Data and Information

Model of Natural communication, Elias Zafiris

Marinčić, N. (2017). Towards Communication in CAAD: Spectral characterization and modelling with conjugate symbolic domains(Doctoral dissertation, ETH Zurich).

Model
information
→

encode
Machine Intelligence

world ↻

decode
Intelligent Humans

Map
data
→
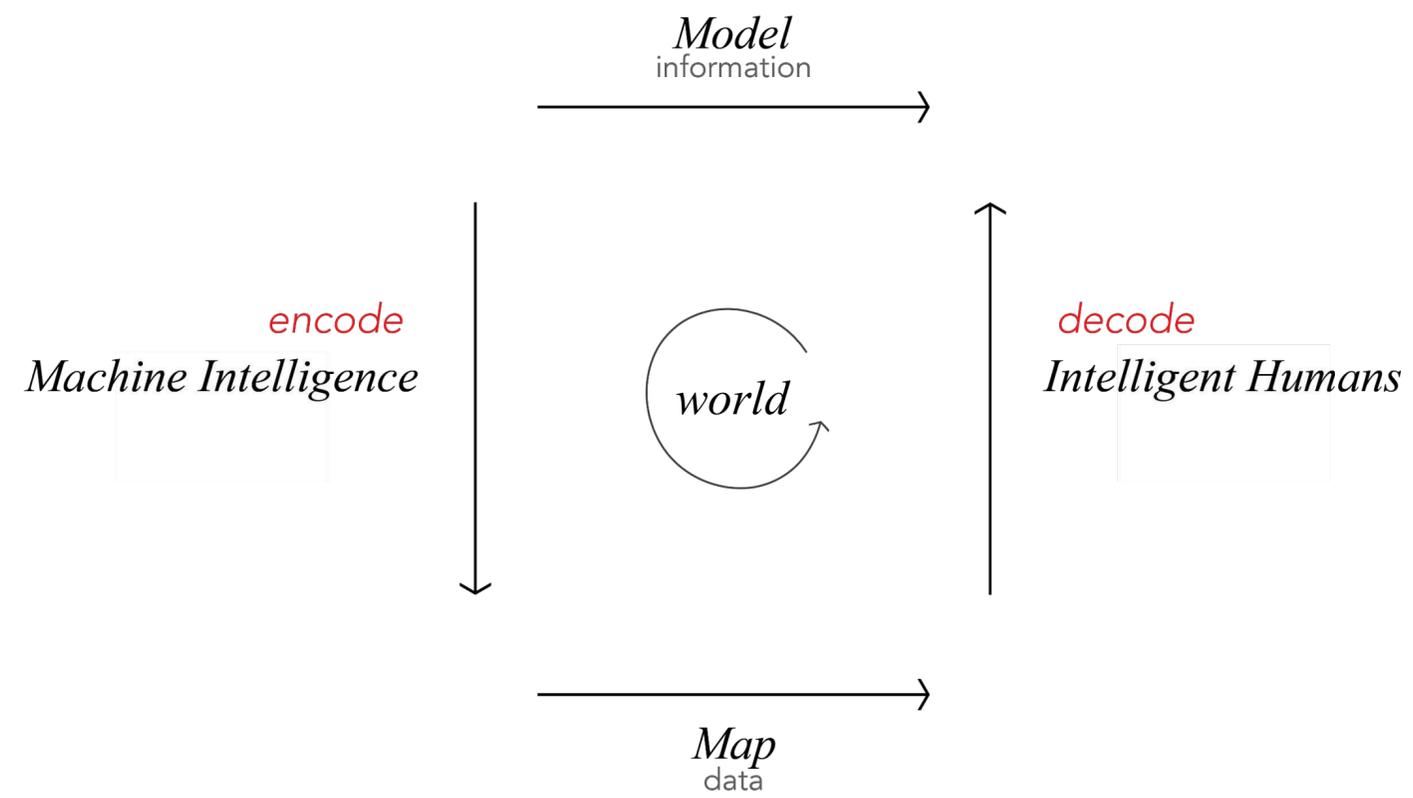
**Harmonics /** *Geometry*

Map / *Model*

Data / *Information*
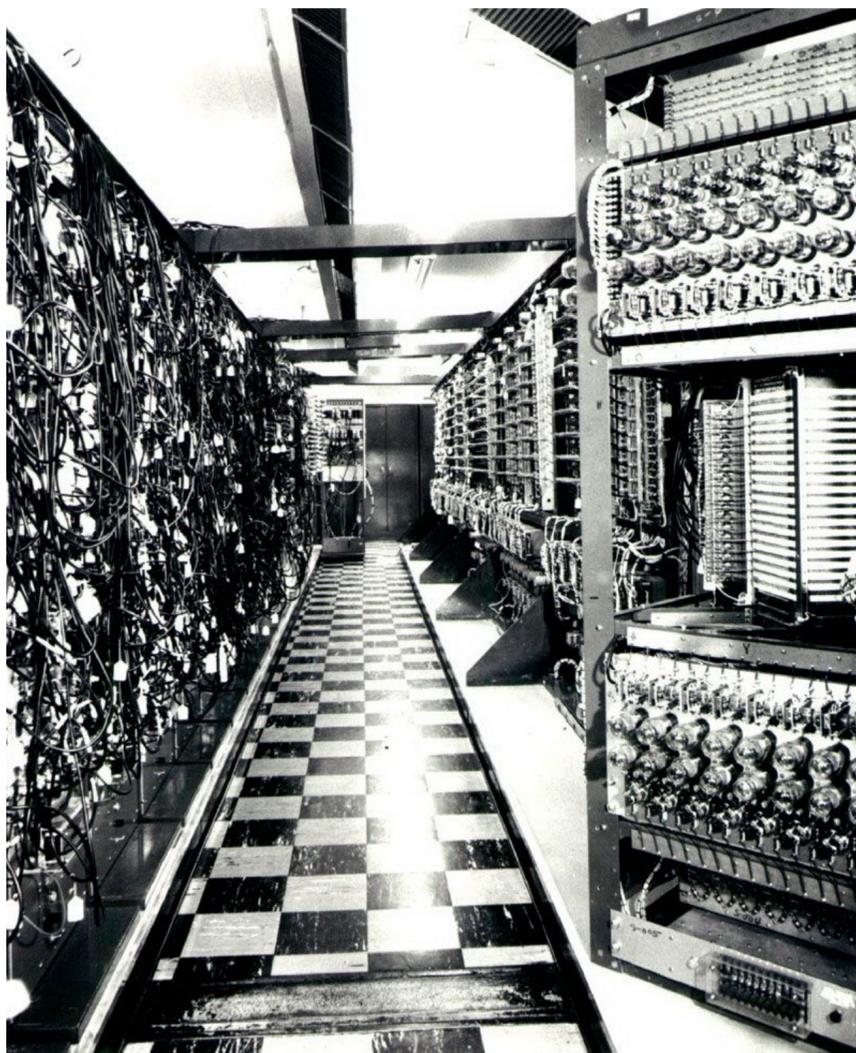
one computer many people > **one person one computer** > **one person many devices**
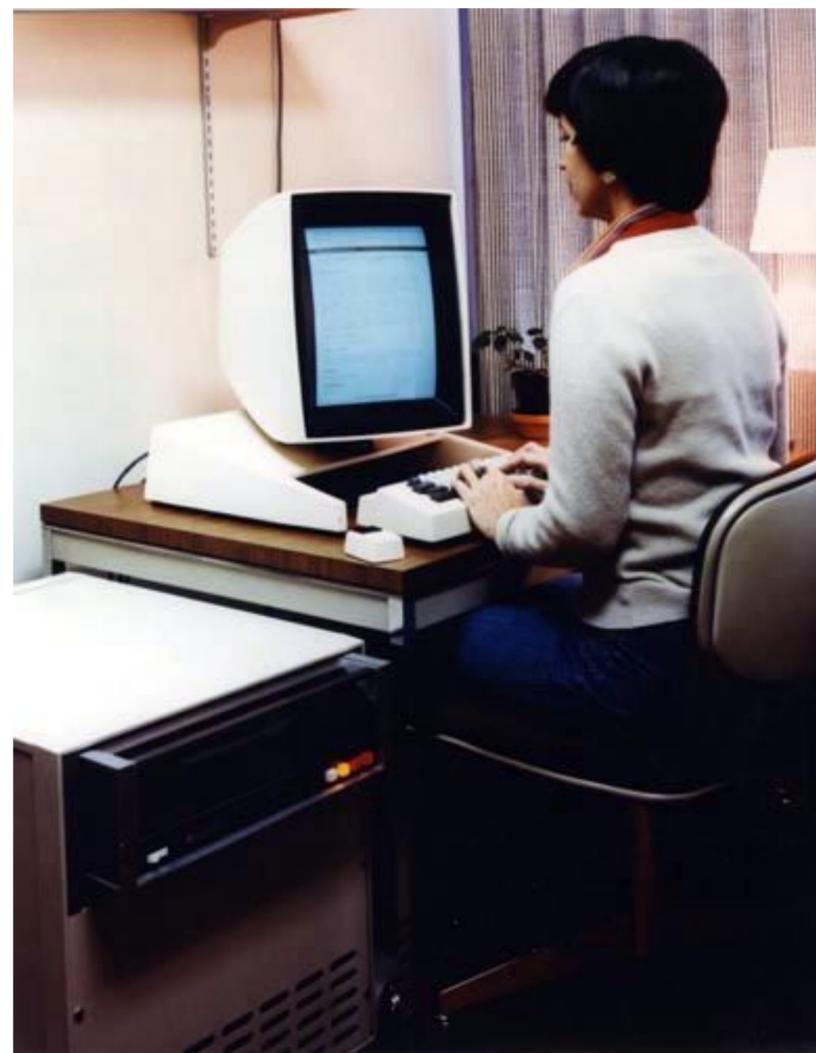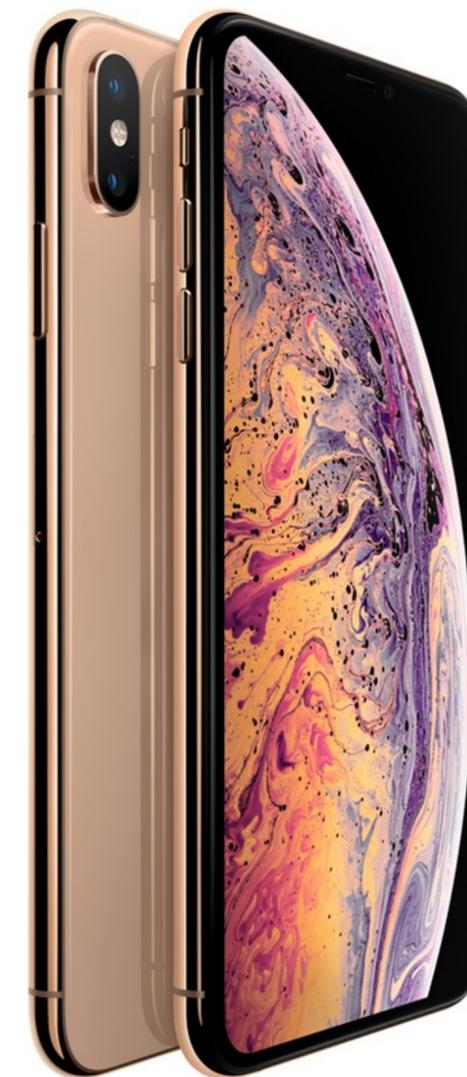
Alvarez D., Saldana K., (2017), City of indexes, Elective course 2017, ETH ITA CAAD Praxis 2018

As people move through the city, **they produce, collect, share, store, leave traces of their daily activities**. In cities, a tremendous amount of data is being produced either intentionally, via blogs, social media, discussion forums, news, messages, or unintentionally via online and offline interactions.

Vahid Moosavi
Modeling urban traffic dynamics in coexistence with urban data streams

Rouvroy, A. (2016). "Of Data and Men" Fundamental Rights and Freedoms in a World of Big Data. Bureau of the consultative committee of the convention for the protection of individuals with regard to automatic processing of personal data. Estrasburgo.

What is *big*-data?

Big-data refers to the digital byproducts of human activity.

