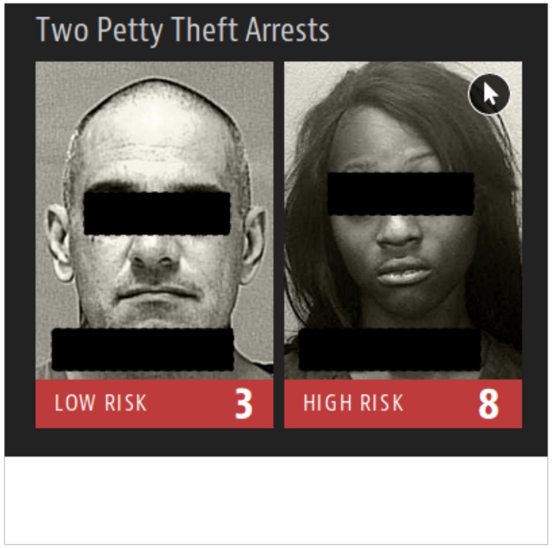
- Country A's army demands a classifier to recognize whether a tanks is from country A or country B. It provides the developers with a series of photos of tanks from both countries.
- After the training, the developers investigate by introspection the activation patterns. They discover that "**daylight**" is a major factor supporting a B-tank classification. Returning on the source data, the developers discovered that there was *no photo of B-tanks at night*.

statistical biases endanger ML predictive abilities

(LOW DATA QUALITY)

On the “artificial prejudice”

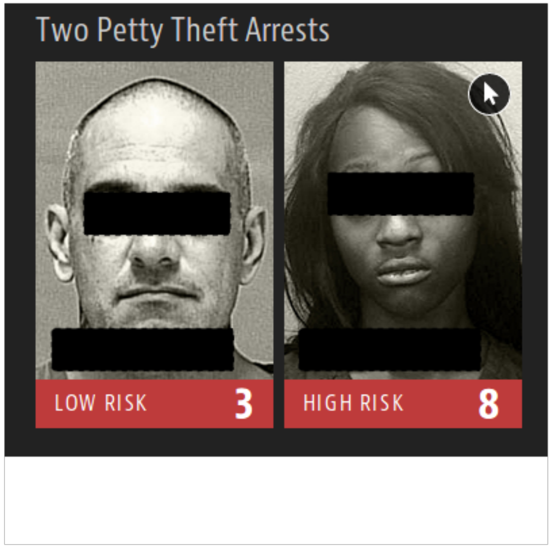
- The large-scale application of statistical-based methods for legally-relevant decisions raises several concerns:
  - COMPAS: software used in the US predicting future crimes and criminals argued to be biased against African Americans (2016)
  - SyRI (System Risk Indication) used in the Netherlands to create risk alerts for welfare frauds by processing and linking personal data of citizens argued to be discriminatory and unlawful (2018)



Angwin J. et al. ProPublica, May 23 (2016). *Machine Bias: risk assessments in criminal sentencing*
https://pilpnjcm.nl/en/proceedings-risk-profiling-dutch-citizens-syri/

On the “artificial prejudice”

- The large-scale application of statistical-based methods for legally-relevant decisions raises several concerns:
  - COMPAS: software used in the US predicting future crimes and criminals argued to be biased against African Americans (2016)
  - SyRI (System Risk Indication) used in the Netherlands to create risk alerts for welfare frauds by processing and linking personal data of citizens argued to be discriminatory and unlawful (2018)



