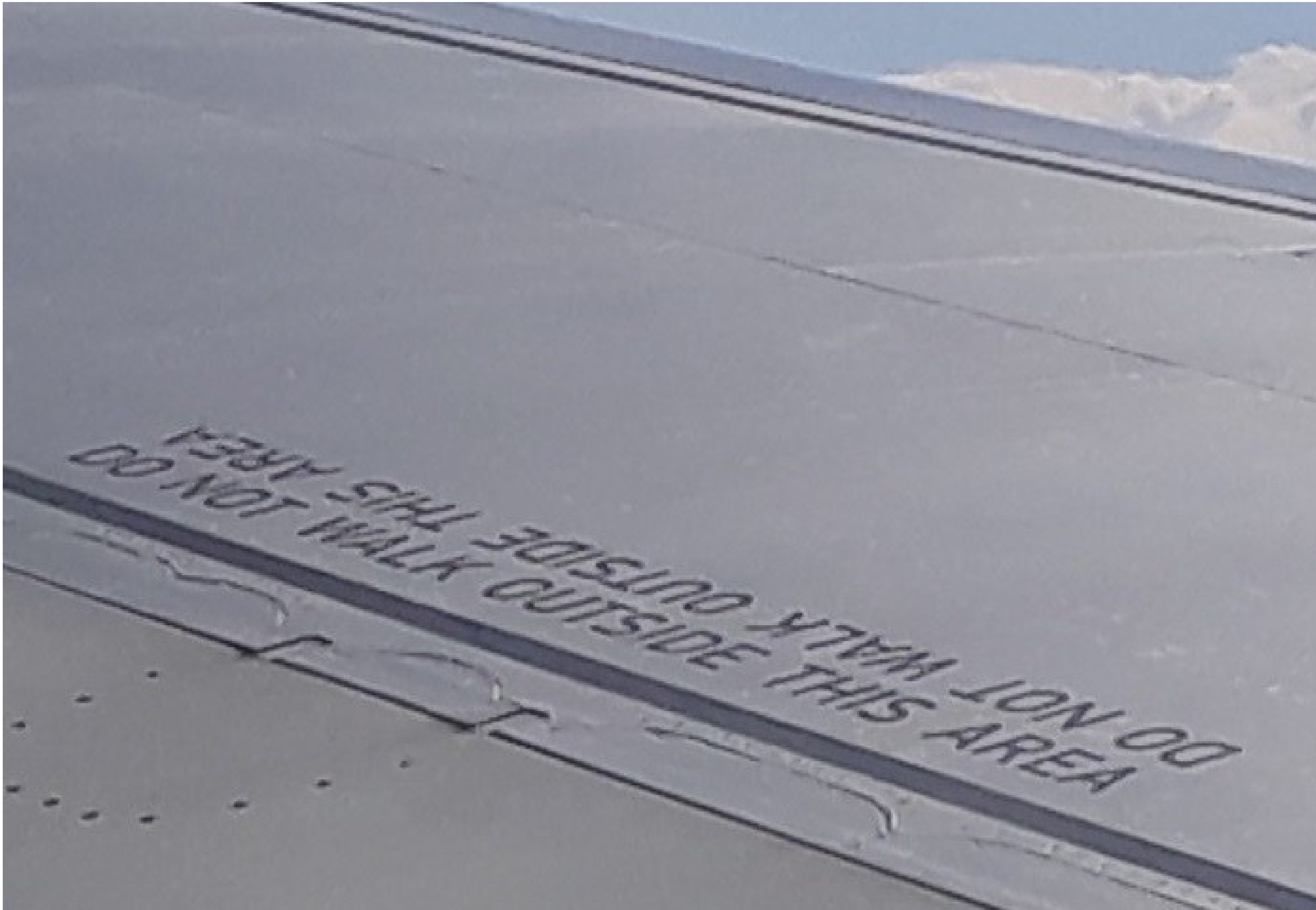
Horse 2017
Queen Mary University of London
20 September 2017

Avoiding Deep Horses: Finding Structure in Deep Networks

Artur d'Avila Garcez
City, University of London
a.garcez@city.ac.uk


DO NOT WALK OUTSIDE THIS AREA
DO NOT WALK OUTSIDE THIS AREA


Context is everything... commonsense!
DO NOT WALK OUTSIDE THIS AREA
DO NOT WALK OUTSIDE THIS AREA

The AI revolution...

