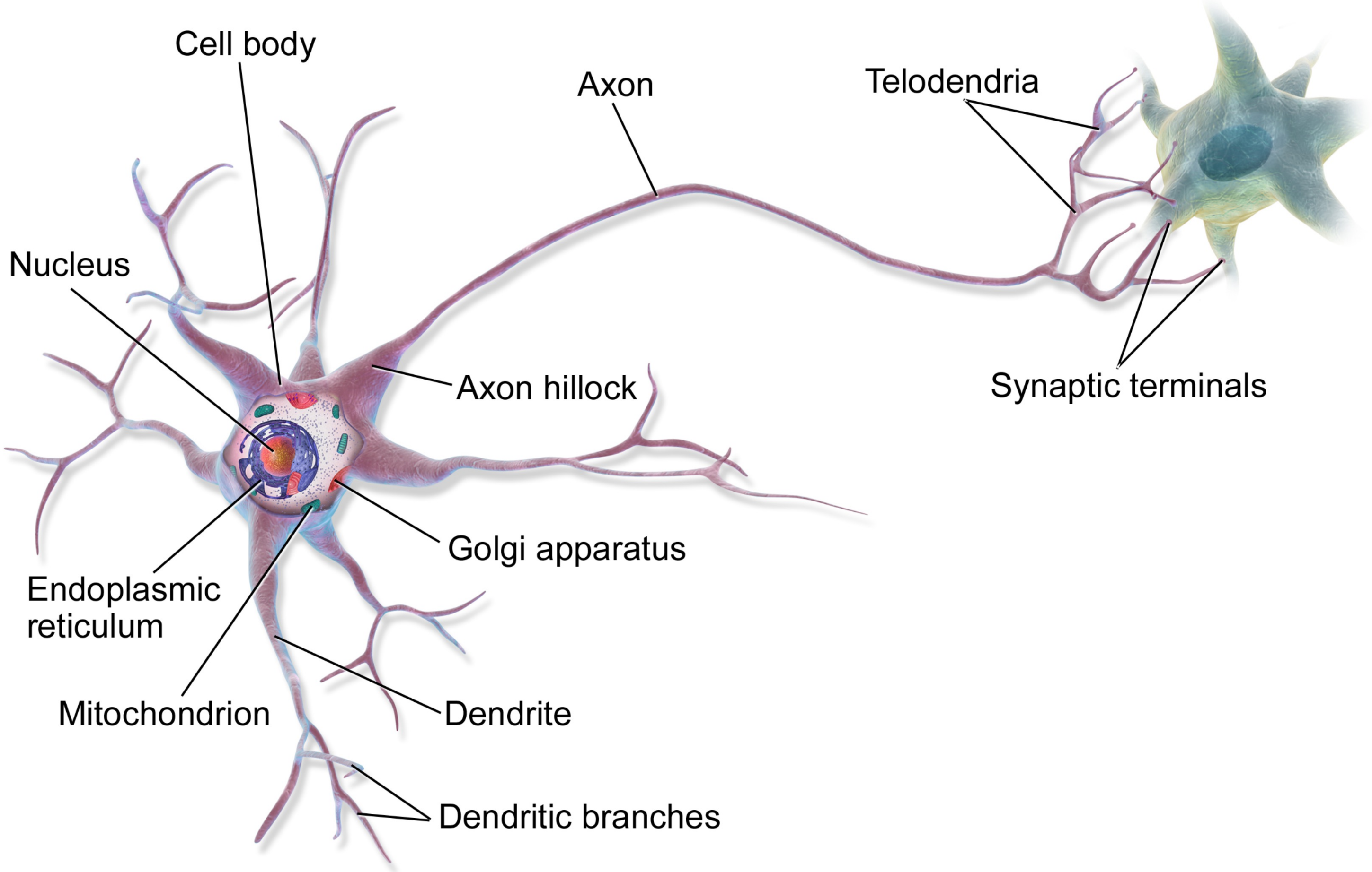
Neuro-Symbolic Artificial Intelligence

Sargur N. Srihari
srihari@buffalo.edu

Human and Biological Intelligence

https://www.youtube.com/watch?v=PIevD5Mjeak

Neurons


Cell body
Axon
Telodendria
Nucleus
Axon hillock
Synaptic terminals
Golgi apparatus
Endoplasmic reticulum
Mitochondrion
Dendrite
Dendritic branches