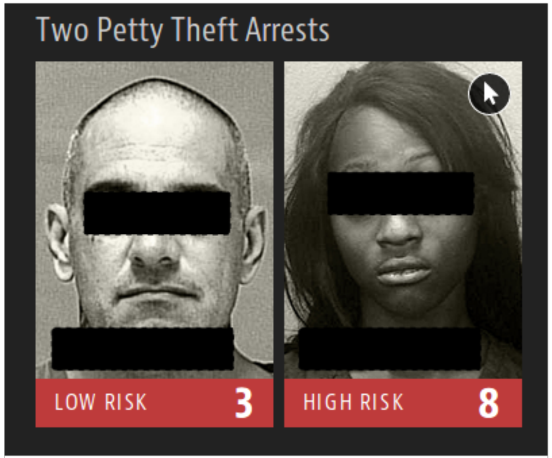
a.o. Human rights

a.o. Data protection law

Angwin J. et al. ProPublica, May 23 (2016). *Machine Bias: risk assessments in criminal sentencing*
https://pilpnjcm.nl/en/proceedings-risk-profiling-dutch-citizens-syri/

On the “artificial prejudice”

- The large-scale application of statistical-based methods for legally-relevant decisions raises several concerns:
  - COMPAS: software used in the US predicting future crimes and criminals argued to be biased against African Americans (2016)
  - SyRI (System Risk Indication) used in the



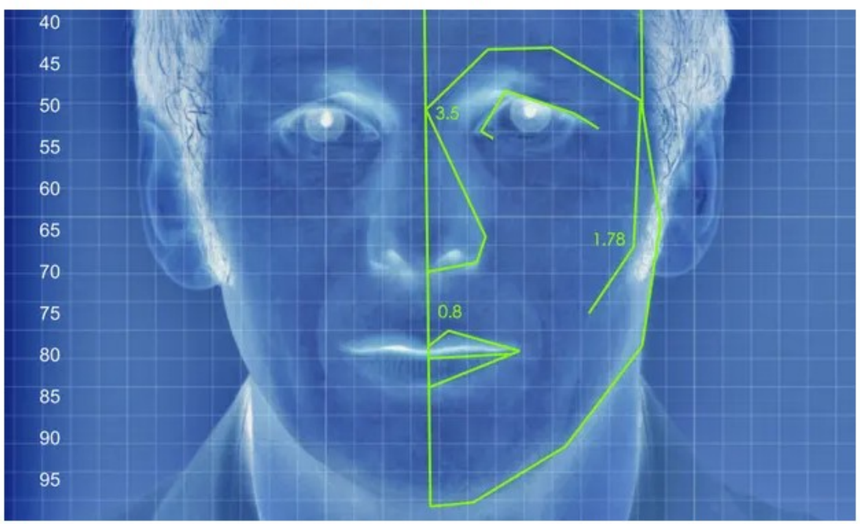
Common pattern:

- existing statistical bias (correct **description**)
- when used for prediction on an individual it is read as *behavioural predisposition*, i.e. it is interpreted as instance of a **mechanism.**
- the judgment introduces here **negative consequences** in society.

Surface vs depth / Acceptability

New AI can guess whether you're gay or straight from a photograph

An algorithm deduced the sexuality of people on a dating site with up to 91% accuracy, raising tricky ethical questions



▲ An illustrated depiction of facial analysis technology similar to that used in the experiment. Illustration: Alamy

Artificial intelligence can accurately guess whether people are gay or straight based on photos of their faces, according to new research that suggests machines can have significantly better "gaydar" than humans.

The study from Stanford University - which found that a computer algorithm could correctly distinguish between gay and straight men 81% of the time,

WIRED Technology | Science | Culture | Gear | Business | Politics | More

Privacy

Co-op is using facial recognition tech to scan and track shoppers

Branches of the Southern Co-op are using facial recognition to look for potential shoplifters. The roll-out raises concerns about the creep of surveillance tech in the private sector

https://www.wired.co.uk/article/coop-facial-recognition

(2020)