The promise of AI:

- Education (adaptive learning)
- Finance (time series prediction)
- Security (image and speech recognition)
- Health (monitoring - IoT, drug design)
- Telecom (infrastructure data analysis)
- Games (interactive learning)
- Transport (logistics optimization)
- etc.

Brain/Mind dichotomy

Symbolic AI: a symbol system has all that is needed for general intelligence

Sub-symbolic AI: intelligence emerges from the brain (neural networks)

Machine Learning (ML)

Systems that improve their own performance from experience

Systems that, in addition, enable humans to improve their performance (human-machine interaction and human computation needed here)

Explainable AI: accountability, trust and transfer learning... c.f. EU GDPR Reg. 71

System Verification

Whose fault is it when a self-driving car gets into an accident?

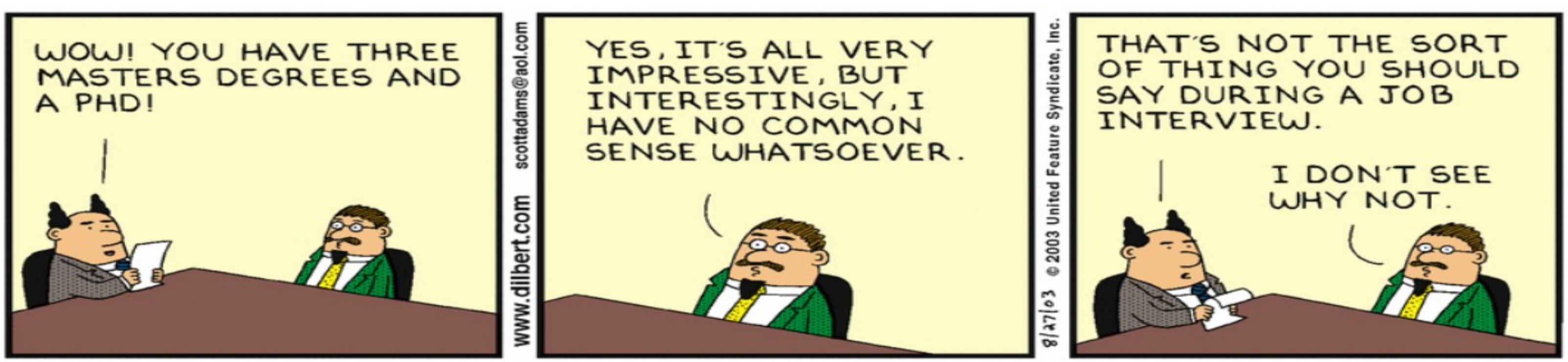
In the news recently: Tesla self-driving system cleared in deadly crash

There will be fewer deaths from self-driving cars...

Deep Learning

Given Big Data, deep learning (based on neural nets) works better than symbolic ML!