volume

the digital universe today is made up of more than 1,200 billion billion bytes, 90% of which would appear to have been produced in the last two years
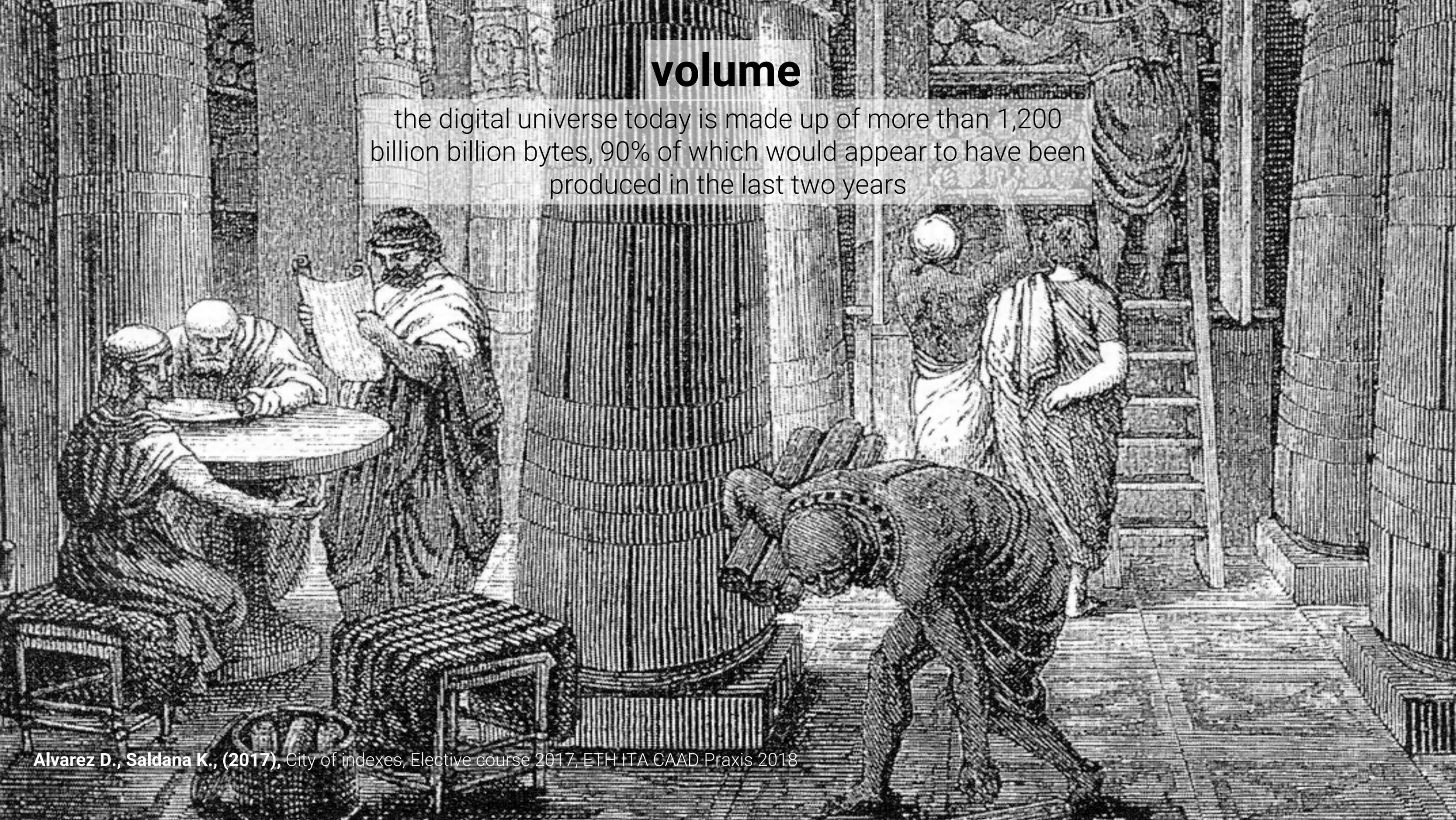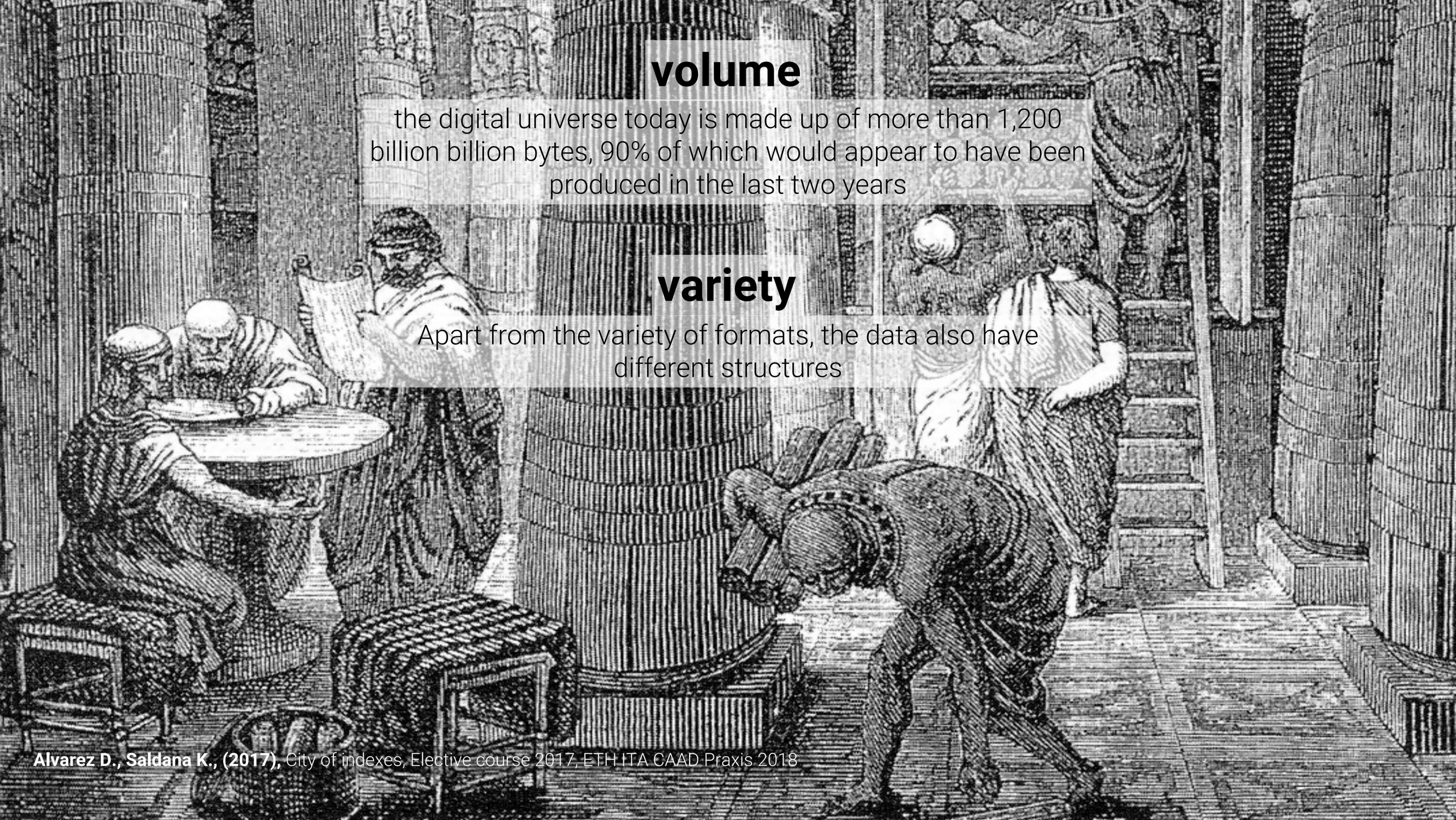
volume

the digital universe today is made up of more than 1,200 billion billion bytes, 90% of which would appear to have been produced in the last two years

variety

Apart from the variety of formats, the data also have different structures

volume
the digital universe today is made up of more than 1,200 billion billion bytes, 90% of which would appear to have been produced in the last two years

variety
Apart from the variety of formats, the data also have different structures

velocity
Accumulated in real time, all the time

Alvarez D., Saldana K., (2017), City of indexes, Elective course 2017, ETH ITA CAAD Praxis 2018
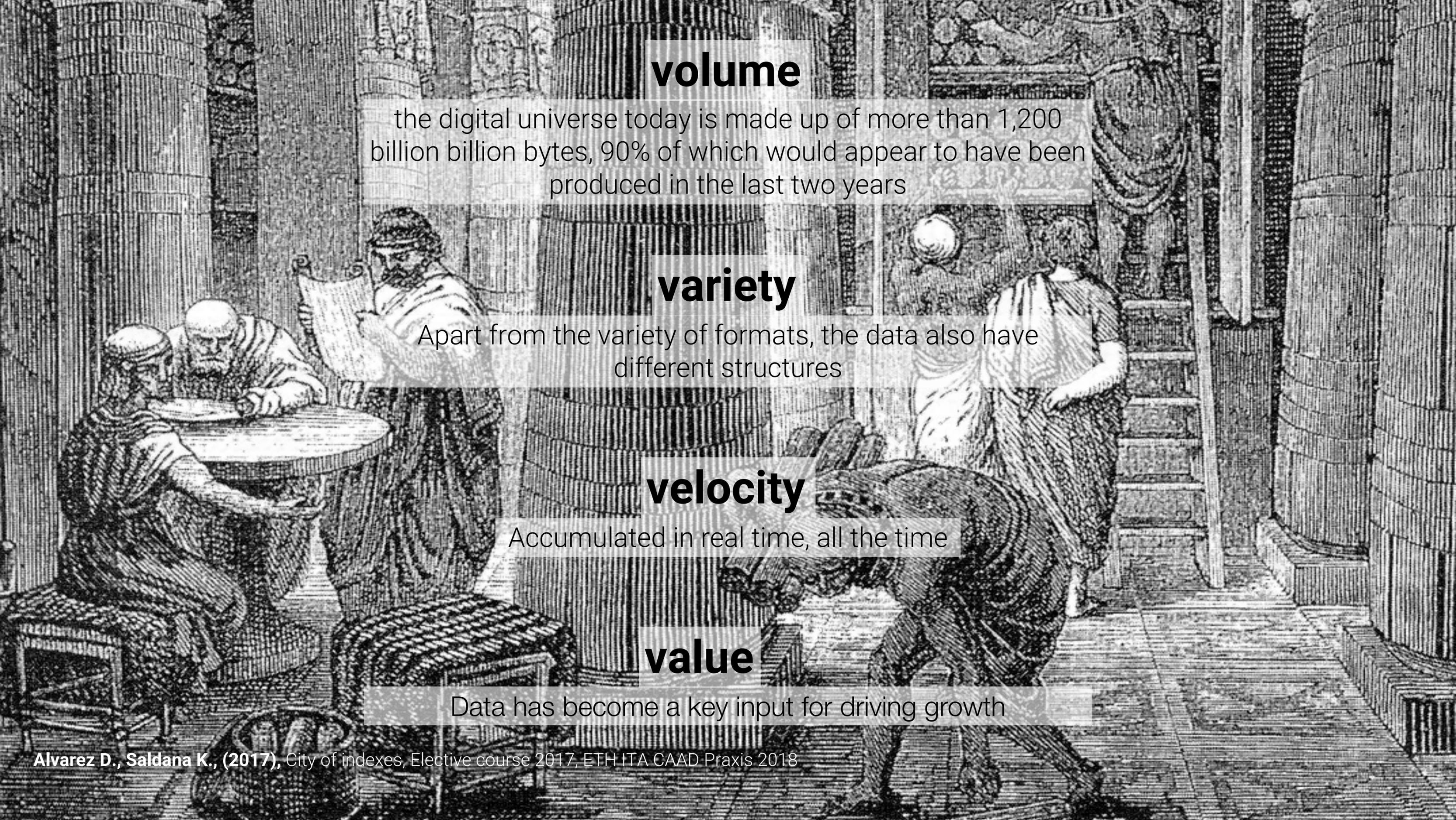
volume
the digital universe today is made up of more than 1,200 billion billion bytes, 90% of which would appear to have been produced in the last two years

variety
Apart from the variety of formats, the data also have different structures

velocity
Accumulated in real time, all the time

value
Data has become a key input for driving growth

Alvarez D., Saldana K., (2017), City of indexes, Elective course 2017, ETH ITA CAAD Praxis 2018

we are being flooded with data

Alvarez D., Saldana K., (2017), City of indexes, Elective course 2017, ETH ITA CAAD Praxis 2018

What is *information?*

late 14c., informacion, "**act of informing**, communication of news," from Old French informacion, enformacion "**advice, instruction**," from Latin informationem (nominative informatio) "**outline, concept, idea**," noun of action from past participle stem of informare **"to train, instruct, educate; shape, give form to"**

"I do not know any living being, cell, tissue, organ, individual, or perhaps even species, of which we cannot say that **they store information, that they treat (or process) information, that they emit it and they receive information.** [...] I know of no object in the world, atom, crystal, mountain, planet, star, galaxy, of which one could not say again **that it stores information, it treats (or processes) information, it emits and it receives information**. So there's this quadruple characteristic in common between all the objects of the world, living or inert." **These four functions seem then to be universal.**"

Serres

Model
information

encode
Machine Intelligence

world

decode
Intelligent Humans

Map
data
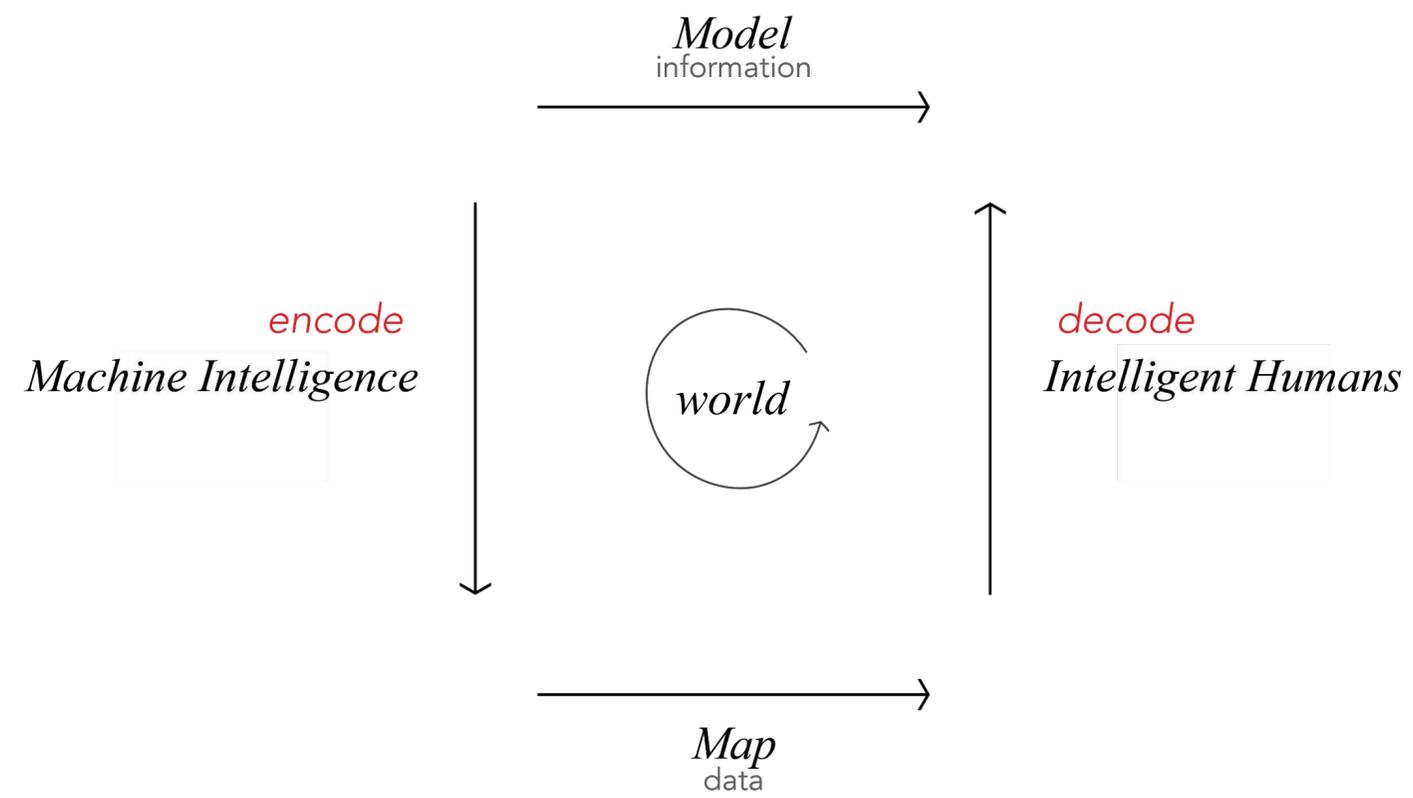
Intelligent Human

"It is almost impossible that our **judgment could be so clear or so solid**, as it might have been, had **we had the entire use of our reason from the time of our birth**, and been always guided by it alone."

Decartes

Intelligent Human

Embracing obstacles

Machine Intelligence

machine intelligence algorithms are seen as a **problem-solving tool**, attuned to a specific set of problems. These algorithms are fast, computationally powerful and grow exponentially in capacity to process, store and communication

What is Intelligence?