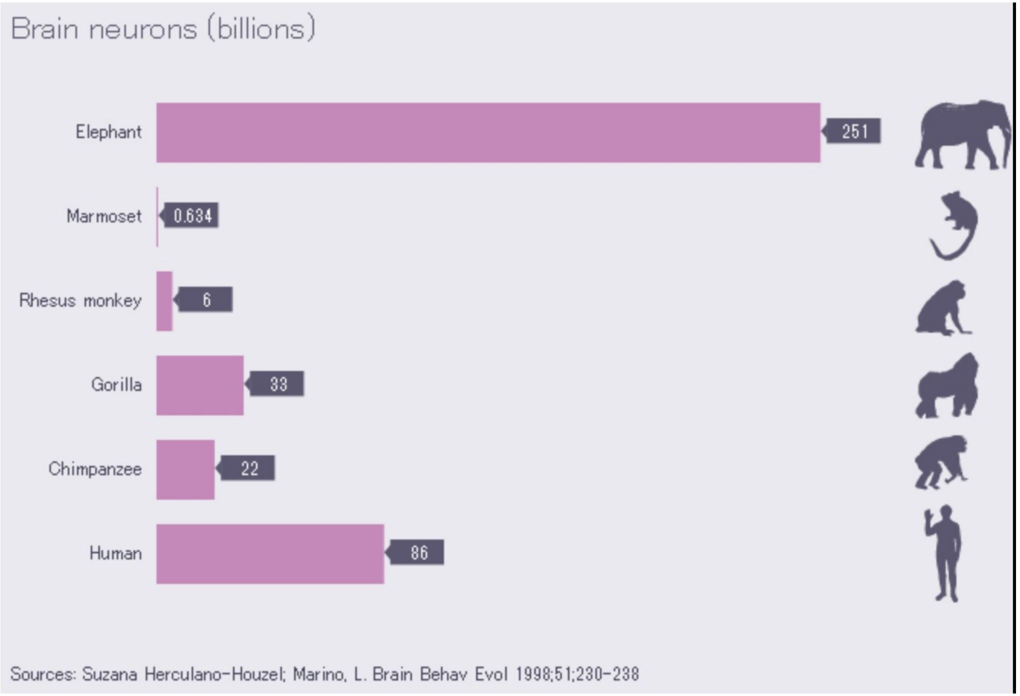
Number of Neurons

Whole nervous system

Tardigrade	about 200		Brain only	
Ciona intestinalis **larva** (sea squirt)	231	8,617 (central nervous system only)		
Caenorhabditis elegans (roundworm)	302	~7,500	is the only organism to have its whole connectome (neuronal "wiring diagram") completed.[10][11][12]	[13]
Jellyfish	5,600		*Hydra vulgaris* (*H. attenuate*)	[14]
Megaphragma mymaripenne	7,400			[15]
Box jellyfish	8,700–17,500		adult *Tripedalia cystophora* (8 mm diameter) – does not include 1000 neurons in each of the four rhopalia	[16]
Medicinal leech	10,000			[17]
Pond snail	11,000			[18]
Sea slug	18,000			[19]
Amphioxus	20,000		central nervous system only	[20] [21]
Larval zebrafish	100,000			[22]
Lobster	100,000			[23]
Fruit fly				

Dog (Komondor)	3.99×10^{9}		
Dog (Transylvanian Hound)	4.39×10^{9}		
Lion	4.667×10^{9}		
Greater kudu	4.91×10^{9}		
Rhesus macaque	6.376×10^{9}		
Brown bear	9.586×10^{9}		
Giraffe	1.075×10^{10}		
Yellow baboon	1.095×10^{10}		
Chimpanzee	2.8×10^{10}		
Orangutan	3.26×10^{10}		
Gorilla	3.34×10^{10}		
Human	8.6×10^{10}	$\sim1.5\times10^{14}$	Neurons for average adult
African elephant	2.57×10^{11}		