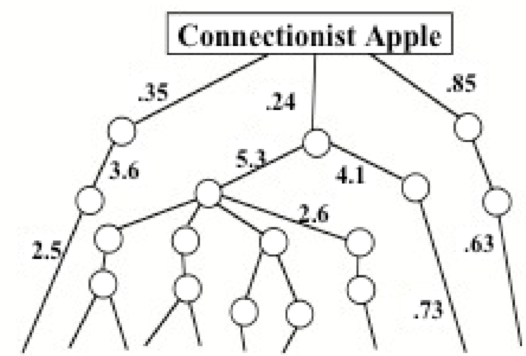
https://www.theguardian.com/technology/2017/sep/07/new-artificial-intelligence-can-tell-whether-youre-gay-or-straight-from-a-photograph

(2017)

Inside the black box

sub-symbolic AI

APPLE

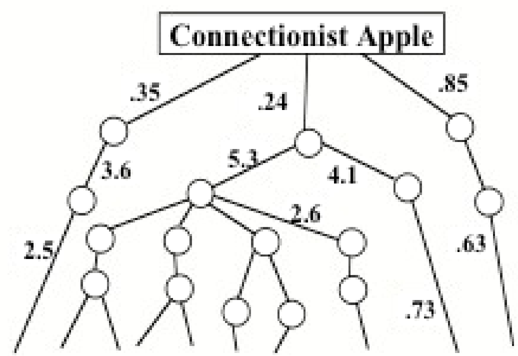


- Given a certain interpretative/behavioural model, we can extrapolate the most important features determining a certain result. e.g. we might discover that color was the main responsible for concluding that the image is about an apple.

Inside the black box

sub-symbolic AI

APPLE



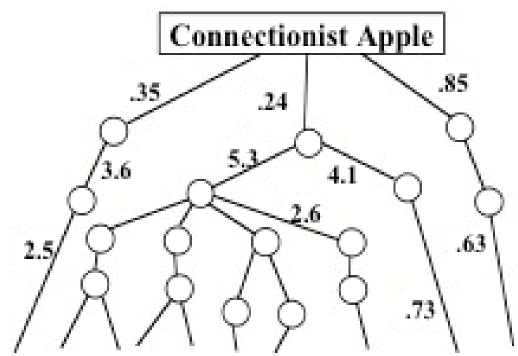
- Given a certain interpretative/behavioural model, we can extrapolate the most important features determining a certain

By “construction”, the model is made to satisfy the training samples. (What is “right” is set during training).

Inside the black box

sub-symbolic AI

APPLE



- Given a certain interpretative/behavioural model, we can extrapolate the most important features determining a certain

But what to do if what is “right” change after training?

• • •

PLOS ONE

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

A general approach for predicting the behavior of the Supreme Court of the United States

Daniel Martin Katz, Michael J. Bommarito II, Josh Blackman

Published: April 12, 2017 • https://doi.org/10.1371/journal.pone.0174698

Article | Authors | Metrics | Comments | Media Coverage

bstract

ıilding on developments in machine learning and prior work in the science of judicial ediction, we construct a model designed to predict the behavior of the Supreme Court of the ıited States in a generalized, out-of-sample context. To do so, we develop a time-evolving ndom forest classifier that leverages unique feature engineering to predict more than 240,000 stice votes and 28,000 cases outcomes over nearly two centuries (1816-2015). Using only ıta available prior to decision, our model outperforms null (baseline) models at both the justice ıd case level under both parametric and non-parametric tests. Over nearly two centuries, we :hieve 70.2% accuracy at the case outcome level and 71.9% at the justice vote level. More cently, over the past century, we outperform an *in-sample optimized* null model by nearly 5%. ır performance is consistent with, and improves on the general level of prediction ımonstrated by prior work; however, our model is distinctive because it can be applied out-of- ımple to the entire past and future of the Court, not a single term. Our results represent an ıportant advance for the science of quantitative legal prediction and portend a range of other ıtential applications.

https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0174698

JURI SAYS:

I'm predicting judgments of the European Court of Human Rights with an accuracy of 87.2% over the *last month*.