Annoyed by fruitless disputes over word meanings, Alan Turing (1950)
proposed a criterion that satisfies nearly everyone – The Turing Test
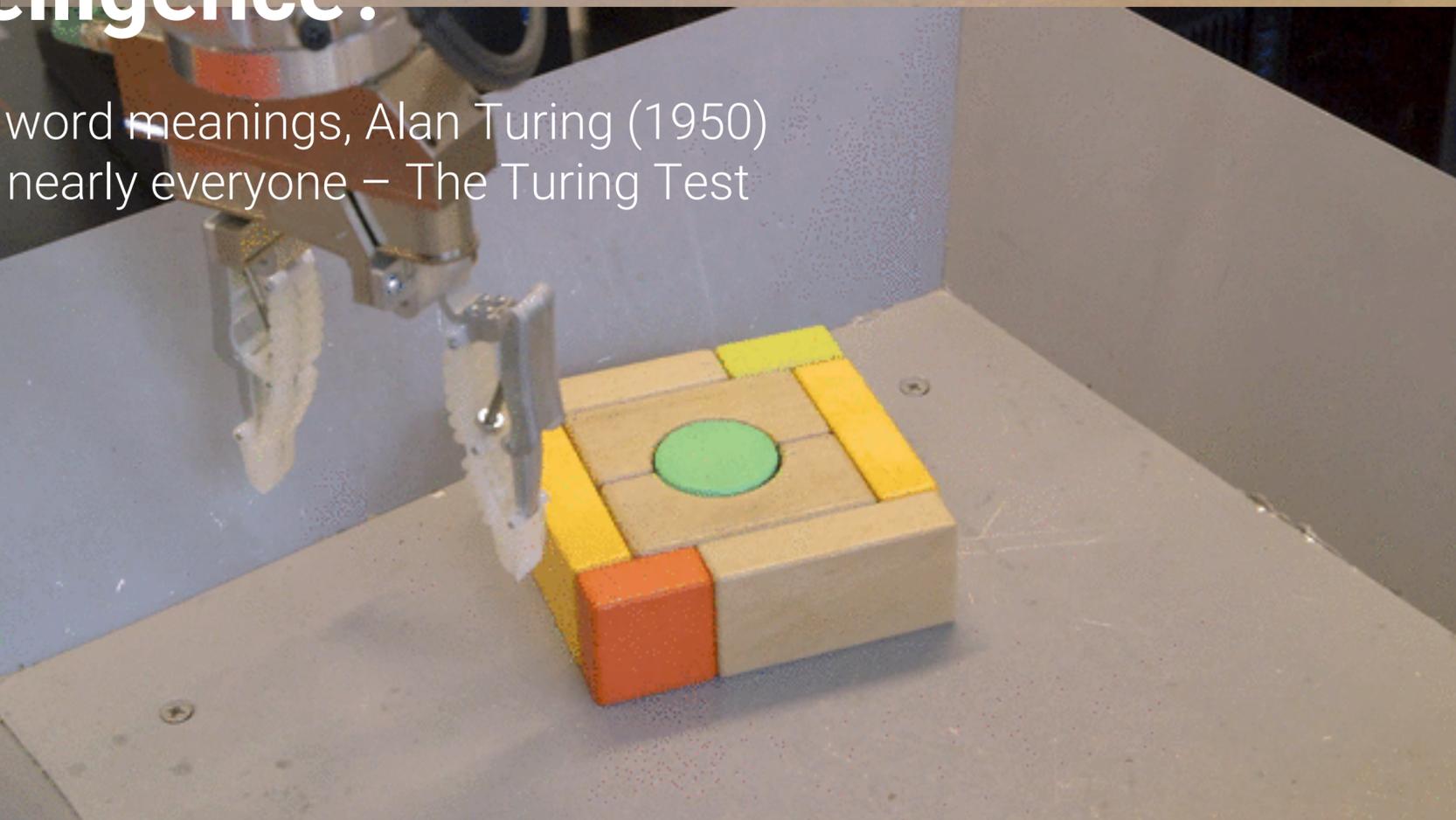
"It is played with three people, **a man (A), a woman (B), and an interrogator (C)** who may be of either sex. The interrogator stays in a room apart front the other two. **The object** of the game for the interrogator is **to determine which of the other two is the man and which is the woman**"

"What will happen when **a machine takes the part of A** in this game?"

**Turing, Alan (1950).** "Computing Machinery and Intelligence," Mind, 59, 434-460 (reprinted in Feigenbaum and Feldman, 1963, and in Anderson, 1964).
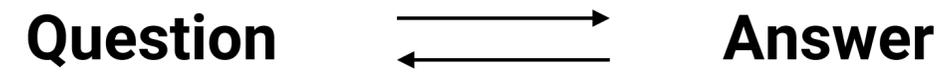
Question ⇄ **Answer**

"The question and answer method seems to be suitable for introducing almost any one of the fields of human endeavor that we wish to include"

Turing, Alan (1950). "Computing Machinery and Intelligence," Mind, 59, 434-460 (reprinted in Feigenbaum and Feldman, 1963, and in Anderson, 1964).

Question ⇄ **Answer**

How well can computer

- play chess

- classify pictures

- drive cars

......

solve particular problem

Approaches of Machine Intelligence

Instead of being an interrogator directly, we compare the outputs with the human-level results

Question $\longrightarrow$ **Answer**

$$\boldsymbol{a} = f(\boldsymbol{q})$$

can we find a mathematical function that always give
us satisfying answers for different questions?
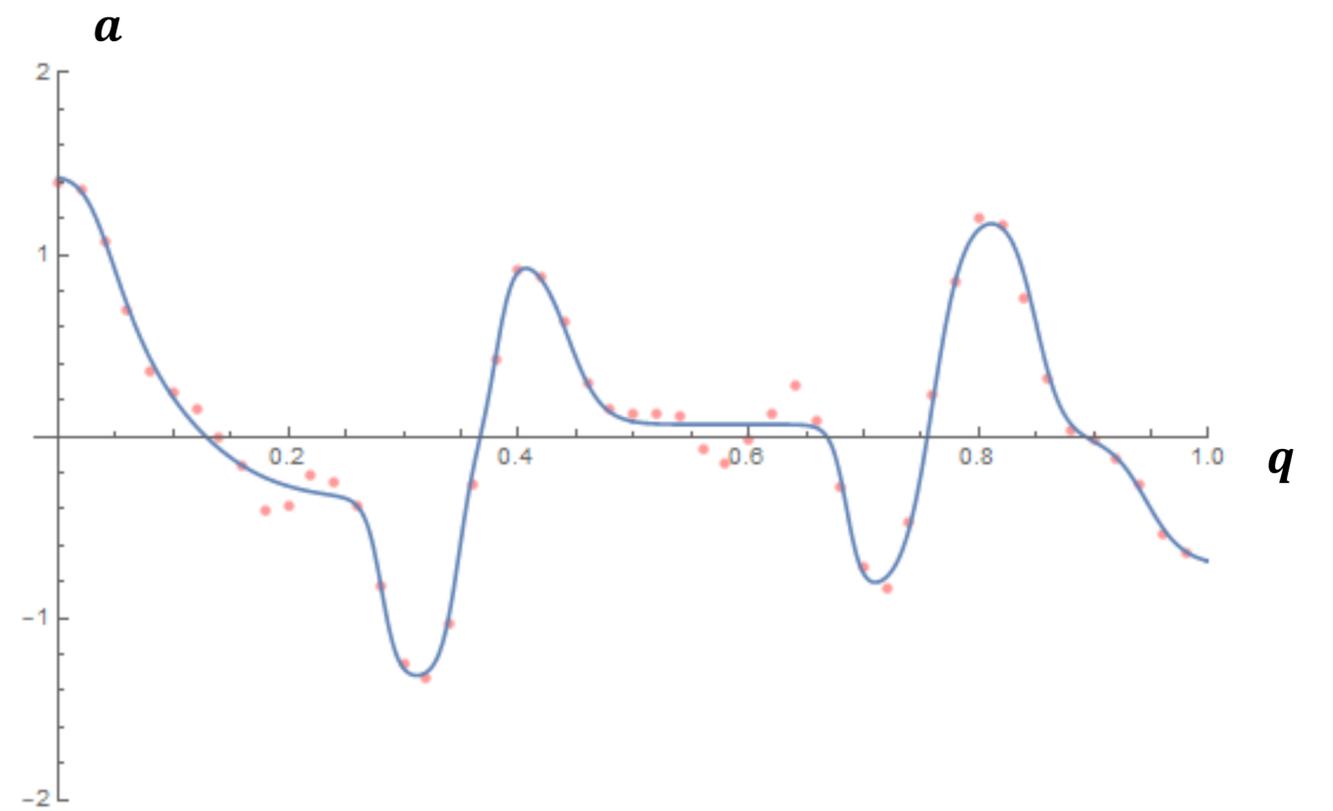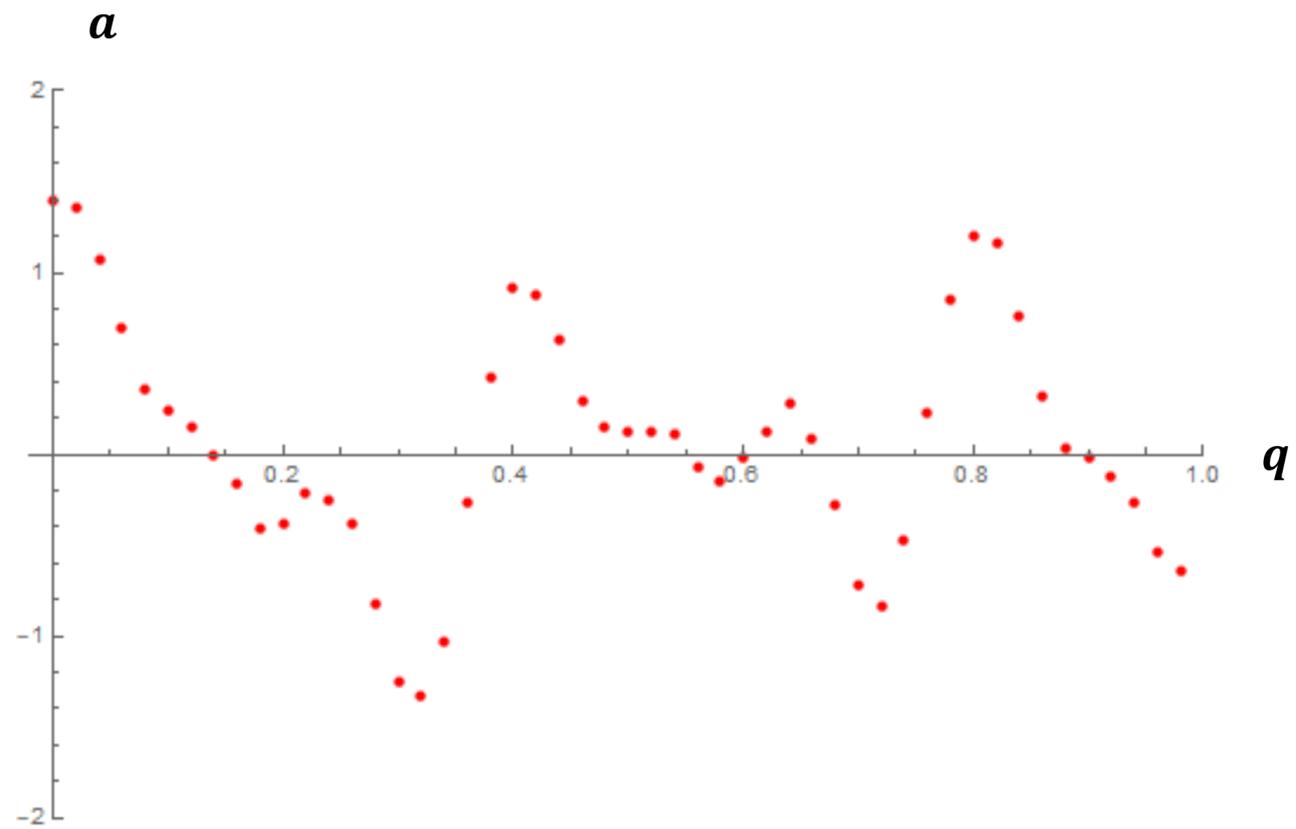
$$a = f(q)$$

Towards Symbolic Manipulation

rule system
decision tree

……

↓

Logic Based

Towards Neuro Simulation

artificial neural network

……

Mathematical Abstraction   Explainable

Alpha GO   neuro science

……   ……

↓

Probabilistic Based

# Towards Symbolic Manipulation

rule system
decision tree

……

↓

Logic Based

Am I Hungry?