No of neurons



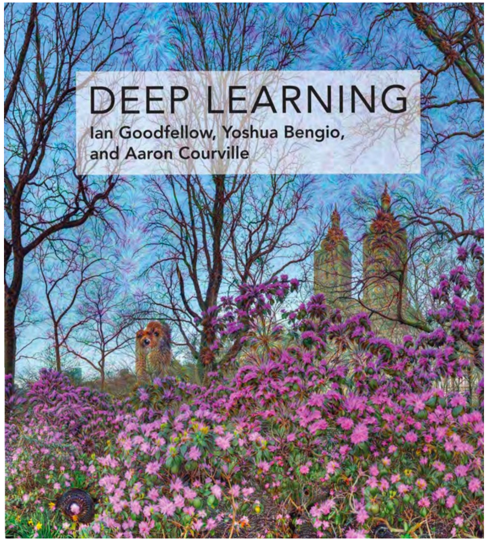
List of animal species by forebrain (cerebrum or pallium) neuron number

Asian elephant		6,775,000,000			
Human	16,340,000,000 21,000,000,000*		Isotropic fractionator Optical fractionator	Pallium (cortex)	*Homo sapiens*

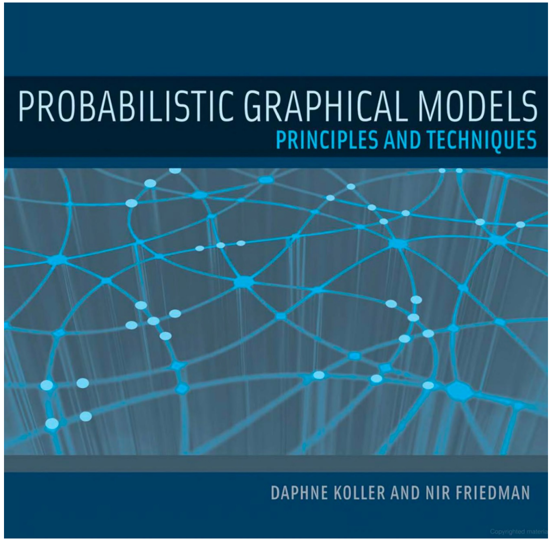
Stories and Abstraction (YN Harari)

- Success of Homo-Sapiens over other primates
  - Ability to abstract high-level *stories*
    - Allow collaboration with others at scale
    - Thus have a broader impact/influence
- Abstracting concepts is necessary for stories
  - Conscious decision beyond involuntary reaction
    - To abstract, generalize, and reason about scenarios and concepts that do not physically exist
    - To focus attention to certain, limited set of features and process them in depth, while disregarding others