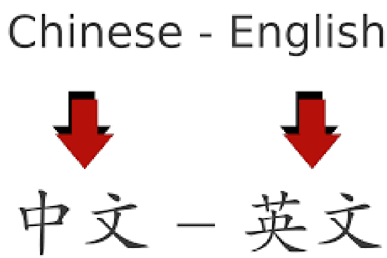
JURI reads published documents from previous years and decisions of the cases judged by the European Court of Human Rights and predicts decisions the Court will make. Every month it learns from its mistakes.

https://jurisays.com/

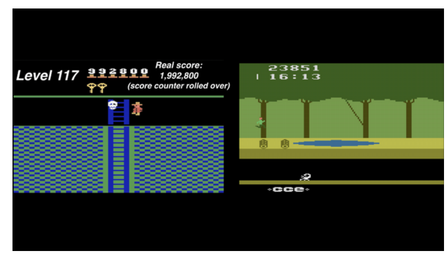
A continuous progress?

- By using a mixture of ML techniques, several human or super-human performances are achieved every year in specific tasks (*mostly by corporation-driven research*).

Google DeepMind (2016)



Microsoft (2018)



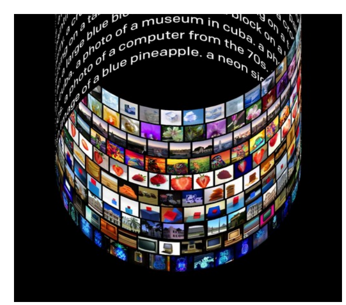
Uber (2019)

Google AlphaStar (2019)

https://www.youtube.com/watch?v=6EQAsrfUIyo

OpenAI GPT-3 (2020)

https://beta.openai.com/dashboard



OpenAI DALL·E (2021)

https://openai.com/blog/dall-e/

OpenAI GPT-3

AI Translations (AI Output in Blue)

Legal clause: Upon liquidation of the Company, the Series A Shareholders will receive in preference to all other shareholders an amount in respect of each Series A Share equal to one times the Original Issue Price (the "Liquidation Preference"), plus all accrued but unpaid dividends. To the extent that the Company has assets remaining after the distribution of that amount, the Series A Shareholders will participate with the holders of Ordinary Shares pro rata to the number of shares held on an as converted basis.

Plain English: If the startup is wound up, the Series A investors will be paid back at least what they invested and they will also share any leftover assets with ordinary shareholders.

Legal Clause: Sale of all or substantially all of the assets of the Company or a sale of shares involving a change in control (each, a "Corporate Transaction") will be treated in the same way as a liquidation and the proceeds of sale will be distributed as set out in paragraph 3. If the holders of Series A Shares have received any Special Dividend it shall be set off against their Liquidation Preference.

Plain English: If the company is sold, or a new owner takes control, the proceeds of the sale will be distributed as in the liquidation clause above. Any special dividend paid will be treated as an initial payment towards the Series A investors.

Legal Clause: If holders of at least 51% of the Series A Shares and Ordinary Shares agree to sell their shares, there will be drag along rights so that all remaining shareholders and option holders will be required to sell on the same terms, provided that the dragged shareholders will not be required to provide to the purchaser any representations or warranties except as to title or to agree to any other terms.

Plain English: If a majority of investors agree to sell their shares, any other investors who don't agree to sell will be forced to do so. They won't have to provide any new assurances or guarantees but they will have to accept the same price and terms as the others.