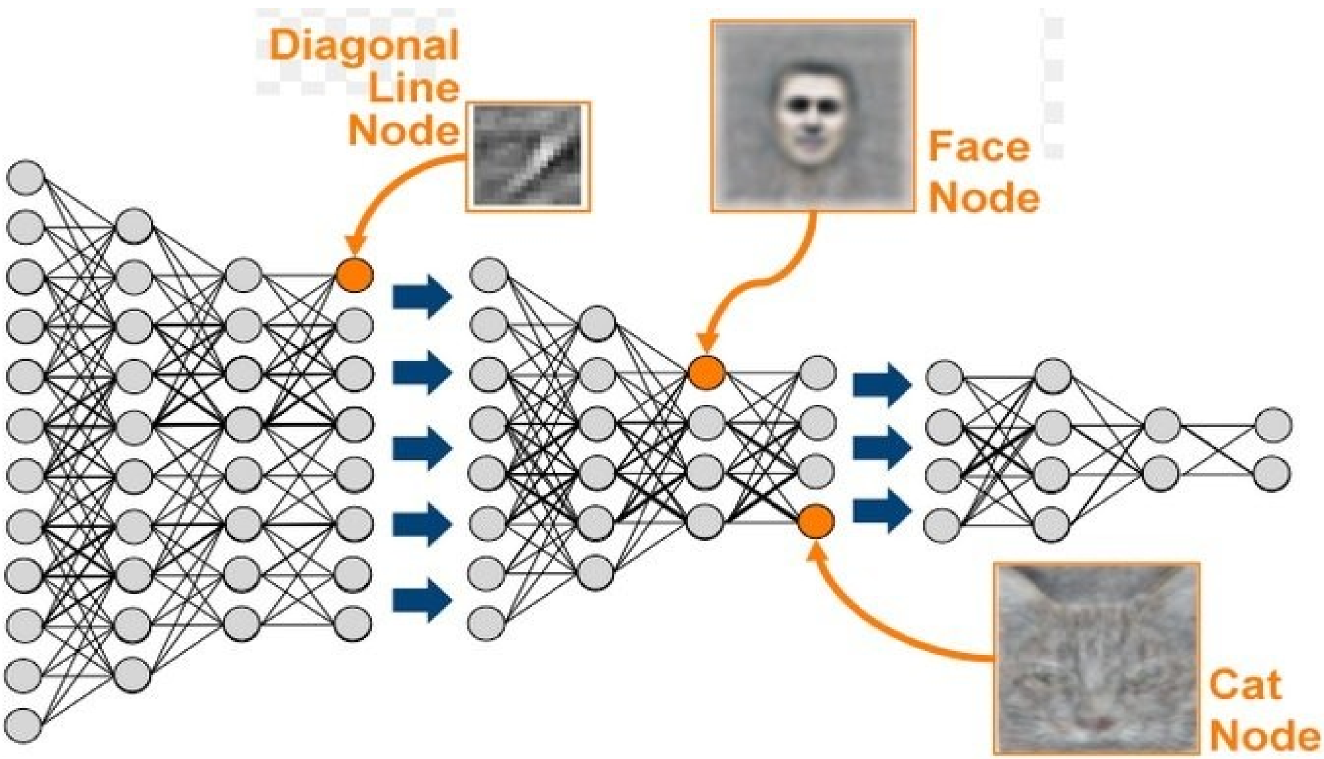
But, where is commonsense?



Deep networks

Very successful on handwritten digit classification, object recognition, speech/audio, games (AlphaGo)

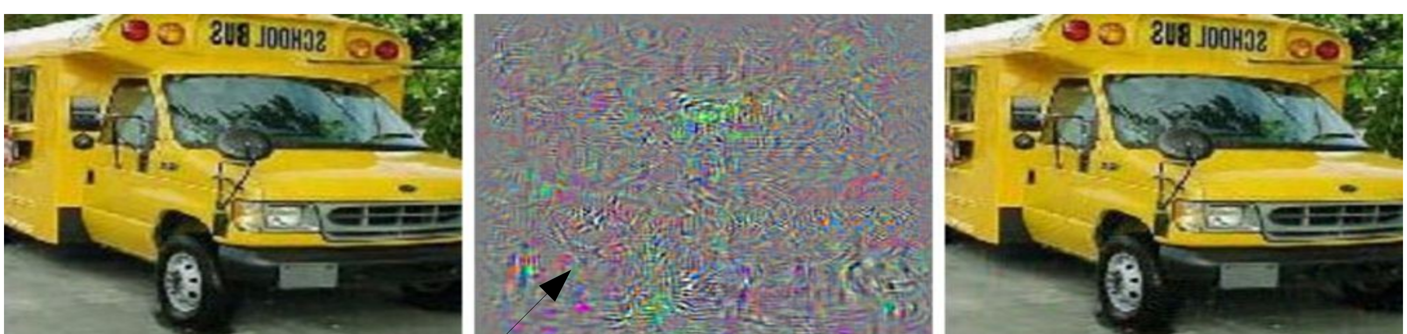
How about language? Video understanding?