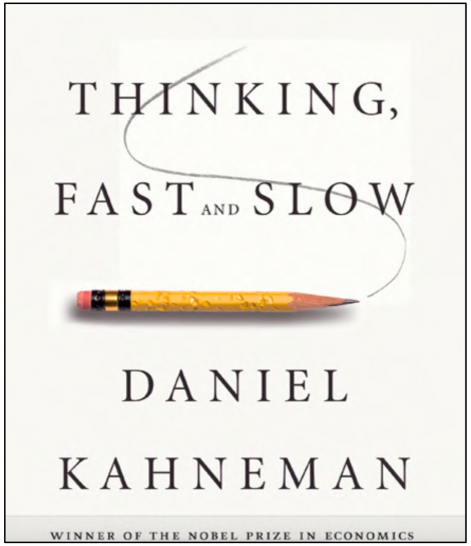
Foundations of Neurosymbolic AI



775 pages



1,232 pages

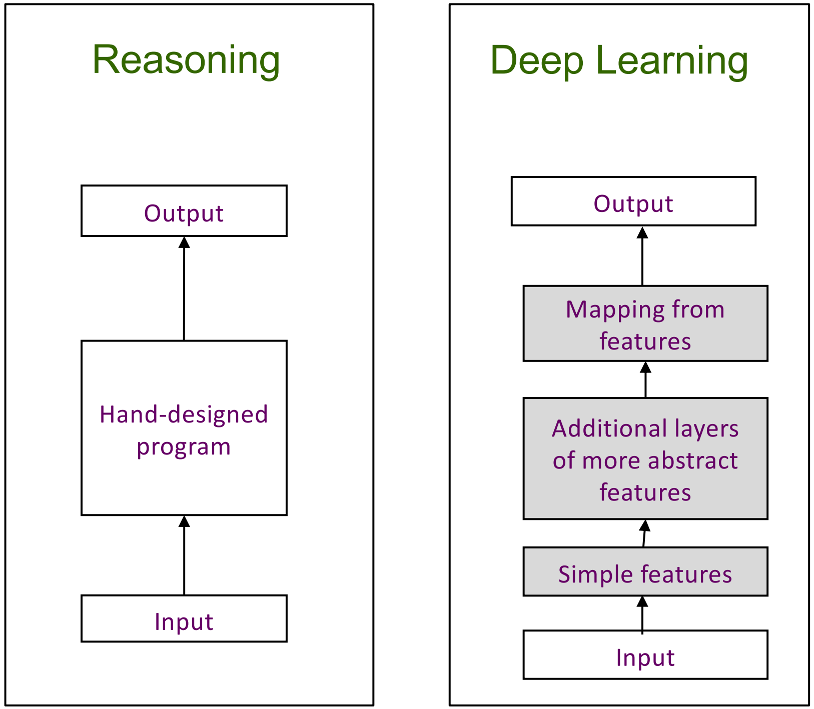


499 pages

Topics

- Paradigms of Artificial Intelligence (AI)
- Causal Reasoning
- Deep Learning

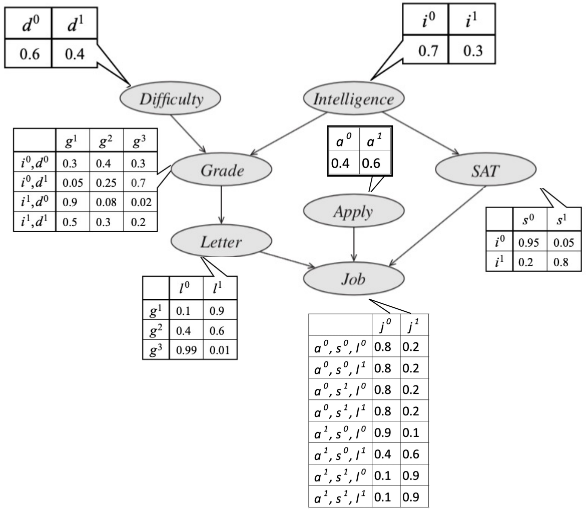
Two Paradigms of AI



Shaded boxes indicate components that can learn from data

Reasoning

- It is intuitive and subjective
- Knowledge-based Approach
  - Hard-code knowledge
  - Need to have deep knowledge of the problem domain



$$P(I=1|J=1) = 0.0629/(0.0629+0.0387) = 0.619$$

Deep Learning

1. Neural networks:
   - Inspired by neural networks of the brain
2. Learning:
   - Machines to discover rich/useful representations
3. Depth:
   - As a composition of learned features and functions

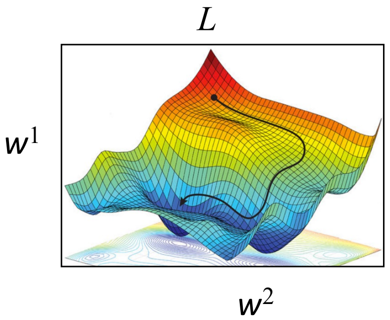
Deep Learning

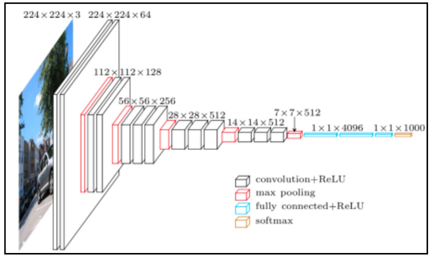
Understand the world as a hierarchy of concepts