OpenAI DALL·E

TEXT PROMPT

an illustration of a baby daikon radish in a tutu walking a dog

AI-GENERATED IMAGES

Edit prompt or view more images ↓

TEXT PROMPT

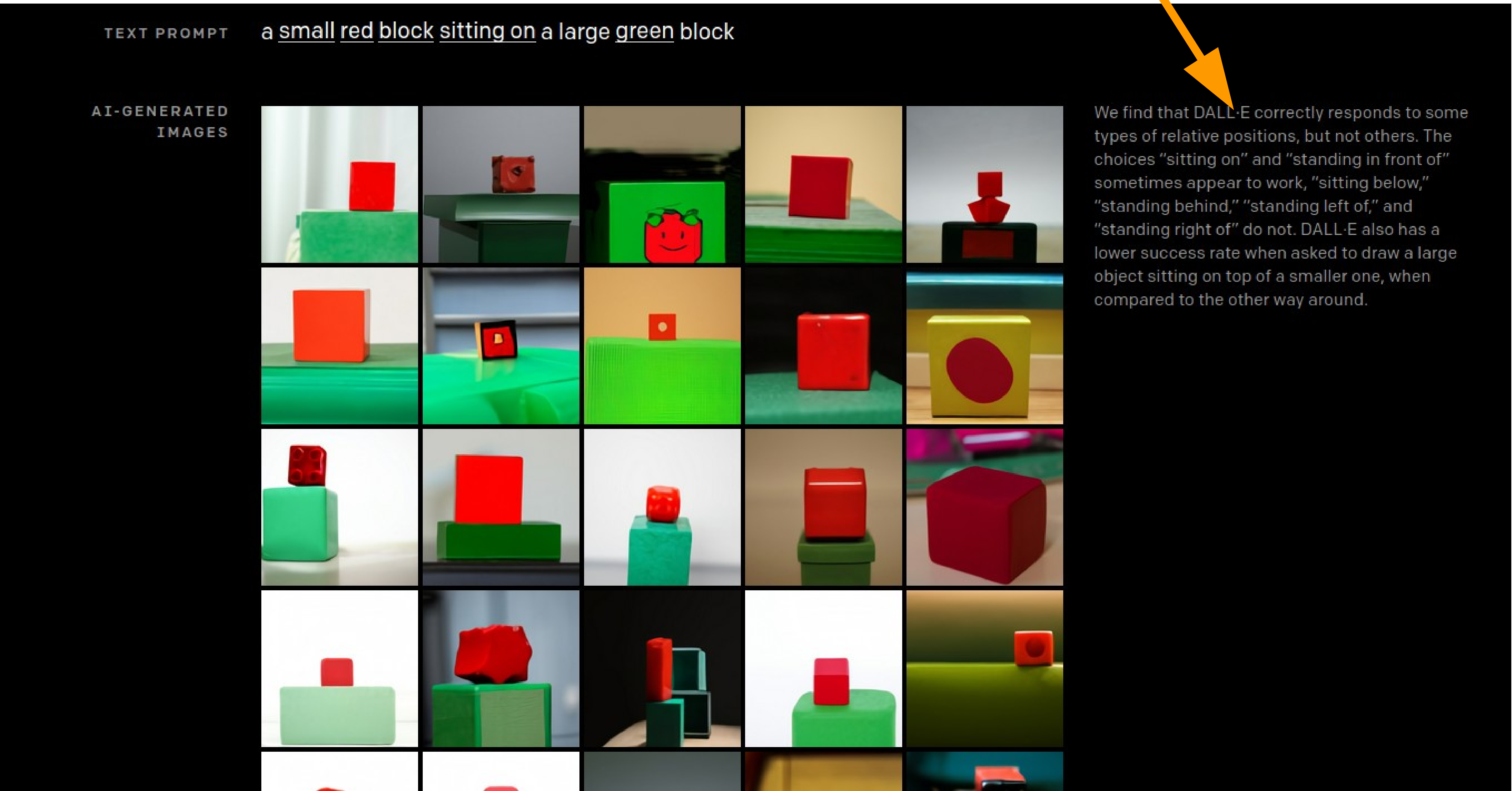
an armchair in the shape of an avocado [...]

AI-GENERATED IMAGES

Edit prompt or view more images ↓

OpenAI DALL·E

"We find that DALL·E correctly responds to some types of relative positions, but not others. The choices "sitting on" and "standing in front of" sometimes appear to work, "sitting below," "standing behind," "standing left of," and "standing right of" do not. DALL·E also has a lower success rate when asked to draw a large object sitting on top of a smaller one, when compared to the other way around."