Have I 50 Fr. ?   Go to Sleep

Go to Restaurant   Buy a Hamburger
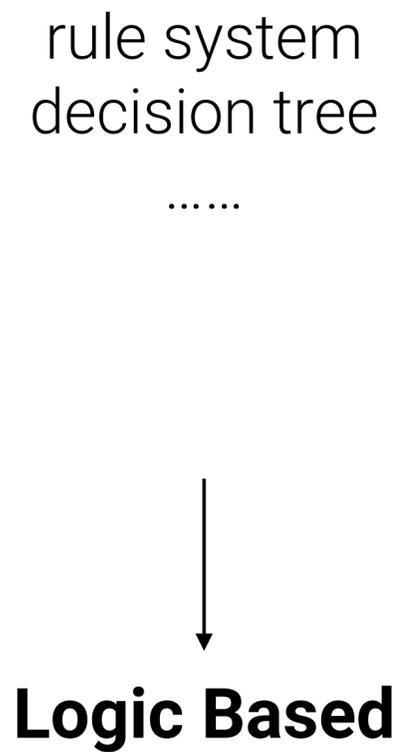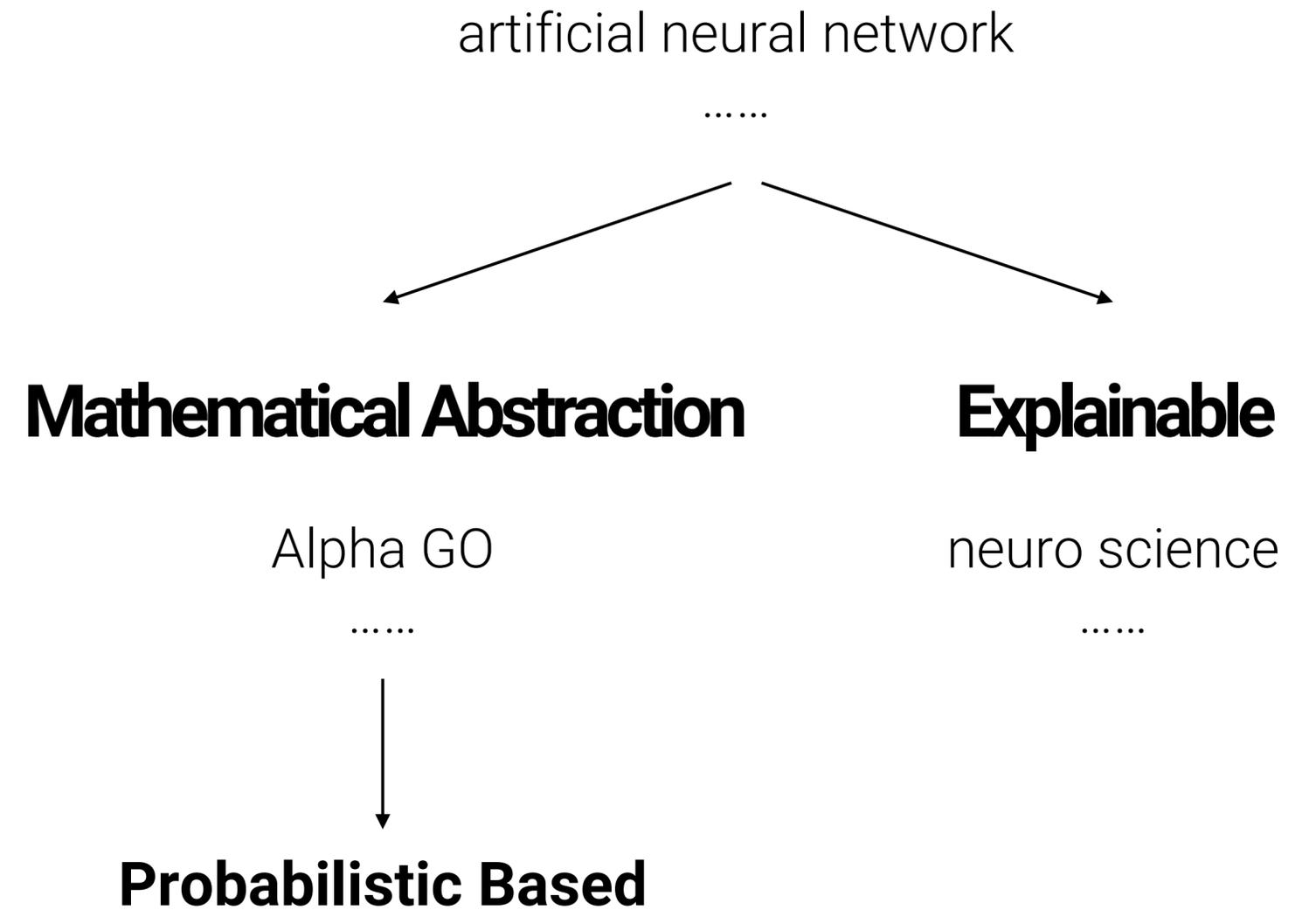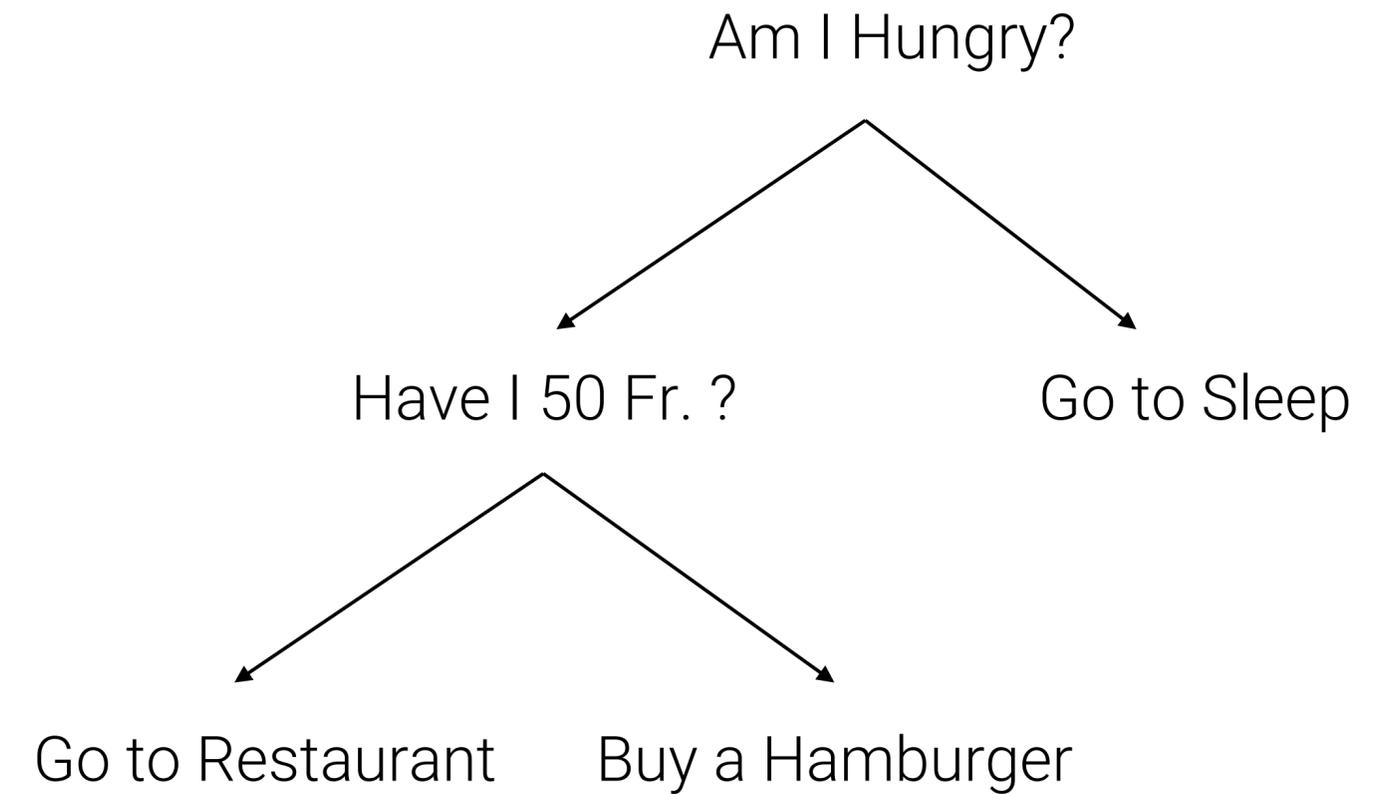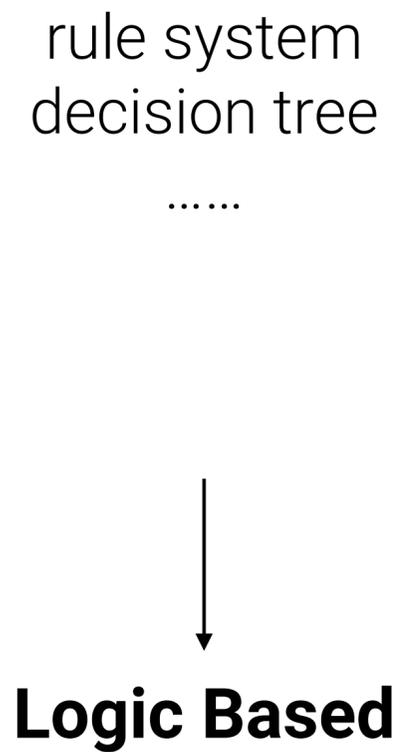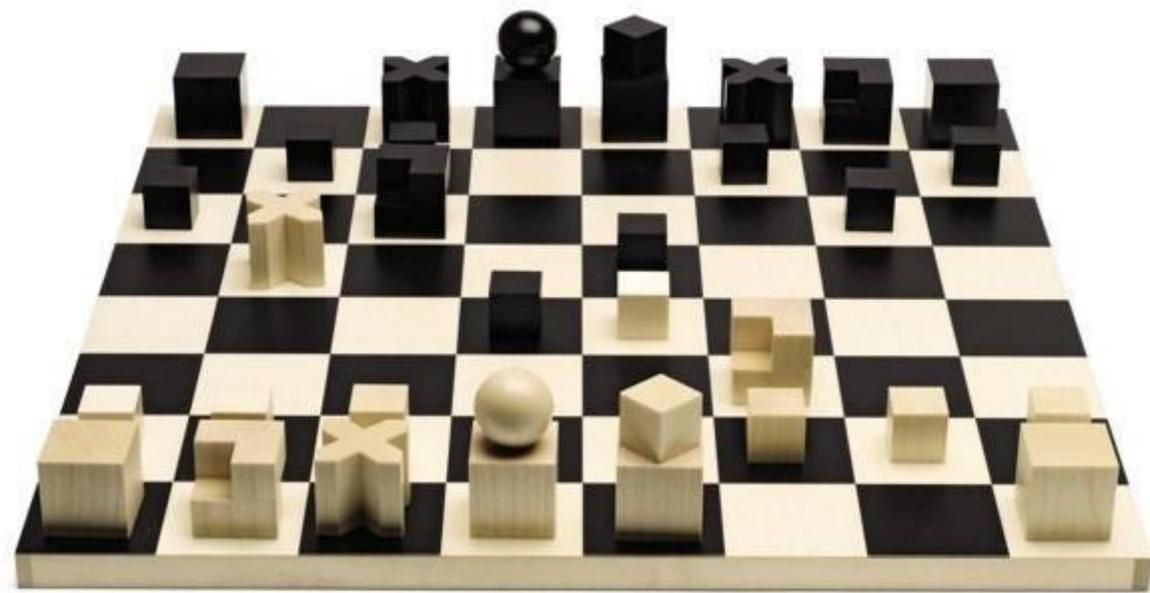
Deep Blue

Brutal search for all possible moves

10^{50}

https://en.wikipedia.org/wiki/Deep_Blue_versus_Garry_Kasparov

Game AI

Predefined State-Action rules

"Good Old Fashioned Artificial Intelligence, as a branch of cognitive science, rest on a particular theory of intelligence and thought – essentially Hobbes's idea that ratiocination is computation"

Haugeland, J. (1989). Artificial intelligence: The very idea. MIT press.

Explosion of the search space

10^{50}

2×10^{170}

Allis, Louis Victor (1994), Searching for solutions in Games and Artificial Intelligence, Maastricht: Proefschrift Rijksuniversiteit Limburg, ISBN 90-9007488-0

airplane
automobile
bird
cat
deer
dog
frog
horse
ship
truck

http://www.cs.toronto.edu/~kriz/cifar.html
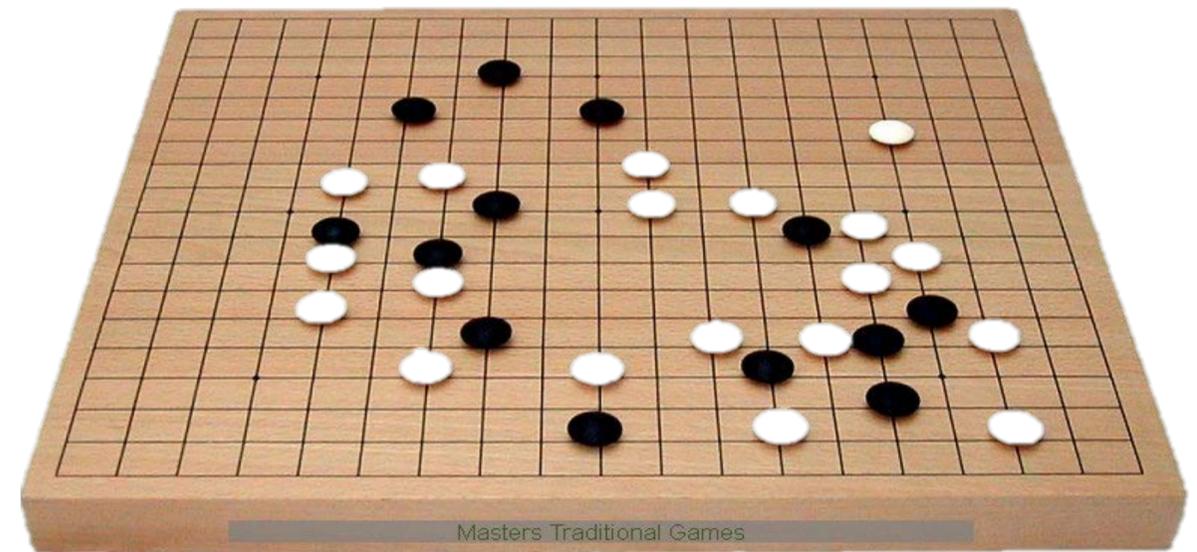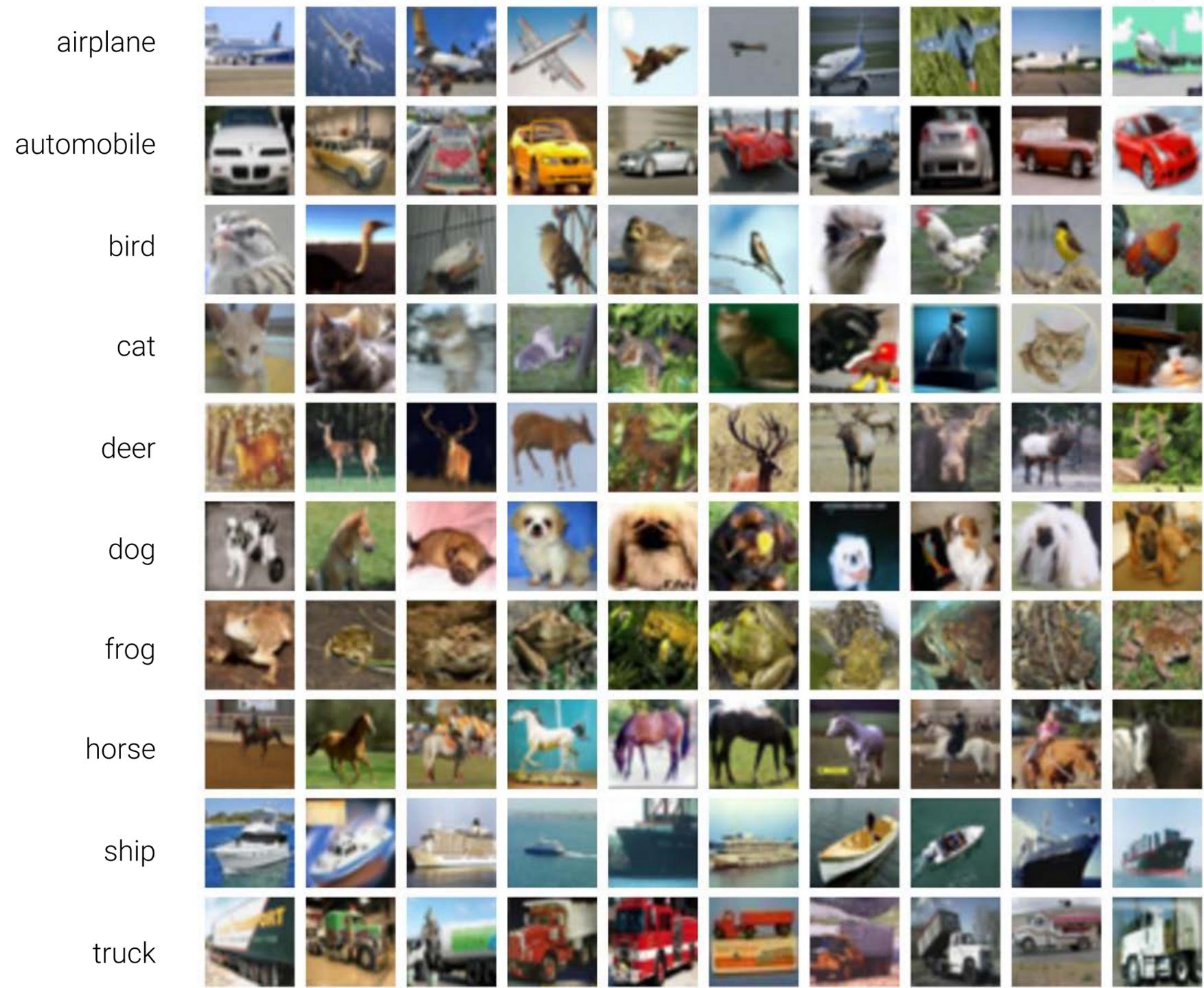
viewpoint variation

scale variation

deformation

occlusion

illumination condition

background clutter

intra-class variation

http://cs231n.github.io/assets/challenges.jpeg

Explicit Rules

Inexplicit Correlations

Cat: Yes
Dog: No

Towards Neuro Simulation

Probabilistic Based

Finding correlations within data using computational models

Usually data-driven

artificial neural network

……

Mathematical Abstraction

Alpha GO

……

Explainable

neuro science

……

Probabilistic Based

Perceptron?