- How these concepts are built on top of each other is deep, with many layers

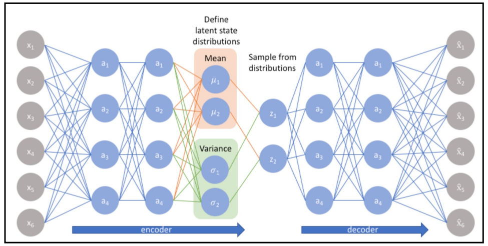
- Weights ***w*** learnt by gradient descent

$$\boldsymbol{w}^{(t+1)} = \boldsymbol{w}^{t} - \varepsilon \, \nabla_{\boldsymbol{w}} L(\boldsymbol{w}^{t})$$



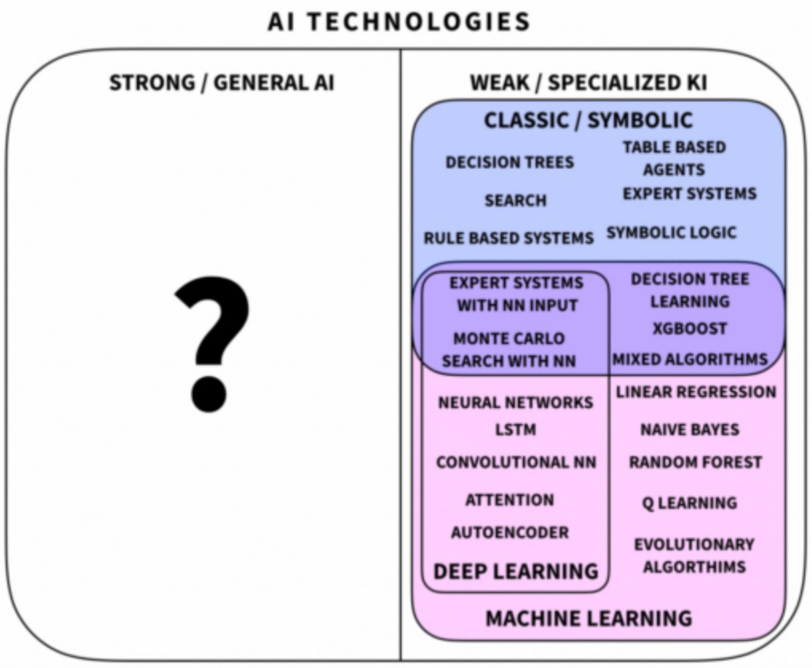


Discriminative Model
CNN



Generative Model
VAE

AI Technologies



Current status of AI

- Human capabilities still lacking in AI
  - Adaptability
  - Generalizability
  - Robustness
  - Explainability
  - Abstraction
  - Common sense
  - Causal reasoning
- Current research addresses these limitations

Research Goal

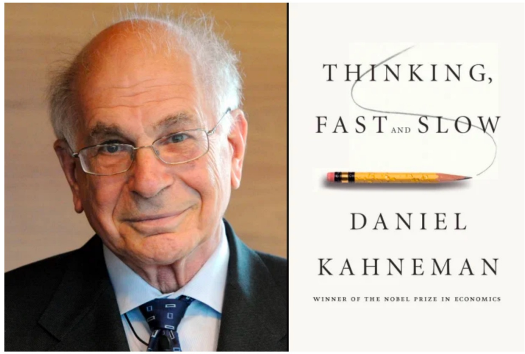
- Can we embed the causal components?
- Complex and seamless integration of learning and reasoning supported by both implicit and explicit knowledge
- Debate: whether end-to-end DL can achieve this goal or whether we need to integrate ML with symbolic and logic-based AI techniques
- Hypothesis: Integration route is most promising