Geoff Hinton says AI needs to start over...

- Deep neural nets (CNNs) do not recognise negative images

- Adversarial networks are not doing much better

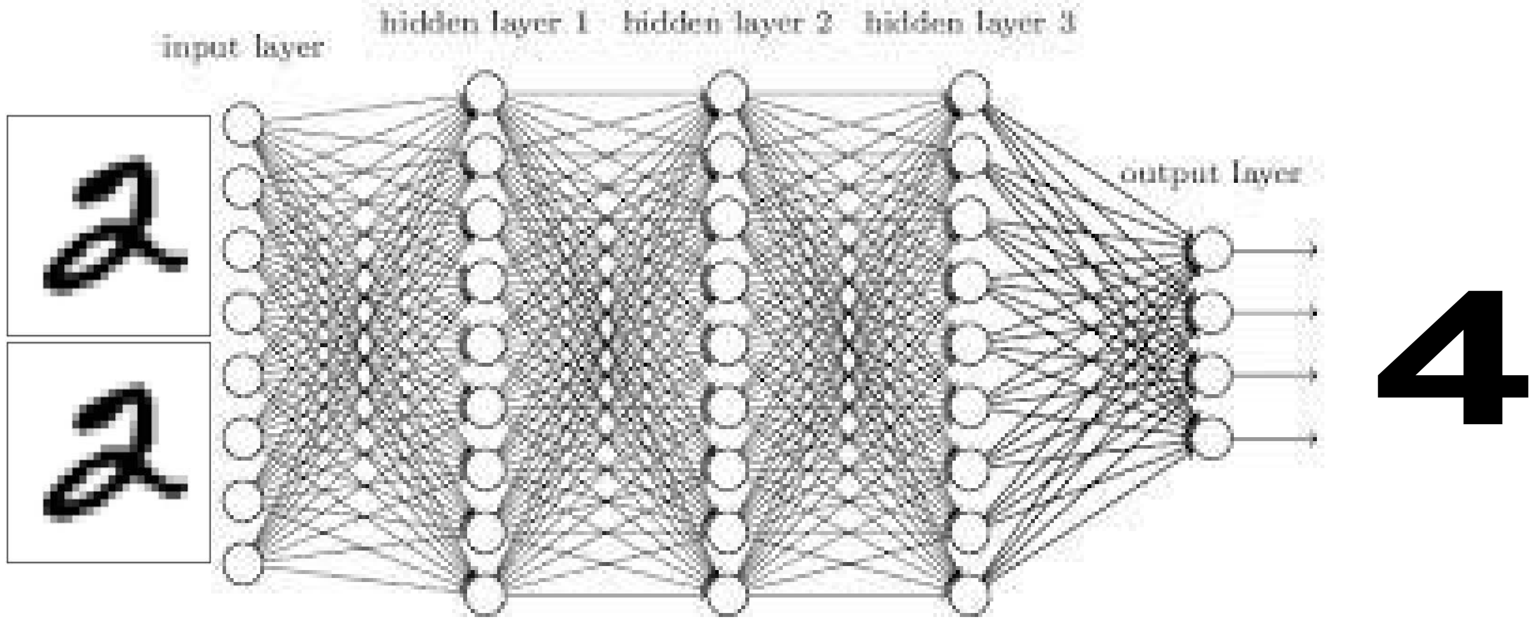


School bus

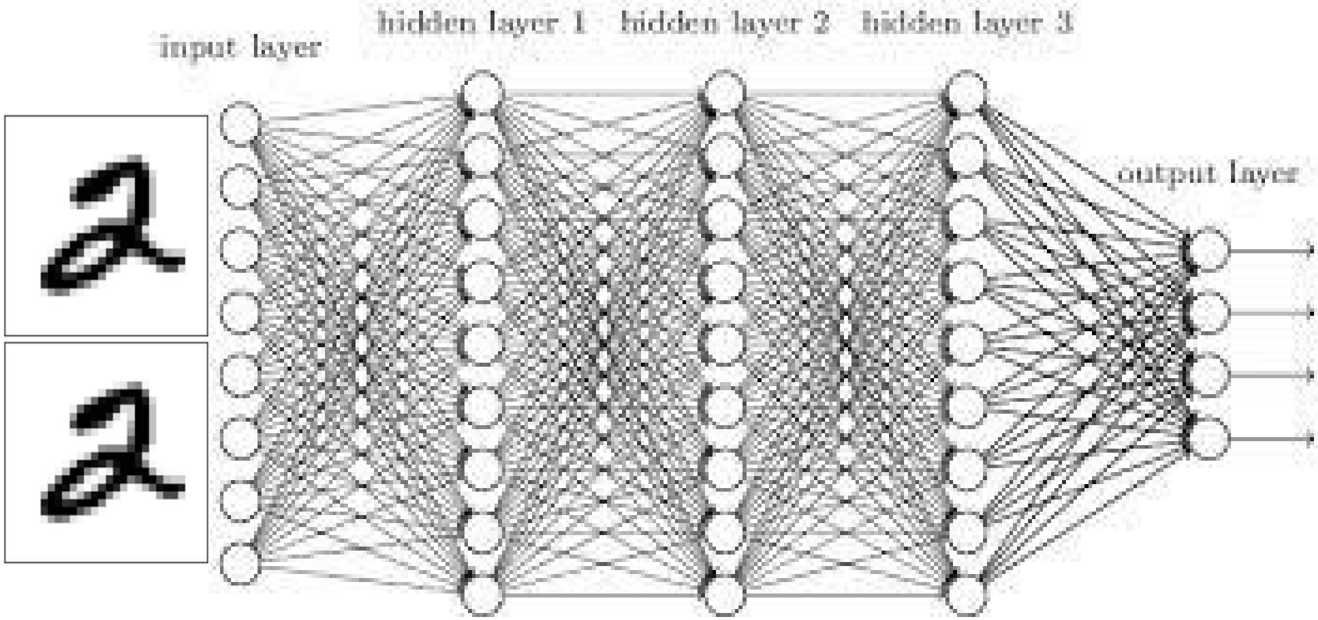
Distortions made to the adversarial example

Ostrich

Knowledge Extraction from Deep Nets



Knowledge Extraction from Deep Nets



4

2+2=4

Neural-Symbolic Systems

Cognitive Science

↓

Logic

Learning

Neural Computation

↑

Neuroscience

One Structure for Learning and Reasoning

(NSS = KR + ML)

Why Neurons and Symbols?

“We need a language for describing the alternative algorithms that a network of neurons may be implementing” L. Valiant

(New) Logic + Neural Computation

GOAL: Learning from experience and reasoning about what has been learned in an uncertain environment in a computationally efficient way

Neural-Symbolic Methodology

high-level symbolic representations
(abstraction, recursion, relations, modalities)

translations

low level, efficient neural structures
(with the same, simple architecture throughout)

Analogy: low-level implementation (machine code) of high-level representations (e.g. java, requirement specs)

A Foundational Approach

(as opposed to the neuroscience or the engineering approach)