A bio-inspired mathematic model

$$= \tanh(\sum x_n w_n + b)$$

x_1

x_2

…

w_1

w_2 b

…

inputs   perceptron   output
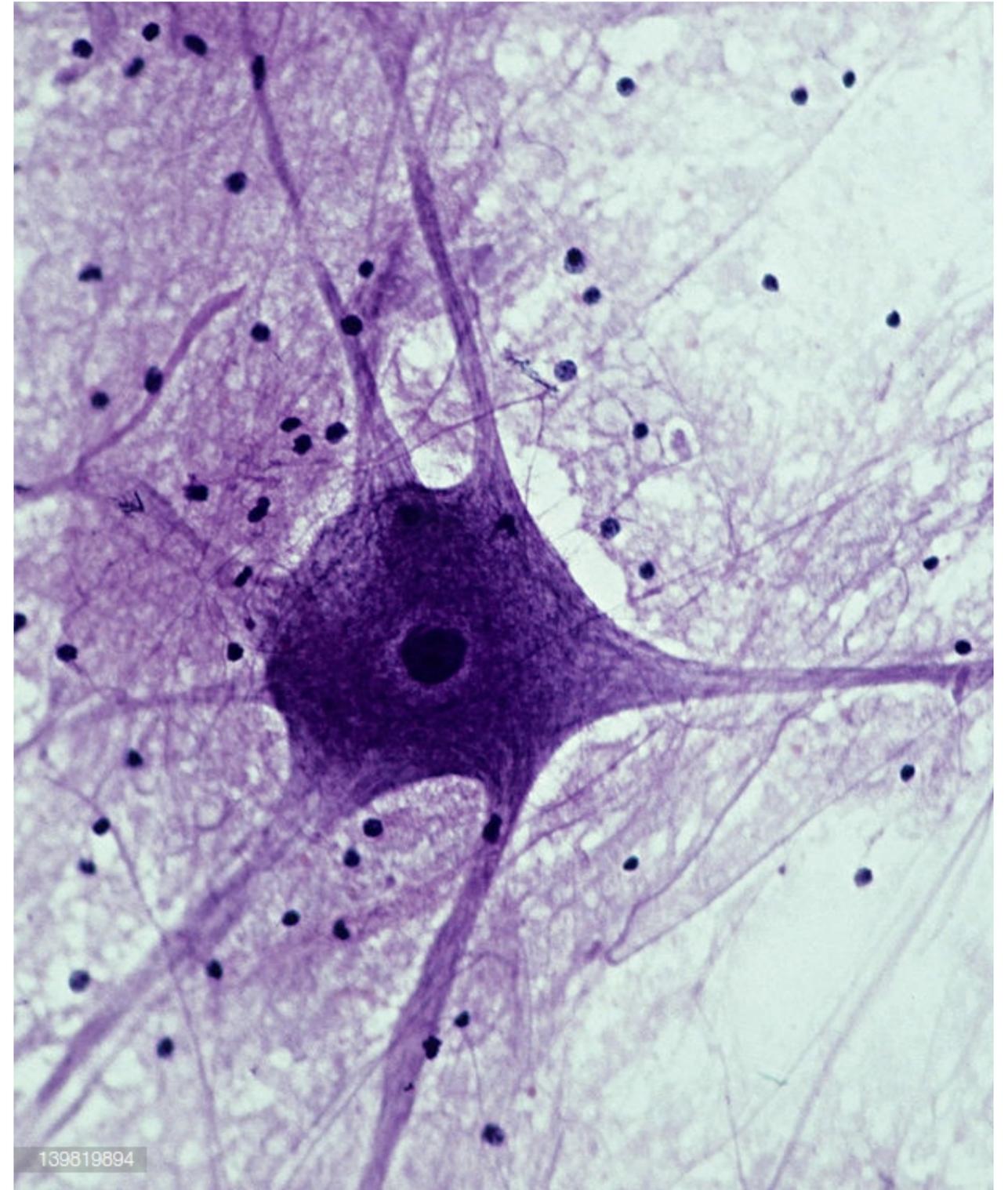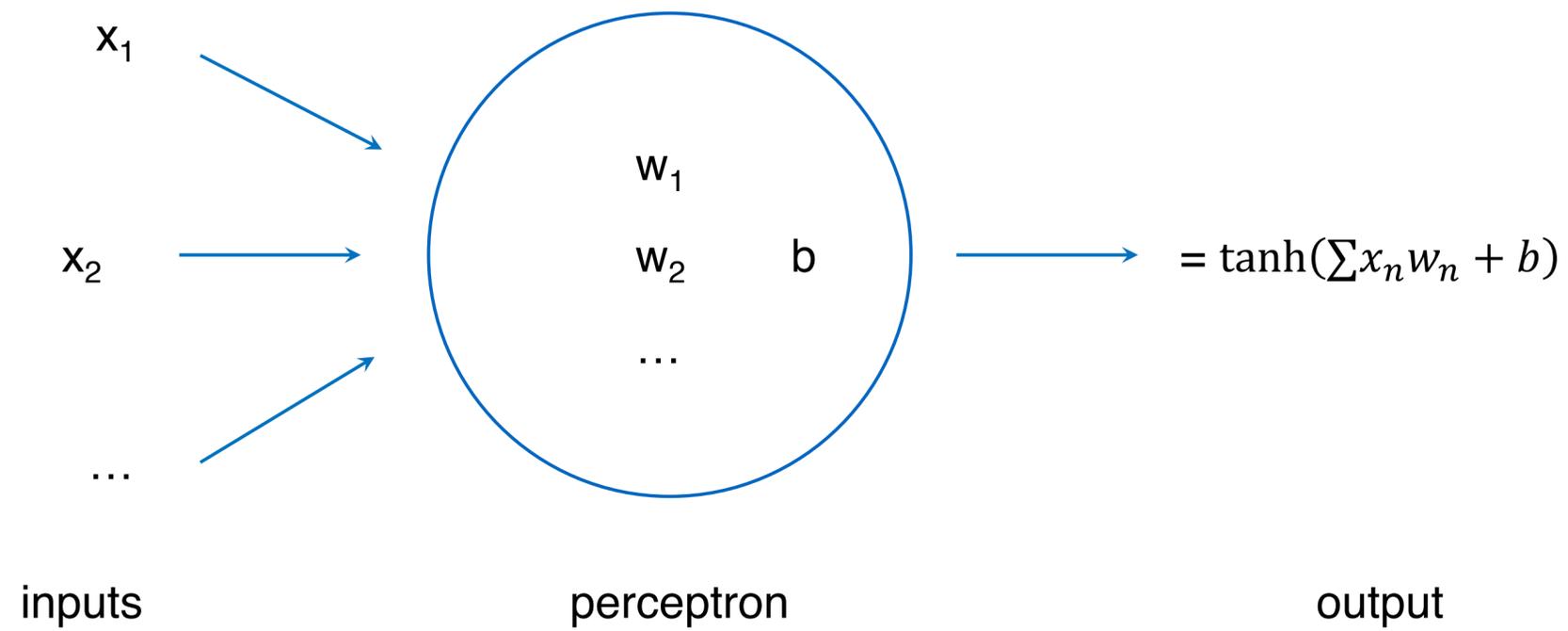
Neural network
Stack of connected Perceptrons

x_1

x_2

...

w_1

w_2 b

...

$= \tanh(\sum x_n w_n + b)$

Hidden

Output

Learning?

x_1

x_2

...

w_1

w_2 b

...

$= \tanh(\sum x_n w_n + b)$
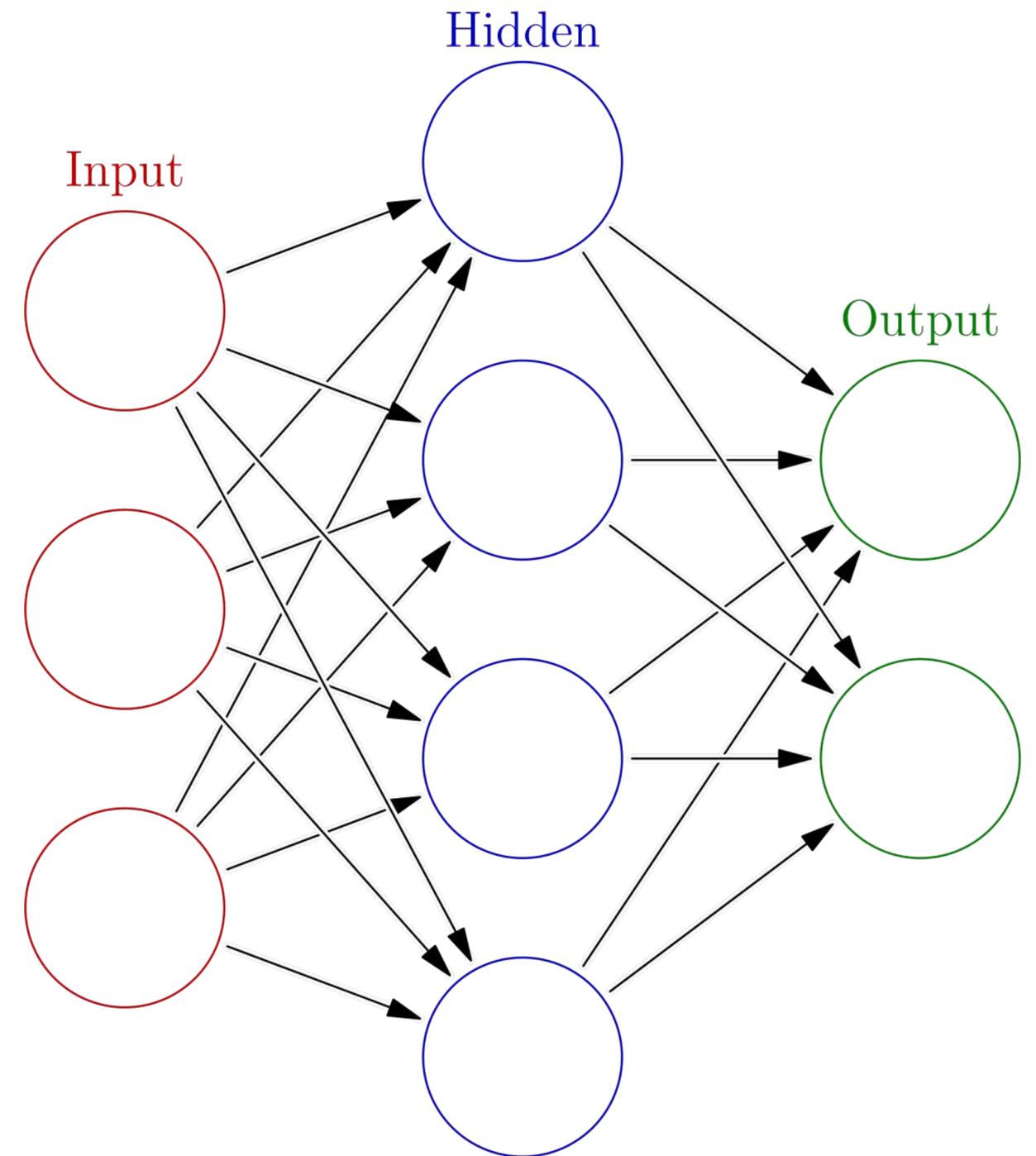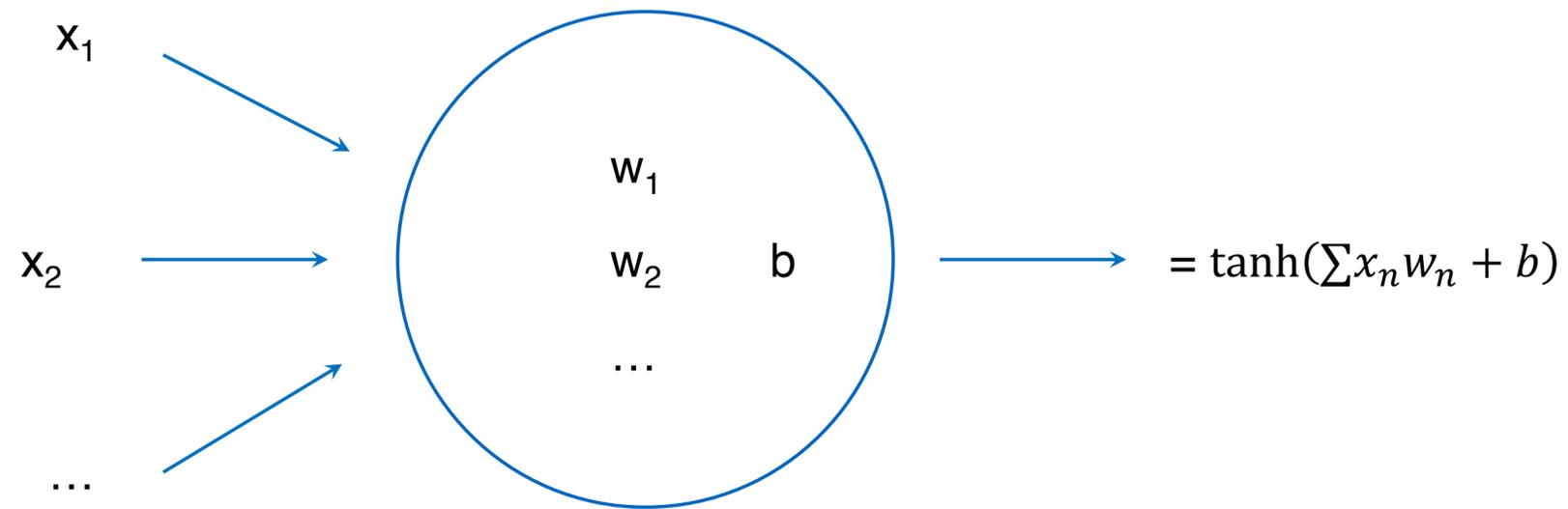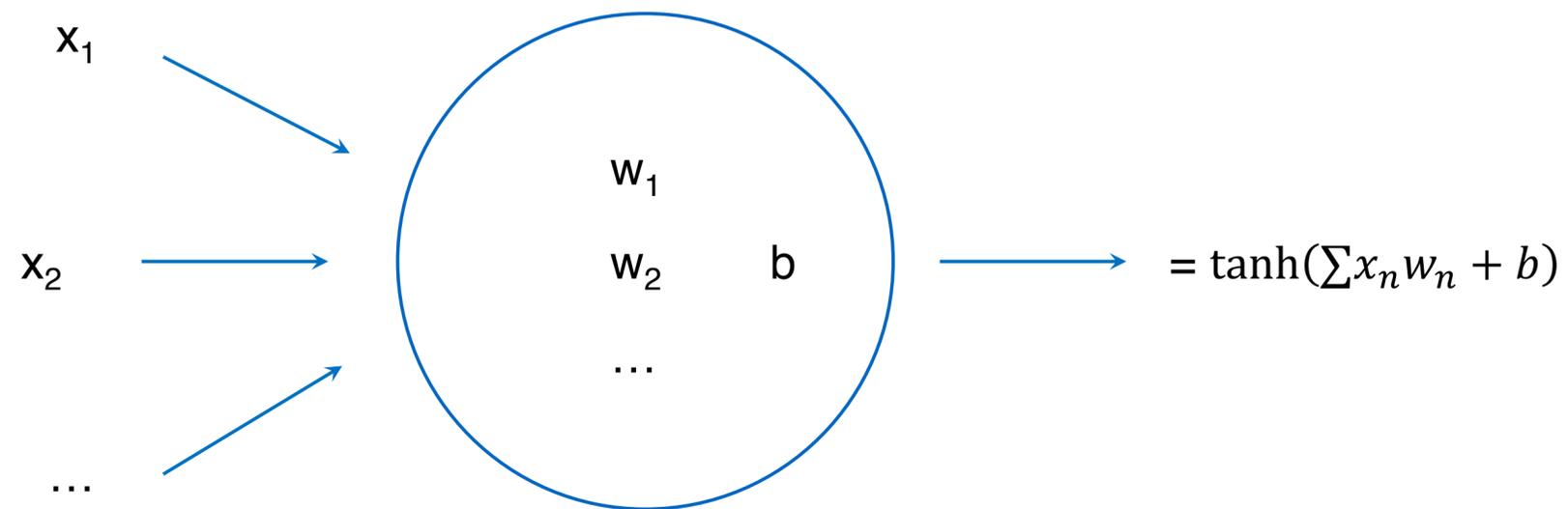
The goal of the learning is to find proper values of w and b so that we get expected outputs from inputs