Cognitive Approach

- Comprehension of how humans have, and have evolved to obtain, these advanced capabilities can inspire innovative ways to imbue AI systems with these competencies
- Study and exploit cognitive theories of human reasoning and decision making as a source of inspiration for the causal source of these capabilities,

Kahneman’s Theory

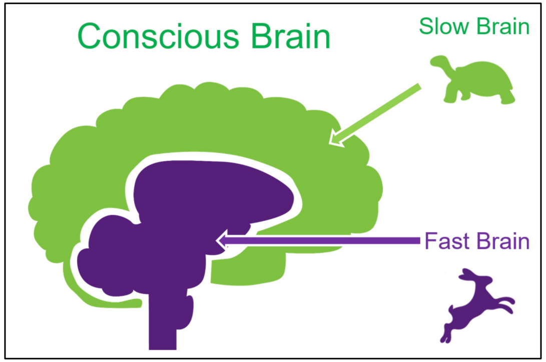
- Human decisions are supported and guided by the cooperation of two main kinds of capabilities
  - System 1: Thinking Fast
    - Provides tools for intuitive, imprecise, fast, and often unconscious decisions
  - System 2: Thinking Slow
    - Handles more complex situations where logical and rational thinking is needed to reach a complex decision

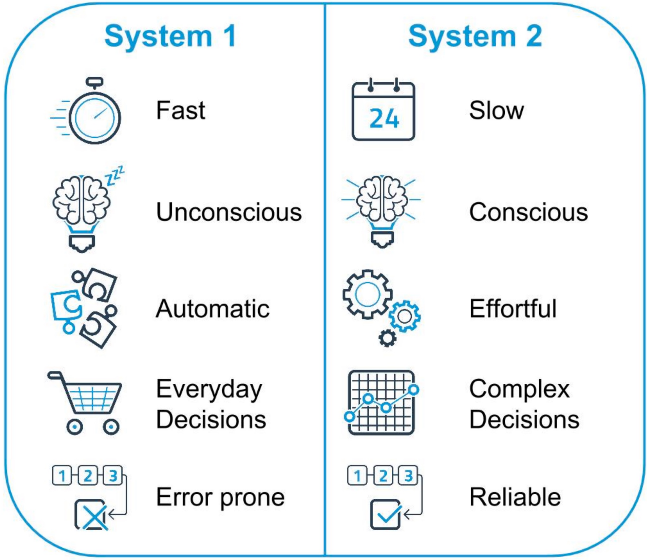


Kahnemann on Two Systems

[Kahnemann]

Summary of two systems

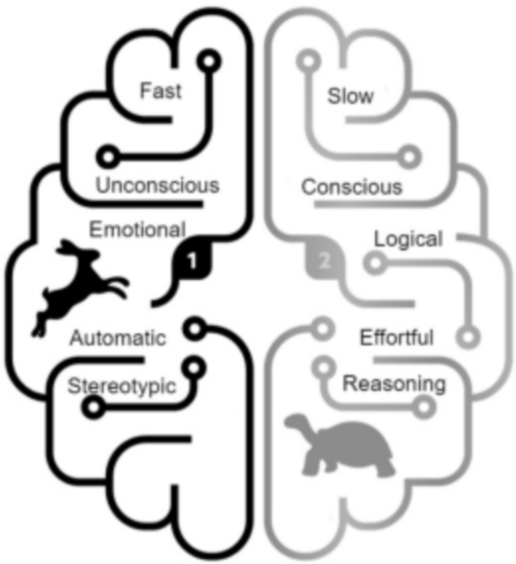




Source:
https://medium.com/@ryansheffer/founders-need-to-think-slow-move-fast-6b683e94c110

Further Summary

System 1	System 2
drive a car on highways	drive a car in cities
come up with a good chess move (if you're a chess master)	point your attention towards the clowns at the circus
understands simple sentences	understands law clauses
correlation	causation
hard to explain	easy to explain