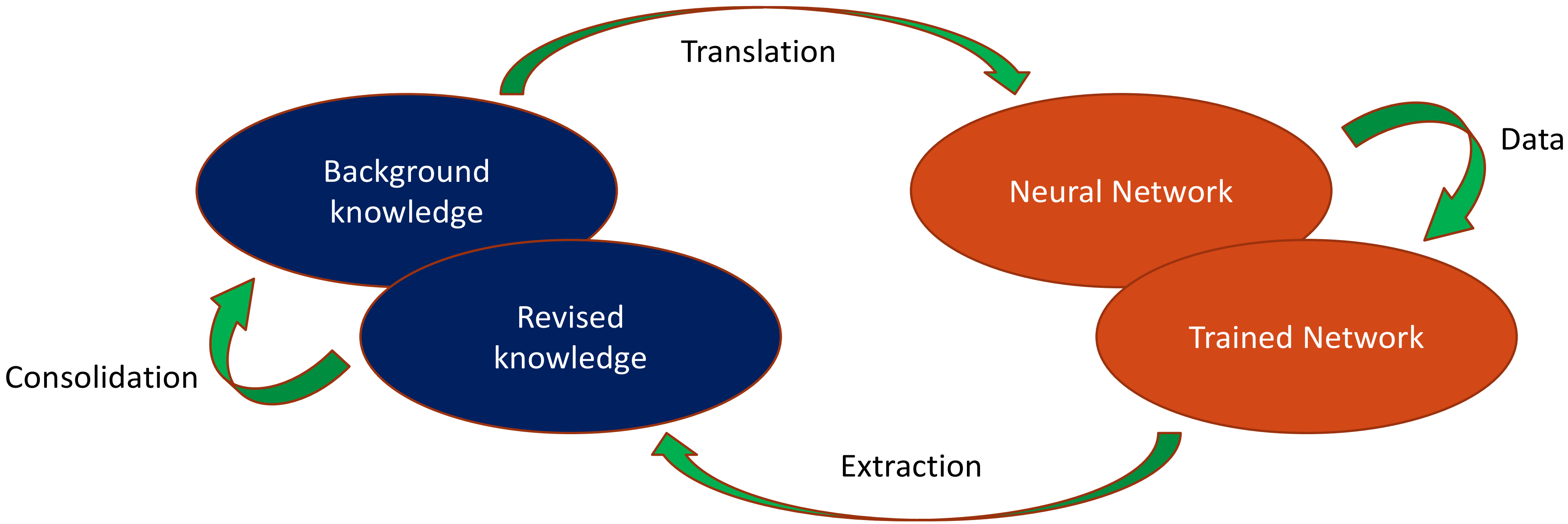
One Structure for Learning and Reasoning:

Take different tasks, consider what they have in common, formalize, evaluate and repeat

KEY: controlling the inevitable accumulation of errors (robustness)

Applications: training in simulators, robotics, evolution of software models, bioinformatics, power systems fault diagnosis, semantic web (ontology learning), general game playing, visual intelligence, finance, business compliance.

Neural-Symbolic Learning Cycle



Connectionist Inductive Logic Programming (CILP System)

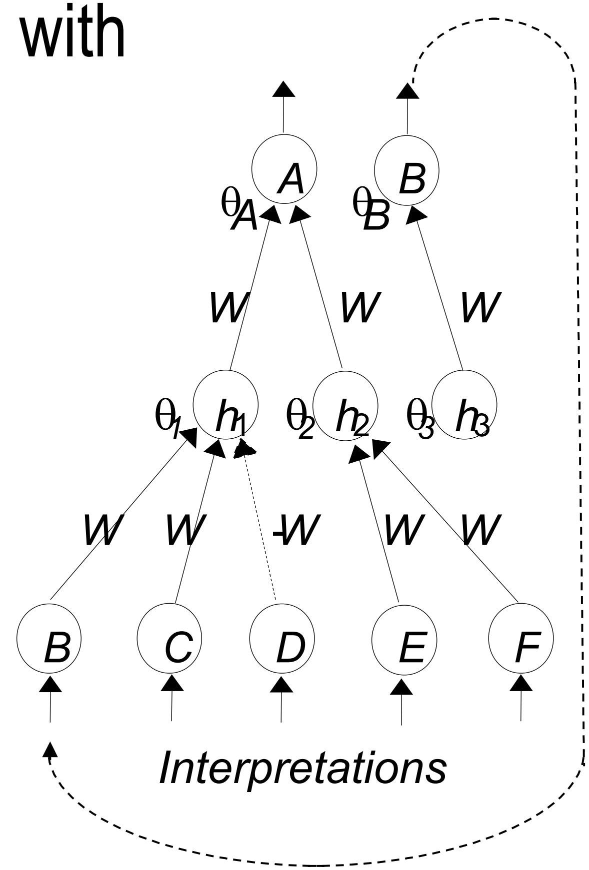
A Neural-Symbolic System for Integrated Reasoning and Learning (neural nets + logic programming)