The bottle neck of the neural network in early stage was **the lack of proper learning method**. This issue was addressed by the algorithm called **Backpropagation** in 1975, by Paul J. Werbos

Backpropagation

partial derivative

x_1

w_1

x_2

w_2 b   actual output

...

error

...
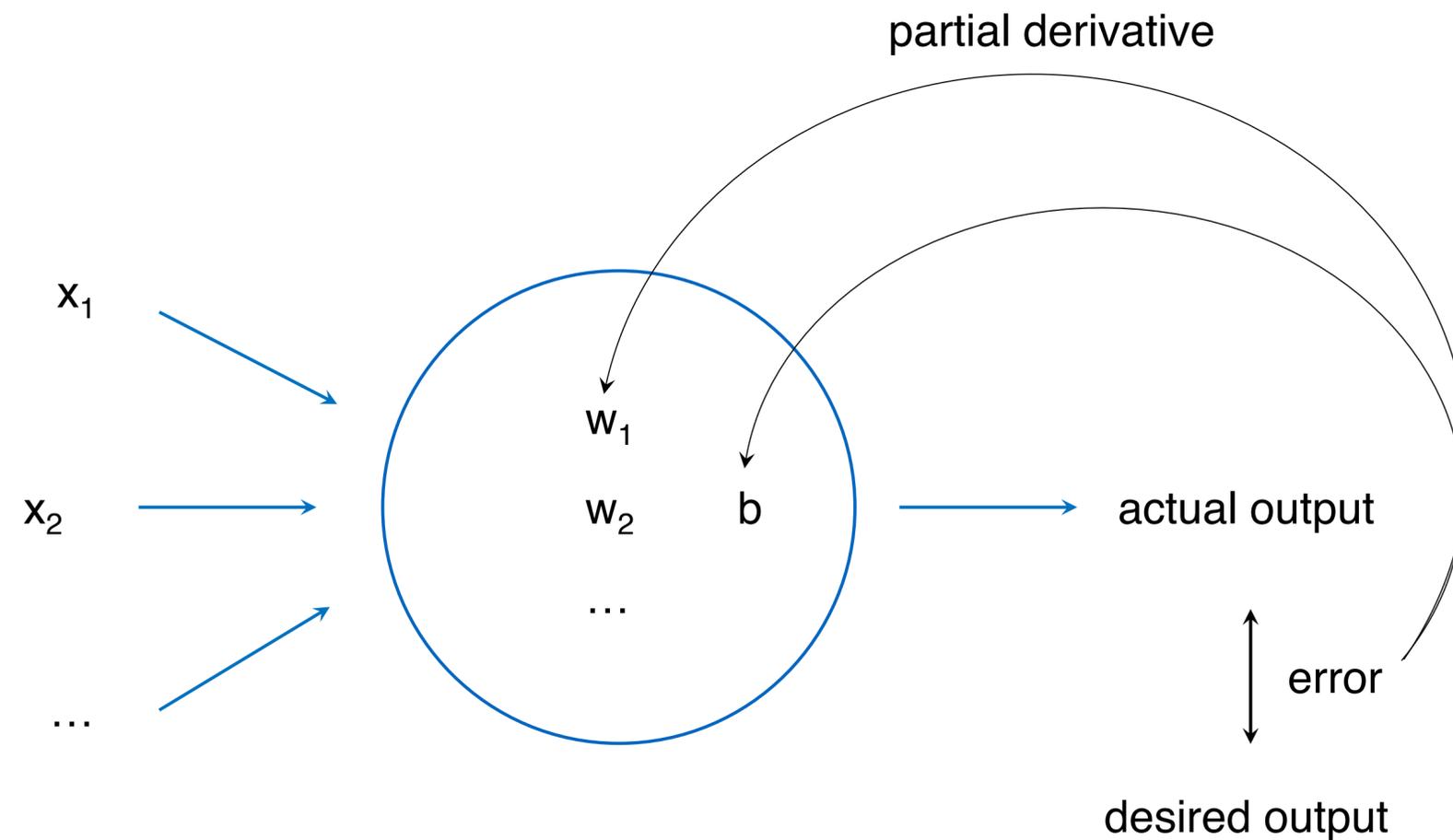
desired output

The idea of Backpropagation is that, since we know the **error between the actual and desired outputs**, and the network is differentiable, we can measure **how much contribution each *w* and *b* has to the error using partial derivative**, and adjust them accordingly.

Why it works?

A function approximator, it finds the **correlation** between the **input** and the **expected output**
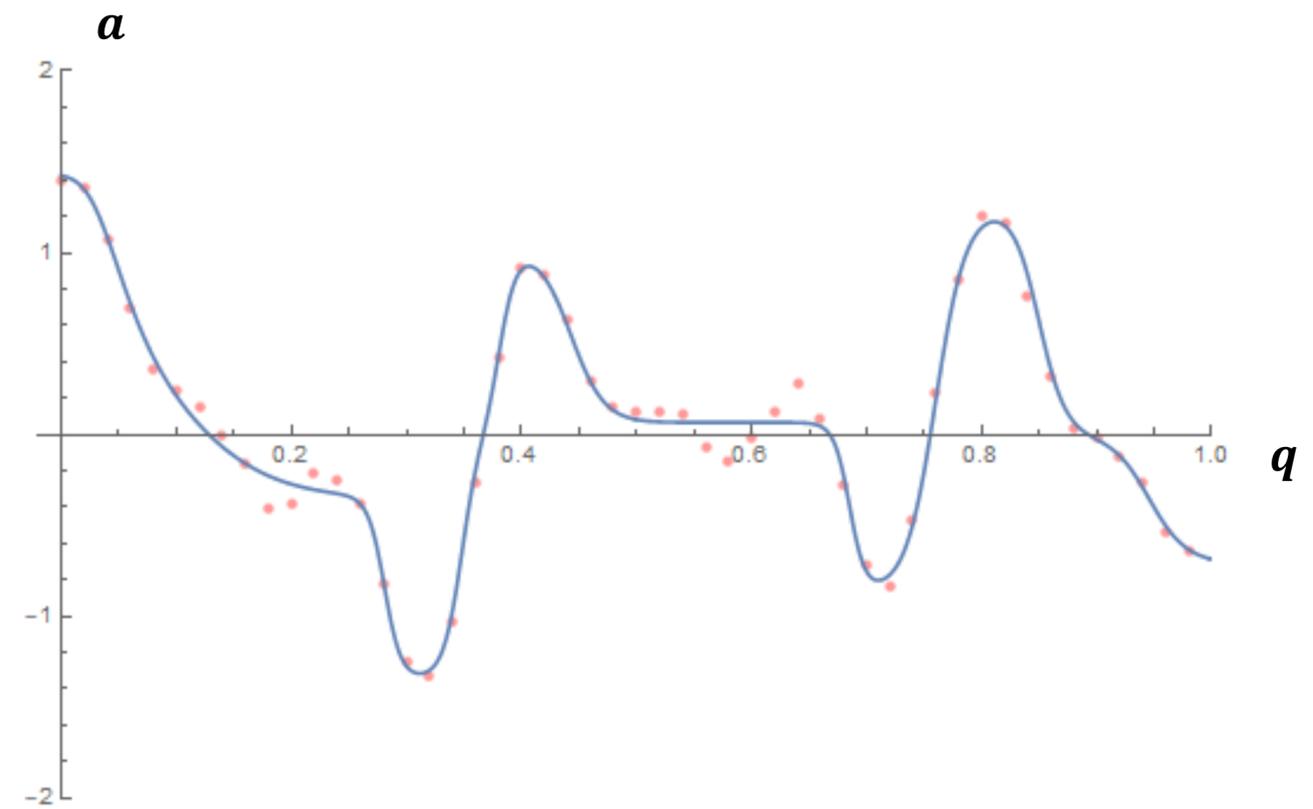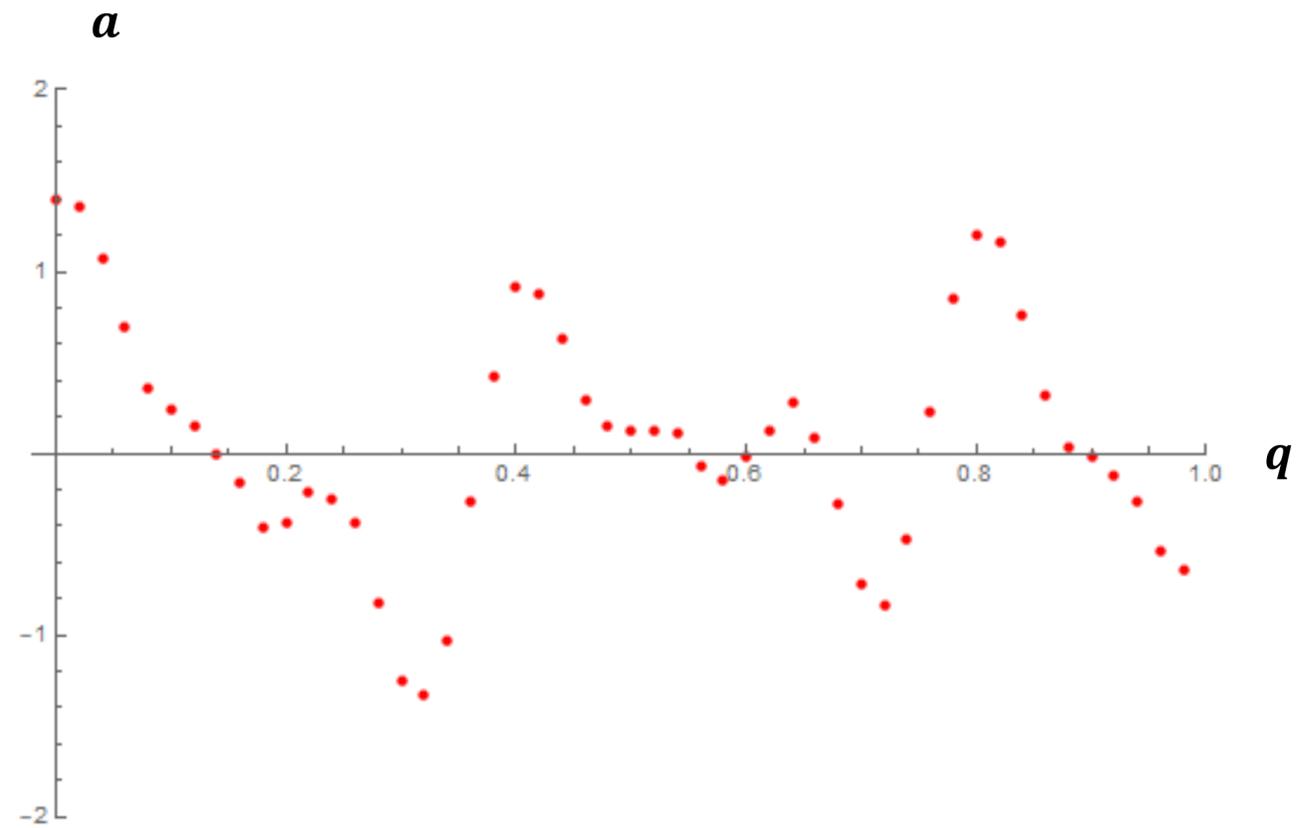
$$\boldsymbol{a} = f(\boldsymbol{q})$$

Deep Learning?