System 1 Characteristics

- Guided mainly by intuition not deliberation
  - Fast answers to very simple questions.
  - Answers are sometimes wrong, due to unconscious bias or reliance on heuristics
- Usually do not have an explanation
- Able to build models of the world
  - Although inaccurate/imprecise, can fill knowledge gaps through causal inference
  - Allow us to respond well to the many stimuli of our everyday life

Examples of System 1 Tasks

1. Finding the answer to a very simple arithmetic calculation
2. Reaching out to grab something that is going to fall

- We use our system 1 about 95% of the time when we need to make a decision

Role of System 2

- When the problem is too complex for System 1
- System 2 kicks in
- Solves it with access to:
  - Additional computational resources
  - Full attention, and
  - Sophisticated logical reasoning.

Example of System 2 Task

- Solving a complex arithmetic calculation, or a multi-criteria optimization problem
- To do this, humans need to be able to recognize that a problem goes beyond a threshold of cognitive ease and therefore the need to activate a more global and accurate reasoning machinery.
- Hence, introspection is essential in this process

When is System 2 invoked

- A new/difficult problem is handled by system 2
- Certain problems over time pass on to system 1
  - Procedures used by system 2 are used to accumulate examples that system 1 can later use readily with little effort
    - Problems, initially solvable only by system 2 reasoning tools, can become manageable by system 1
    - E.g., Reading text in our own native language
- However, this does not happen with all tasks.
  - Finding correct solution to complex arithmetic

System 2 is supported by System 1

- System 2 does not usually work by itself
  - Supported by system 1 in its elaborate calculations
  - When searching for a solution in a very large solution space, system 2 does not usually explore the whole search space but may employ heuristics that are provided by system 1 and that help in focusing the attention only on the most promising parts of the space.
- This allows system 2 to work with manageable time and space

Data-level causality by System 1

- System 1 also capable of basic causal reasoning
  - Allows it to build a (possibly imprecise and biased) model of the world even if it has incomplete knowledge, and to use it to tackle simple tasks
- System 1 data-level causality skills support more complex/accurate reasoning of system 2 on more complex problems

Multi-tasking by System 1

- System 1 reasoning:
  - At various levels of abstraction
  - Adapt to new environments
  - Generalize from experience to other problems
  - Easily multi-task
- System 2 is sequential
  - Requires full attention to devise/execute appropriate procedure for complex or new problem
  - Only one complex task at a time handled

Local vs Global

- System 1 works at a local level
  - Activates only a specific part of the brain, e.g., recognize a familiar face, or when we speak
- System 2 involves more regions of the brain
  - Combines contributions-- works at a global level
- Multiple specialized agents/components/skills supports the functioning of both systems
  - Not as multi-agent architecture terminology but as metaphors for two broad classes of information processing

Experimental Justification

- Two forms of consciousness (Graziano)
  1. I(Information)-consciousness
     - Related to System 2
  2. M(Mystery)-consciousness
     - Related to System 1

I(Information)-consciousness

- Ability to solve (possibly complex) problems
  - By recognizing necessary processing steps in specific (even new) context, to tackle a desired problem
- Related to System 2’s high-fidelity processing mode
  - Considering a problem and harnessing the relevant faculties of our cognition to devise a plan to solve it

M(Mystery)-consciousness

- Ability to build a simplified, approximate, and subjective model of peoples', both ourselves and others, mind, beliefs and intentions.
- Such low-fidelity model building can be linked to system 1, as system 1 is able to form a rapid but usually inexact model of the world

Humans excel at I-consciouness

- While many animals and primates possess various levels of M-consciousness, and also limited forms of I-consciousness, humans excel at I-consciousness
- M-consciousness can even sustain in presence of limited I-consciousness, yet, a sophisticated I-consciousness needs to rely on M-consciousness:
  - without a model of the mind (of self and other agents), it is difficult to devise how to adapt to new environments and solve complex problems.

Summary

- Neural networks are what appear to make humans intelligent
- Two paradigms of AI : Causal Reasoning and Deep learning
- Can we embed causal reasoning into neural networks?
- I-consciousness and M-consciousness are two levels
- I-consciousness is what humans excel at