Background Knowledge Insertion + Learning with Backpropagation + Knowledge Extraction

r_1: $A \leftarrow B, C, \sim D;$

r_2 : $A \leftarrow E, F;$

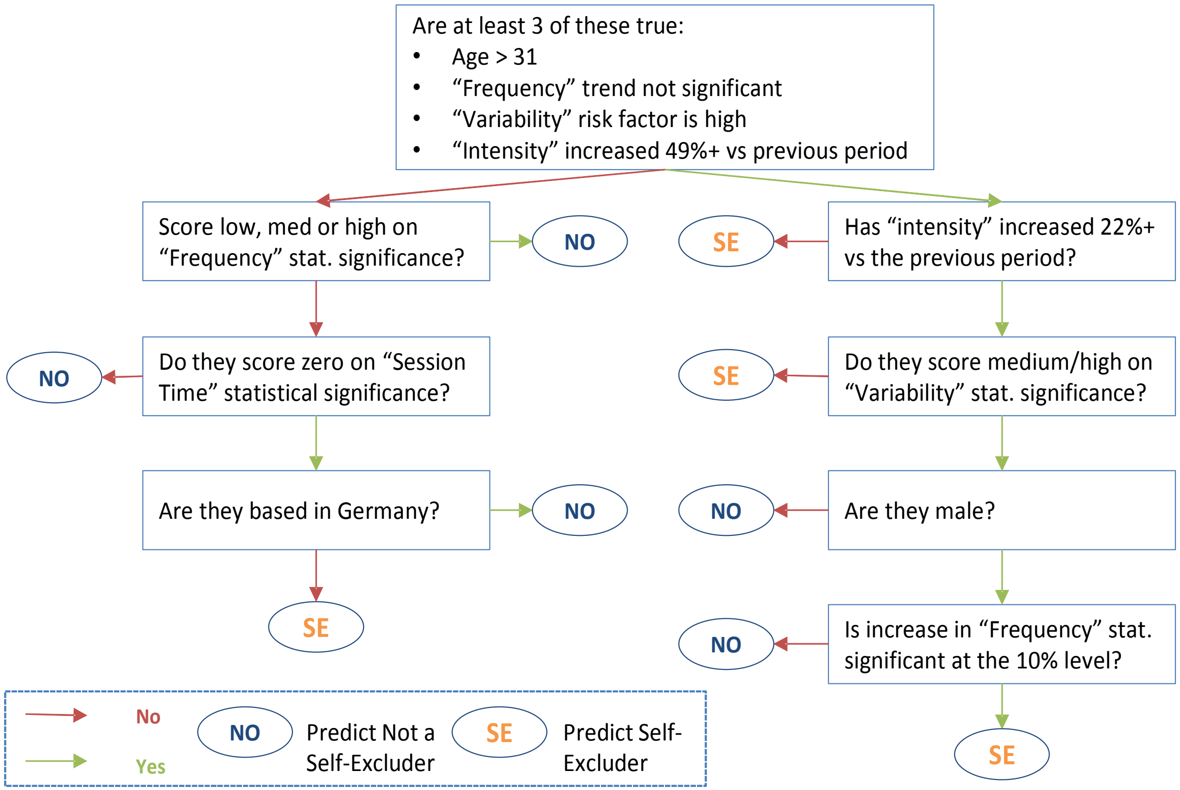
r_3 : $B \leftarrow$



Rule Extraction: Neural Net = Black Box?

- Extracted rules can be visualized in the form of a state transition diagram

- Alternatively, use TREPAN-like rule extraction and variations...



C. Percy, A. S. d'Avila Garcez, S. Dragicevic, M. Franca, G. Slabaugh and T. Weyde. The Need for Knowledge Extraction: Understanding Harmful Gambling Behavior with Neural Networks, In Proc. ECAI 2016, The Hague, September 2016.

Knowledge Consolidation

Challenge: efficient extraction of sound, comprehensible knowledge from large-scale networks (100's of neurons; 1000's of connections)

What makes knowledge comprehensible?

Transfer Learning

S. Tran and A. S. d'Avila Garcez. Deep Logic Networks: Inserting and Extracting Knowledge from Deep Belief Networks. IEEE TNNLS, Nov, 2016

There are many extensions of CILP using richer knowledge...