Machine learning on a model with many
layers, called Deep

Alpha Go
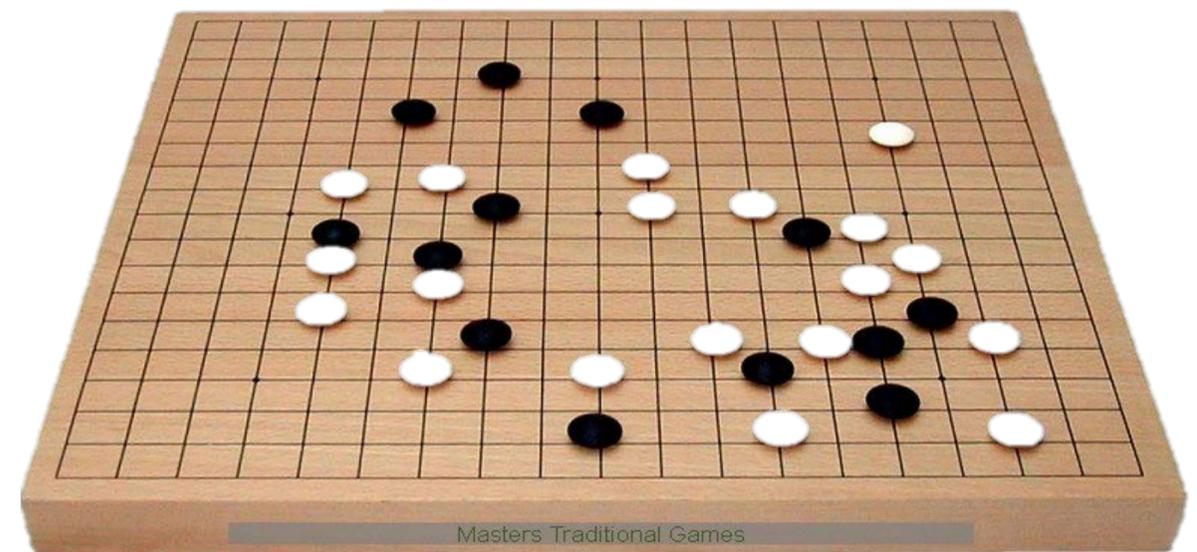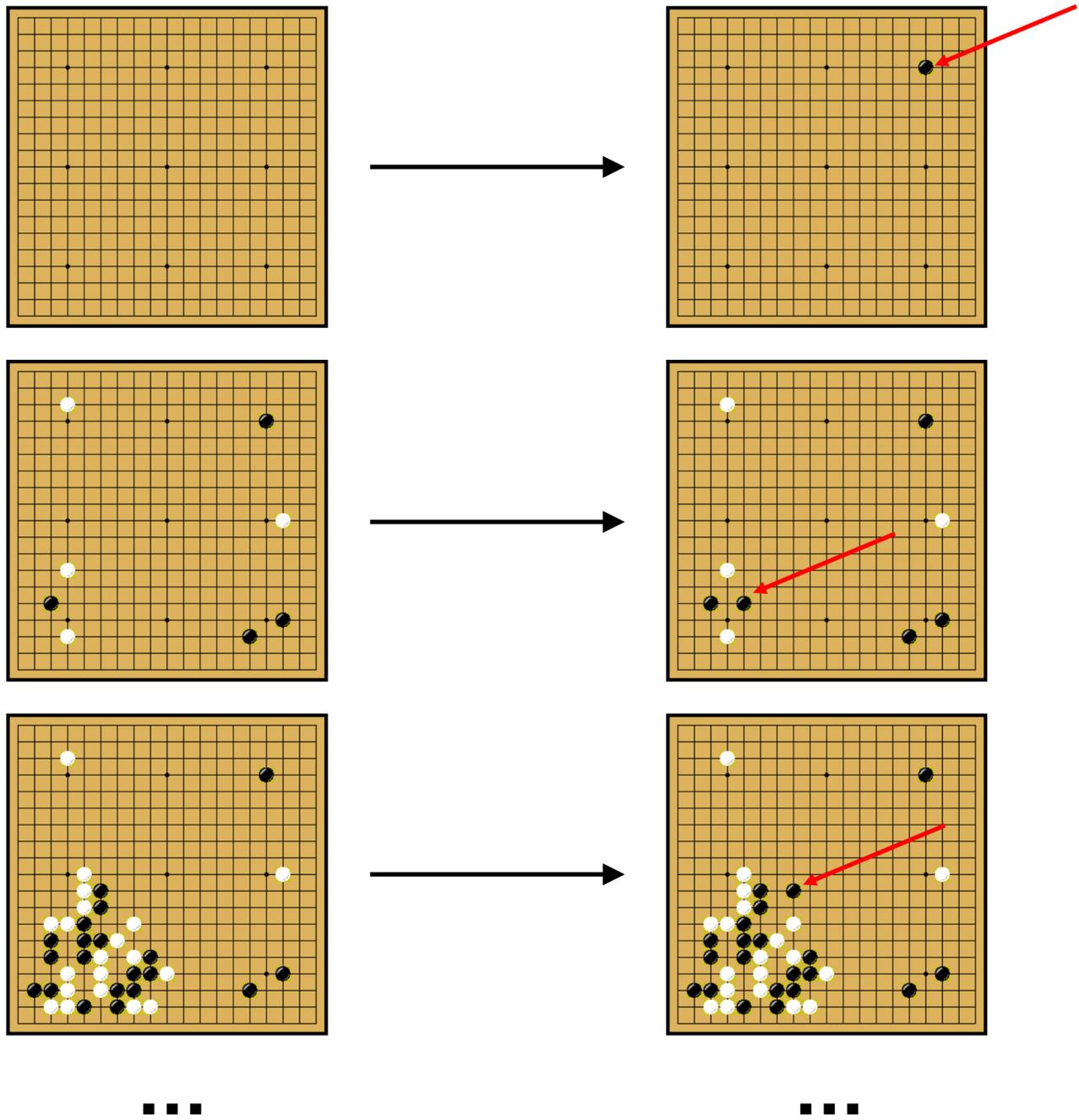
Alpha Go learns the correlation between **state/input** (left) and **decision/output** (right) by observing millions rounds of games by human players.

https://upload.wikimedia.org/wikipedia/commons/f/f3/Go-board-animated.gif

training data

mathematical model

state
(input)

action
(label)

state
(input)

as image

action
(label)
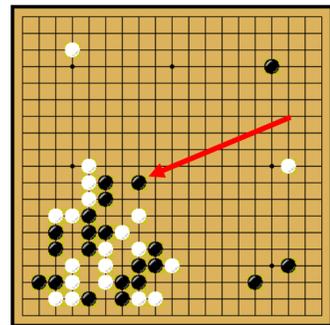
as coordinate
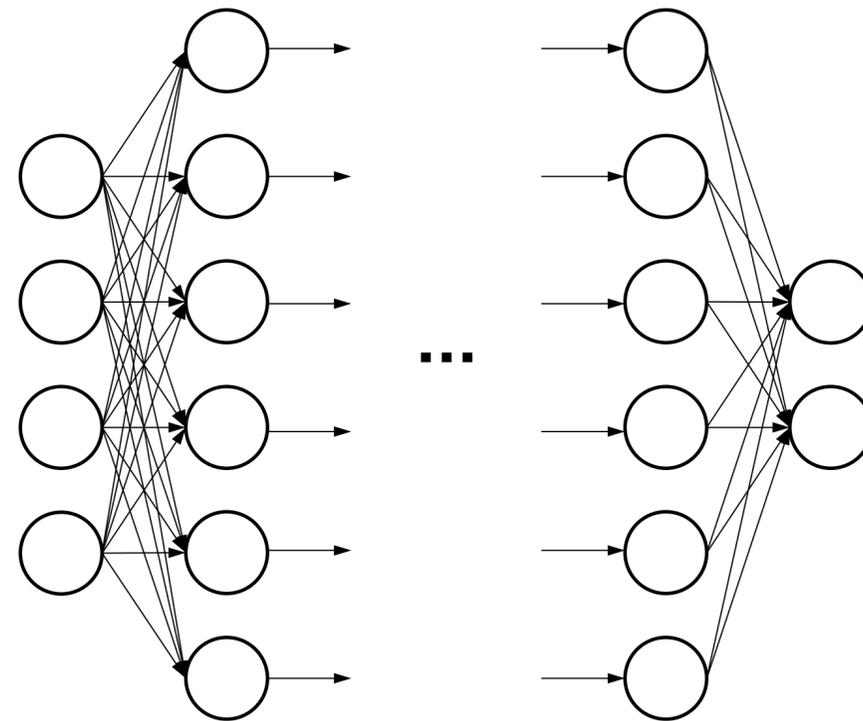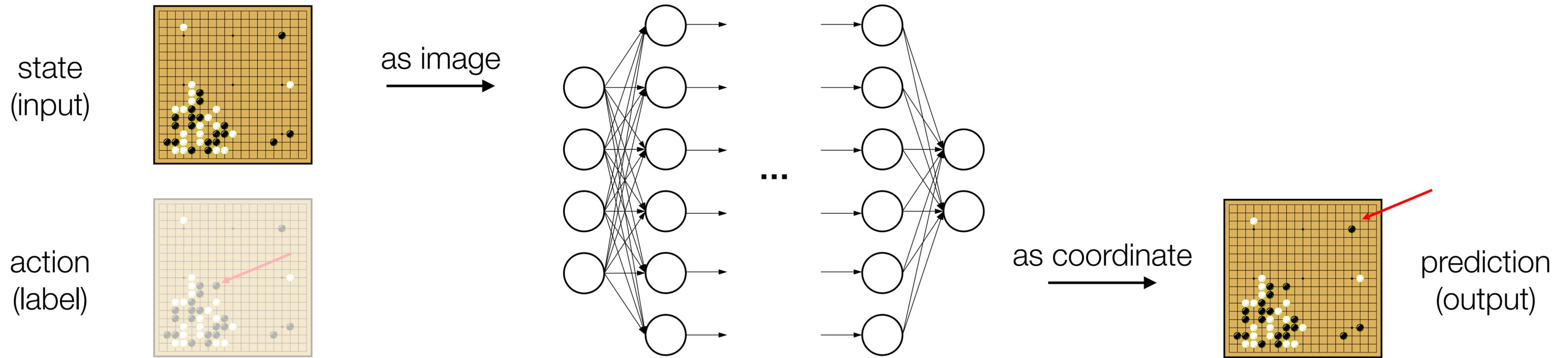
prediction
(output)

As we will show you in todays experiments, both images and coordinates are all numbers, the neural network is the projection from one numerical space (images) to the other (coordinates)

state
(input)
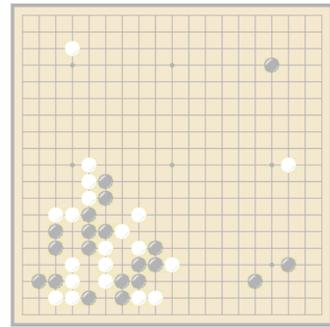
action
(label)
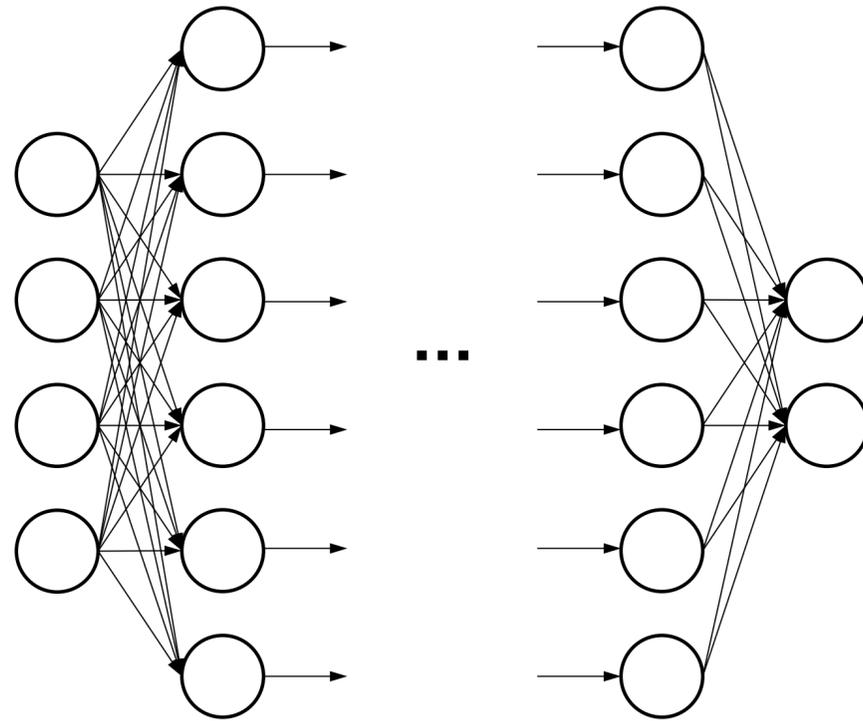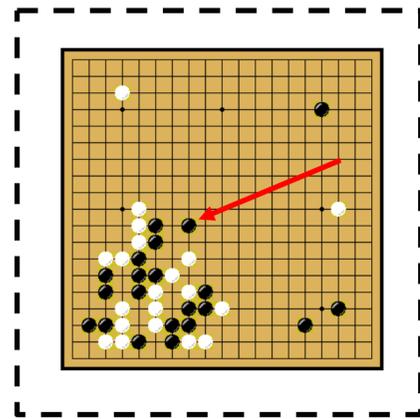
prediction
(output)
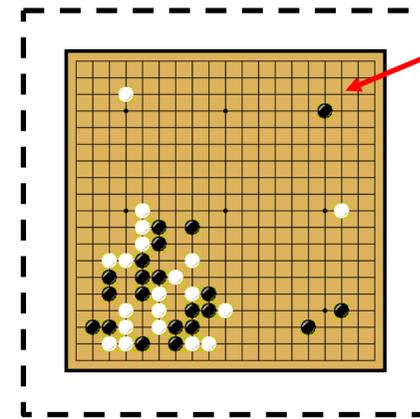
how error the prediction is

Logic-based Approach (that claims to be probabilistic)

Probabilistic Graph – specific model + specific data

Koller, D., & Friedman, N. (2009). Probabilistic graphical models: principles and techniques. MIT press.