A continuous progress?

- By using a mixture of ML techniques, several human or super-human performances are achieved every year in specific tasks (*mostly by corporation-driven research*).
- More and more sensitive applications are been researched and deployed in the wild.
- All the problems of **generalization**, **explainability**, **transparency**, **responsibility**, fairness… are still there.

The present

- New research trends are emerging to face these issues, trying a variety of different approaches.

The present

- New research trends are emerging to face these issues, trying a variety of different approaches.

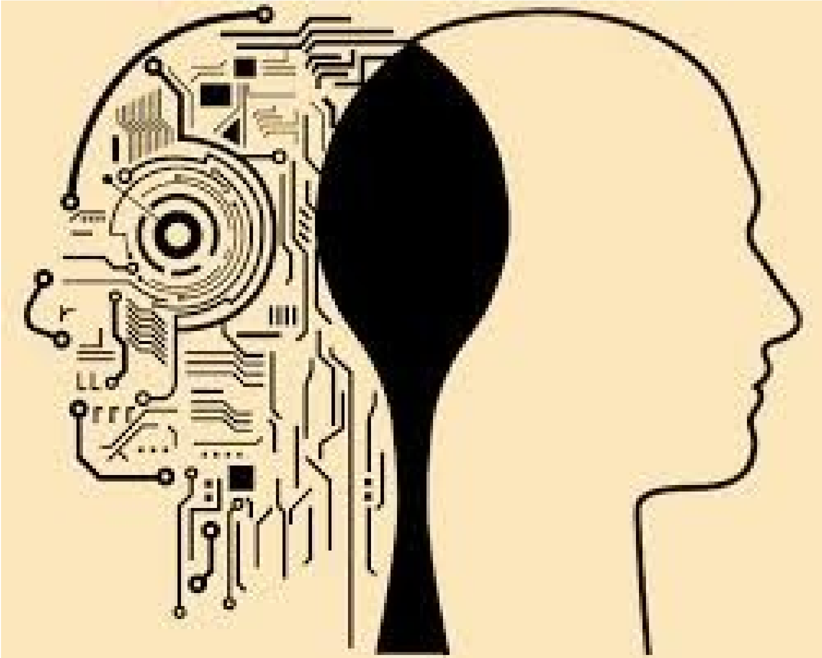
- Still unclear which one will achieve the intent.

Conclusions

artificial general intelligence

No AGI in view

- I believe (with many others) that crucial pieces are still missing to embed *general intelligence* into a single artificial device.



- These pieces might be simple or not, it’s the ML method that is not satisfactory to design them.

Rise of “artificially dumber” systems

- However, already today, the introduction of ubiquitous *cyber-physical connections* in all human activities raises serious concerns at societal and at cognitive level.
  - high risks to be entangled in **artificially dumber** systems.

Rise of "artificially dumber" systems

- However, already today, the introduction of ubiquitous *cyber-physical connections* in all human activities raises serious concerns at societal and at cognitive level.
  - high risks to be entangled in **artificially dumber** systems.

- The potential impact is too critical to be belittled for the belief in technologically-driven '*magnificent and progressive fate*'.