SPRINKLER

RAIN	T	F
F	0.4	0.6
T	0.01	0.99
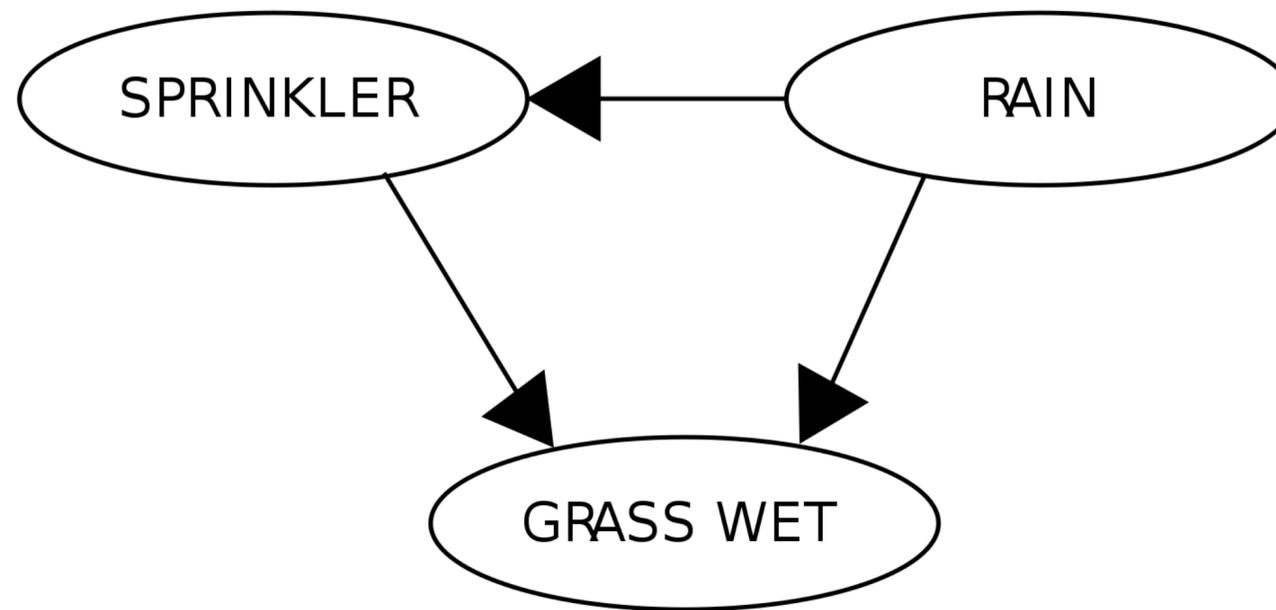
RAIN

	T	F
	0.2	0.8

GRASS WET

SPRINKLER	RAIN	T	F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

https://en.wikipedia.org/wiki/Bayesian_network#/media/File:SimpleBayesNet.svg

Am I Hungry (**H**)?

	S	
H	T	F
T	0.1	0.9
F	0.9	0.1

Have I 50 Fr. (**F**)?

Go to Sleep (**S**)

Go to Restaurant (**R**)

Buy a Burger (**B**)

		R	
F	**S**	T	F
T	T	0.0	1.0
T	F	0.8	0.2
F	T	0.0	1.0
F	F	0.0	1.0

		B	
F	**S**	T	F
T	T	0.0	1.0
T	F	0.2	0.8
F	T	0.0	1.0
F	F	0.8	0.2

Am I Hungry?

Have I 50 Fr. ?

Go to Sleep

Go to Restaurant

Buy a Hamburger

# Applications of Probabilistic Approach

The model observes millions of labeled data provided by human and learns the correlation within them, and then is able to predict new cases.

https://waymo.com/tech/
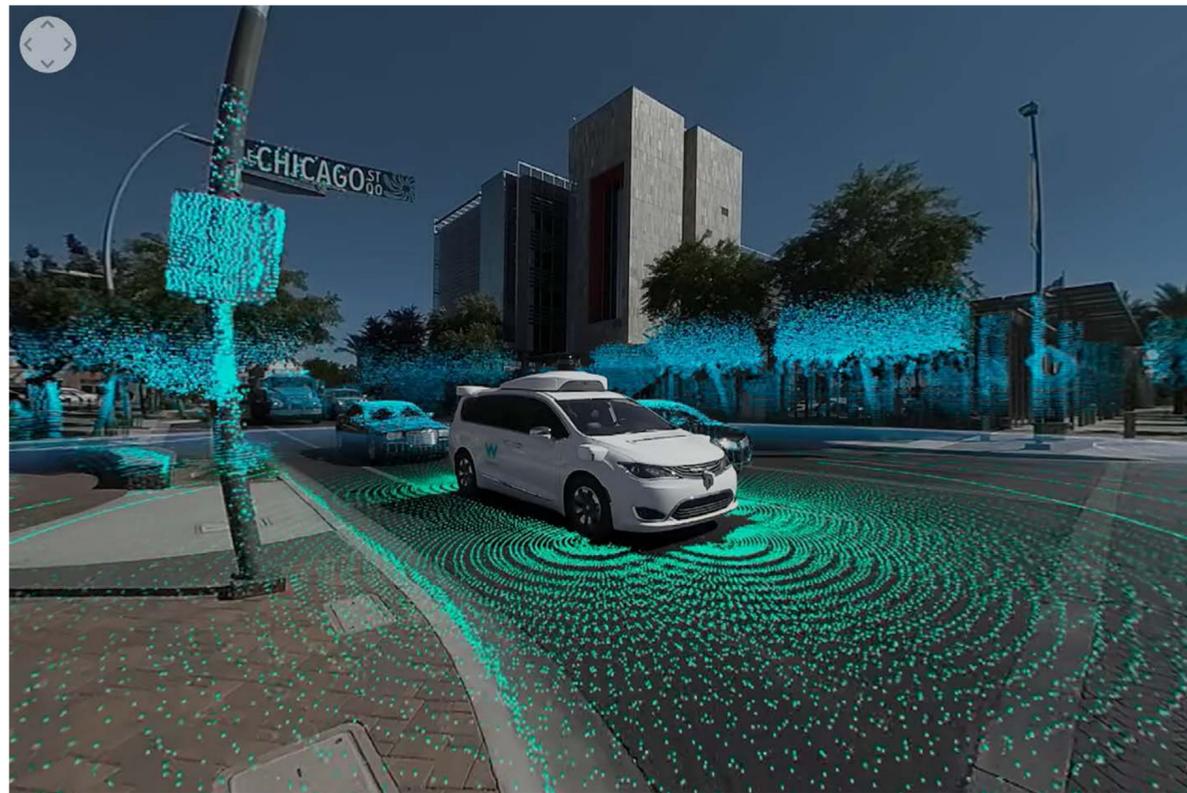
Self-Driving Car?

Equipped with multiple sensors, the self-driving car can "see" its surroundings, just like Alpha Go can "see" the go board.

By observing how human drives for millions of times, the self-driving car learns the correlation between the environment it sees and the actions it should take.
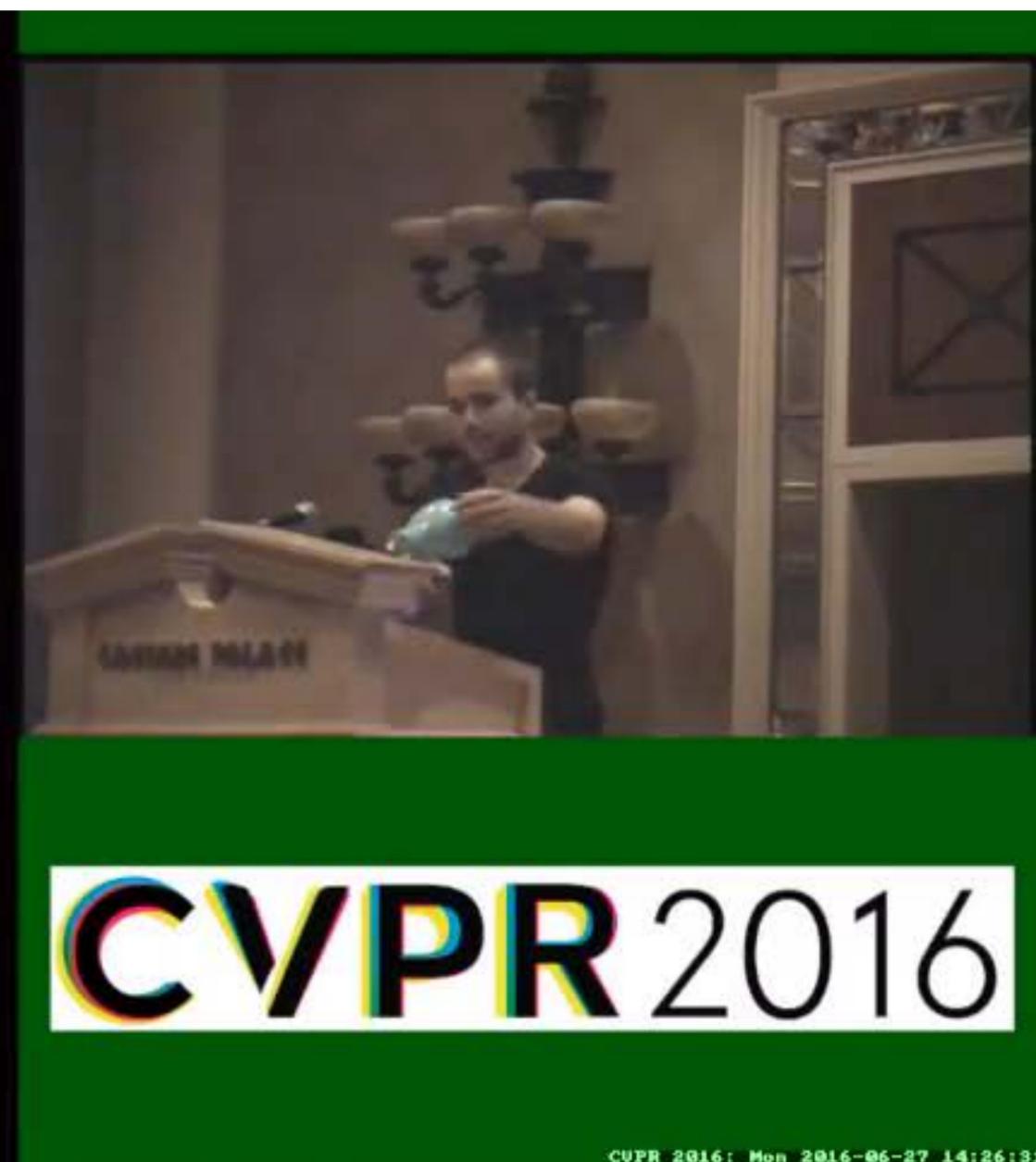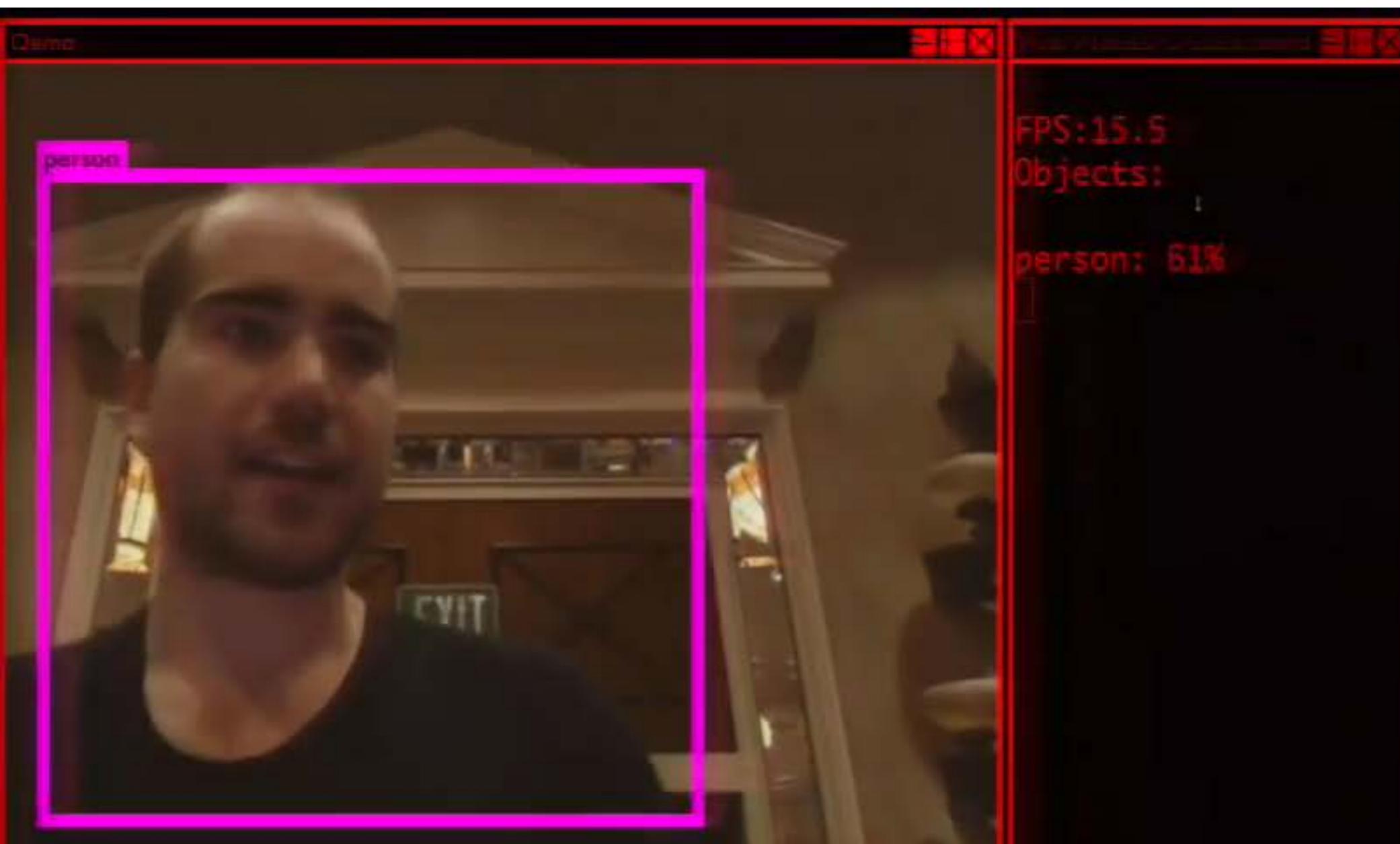
https://www.youtube.com/watch?v=NM6Irxy0bxs

https://www.youtube.com/watch?v=ZJMtDRbqH40

Question ⇄ **Answer**

"The question and answer method seems to be suitable for
introducing almost any one of the fields of human endeavor that
we wish to include"

Turing, Alan (1950). "Computing Machinery and Intelligence," Mind, 59, 434-460 (reprinted in Feigenbaum and Feldman, 1963, and in Anderson, 1964).

Oracle

Let us suppose we are supplied with some unspecied means of
solving number-theoretic problems; a kind of oracle as it were.
. . . this oracle . . . cannot be a machine.
With the help of the oracle we could form a new kind of machine
(call them o-machines), having as one of its fundamental
processes that of solving a given number-theoretic problem."

Alan Turing

The jointing of Intelligent Human and Machine Intelligence

Alvarez D., Saldana K., (2017), City of indexes, Elective course 2017, ETH ITA CAAD Praxis 2018

What is role of Machine Intelligence in design and architecture?

Would it take over?