“Outperforming” humans

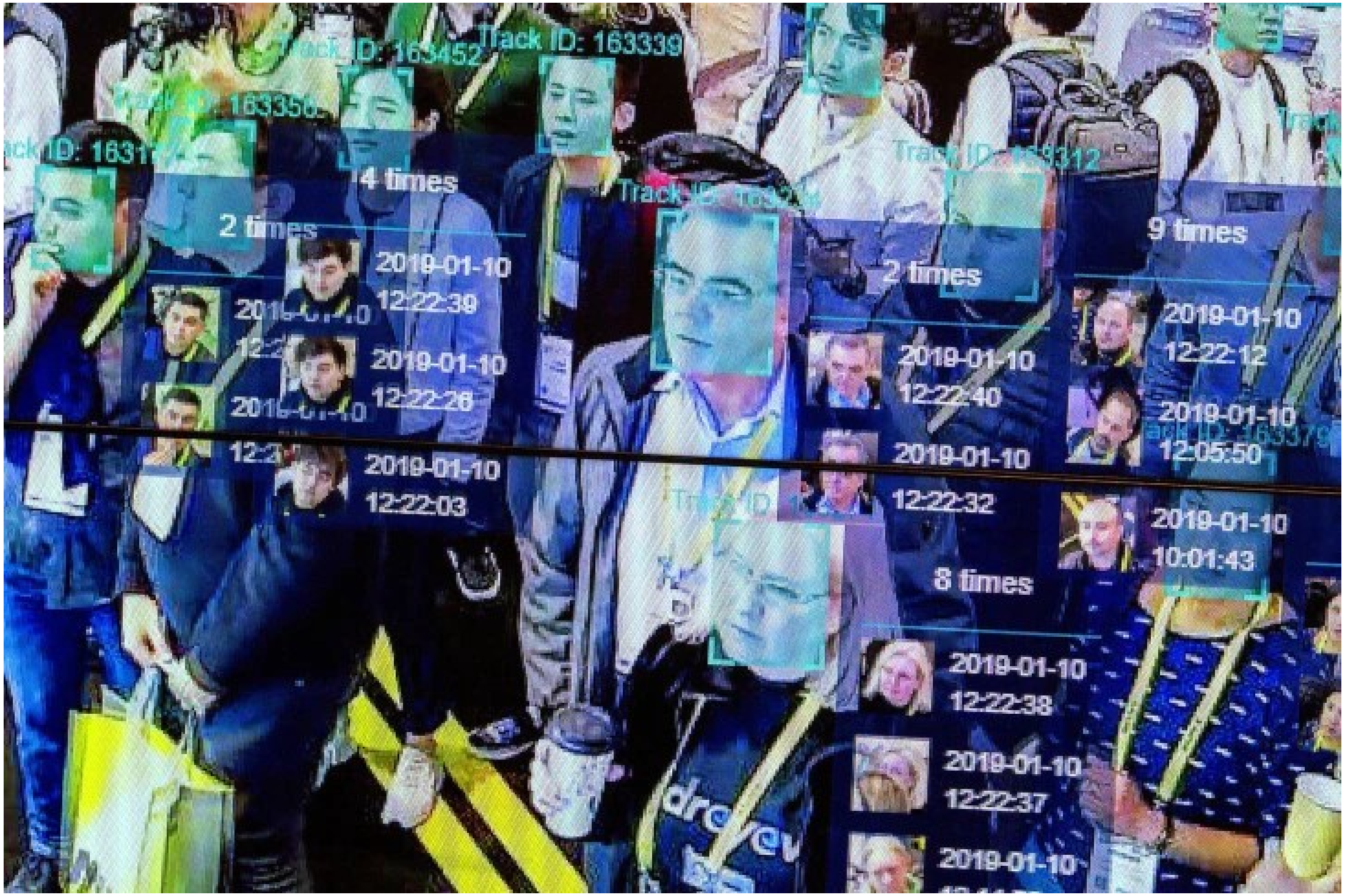
Touching numerals from 1 to 9

Limited-hold memory task (Human, 5 numerals)

Masking task

https://langint.pri.kyoto-u.ac.jp/ai/en/publication/SanaInoue/Inoue2007.html

“Outperforming” humans



training + sufficient memory… can we call it intelligence?

If we are pursuing **rationality** (*rational systems*, *rational institutions*, etc.), it is rather implausible that this will be obtained only by *empirical means*.

Introduction to AI: Understanding the Technology

Winter Academy on Artificial Intelligence and International Law

Asser Institute – 11 February 2021

Giovanni Sileno
g.sileno@uva.nl