- The importance of non-classical reasoning: preferences, nonmonotonic, modal, temporal, epistemic, intuitionistic logic, abductive reasoning, value-based argumentation (dialogues).

- New applications: normative reasoning, temporal logic learning with model checking, software model adaptation (business process evolution), reducing harm from gambling, semantic image interpretation...

Business Process Modelling

Learning = process adaptation

A. Perotti, A. S. d'Avila Garcez and Guido Boella. Neural-Symbolic Monitoring and Adaptation. In Proc. IEEE/INNS IJCNN 2015, Killarney, Ireland, July 2015.

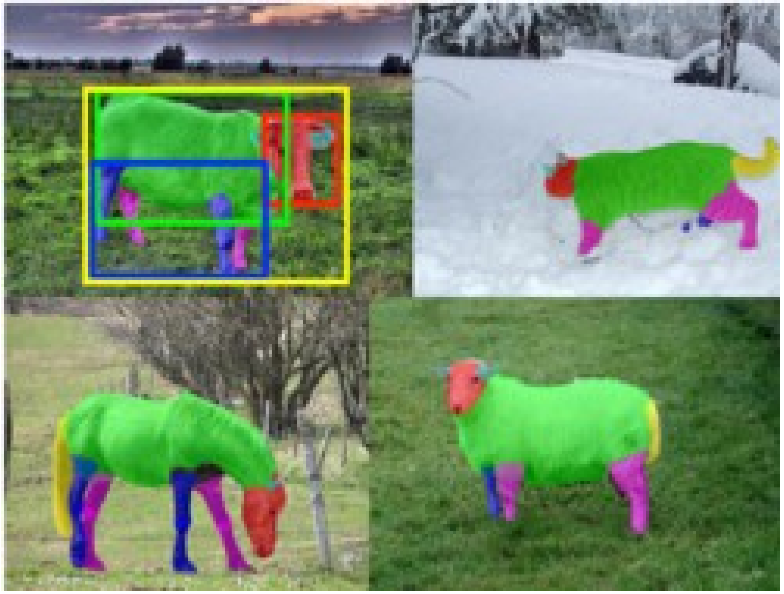
Semantic Image Interpretation

$$\forall xy(\mathsf{partOf}(x,y) \rightarrow \neg\mathsf{partOf}(y,x))$$

Normally, every cat has a tail

Q. Get me the red thing next to the sheep...

A. The horse's muzzle? Yes.

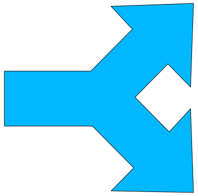


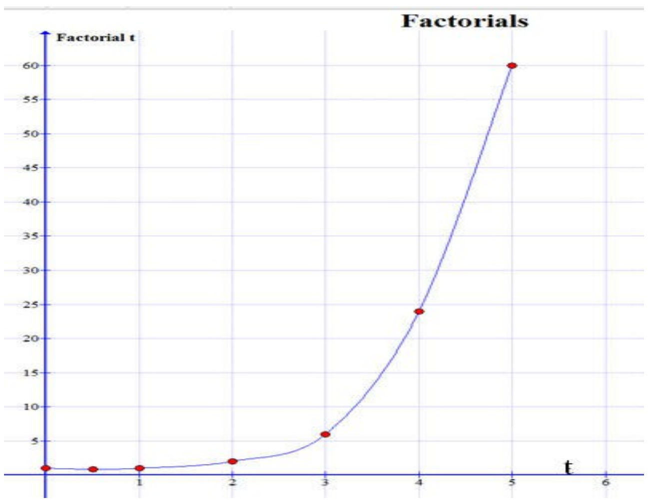
I. Donadello, L. Serafini and A. S. d'Avila Garcez. Logic Tensor Networks for Semantic Image Interpretation. In Proc. IJCAI'17, Melbourne, Australia, Aug 2017.

Deep Learning vs. Symbolic ML

- We want good classification and prediction but also useful descriptions… e.g.: learning factorial (*n!*)

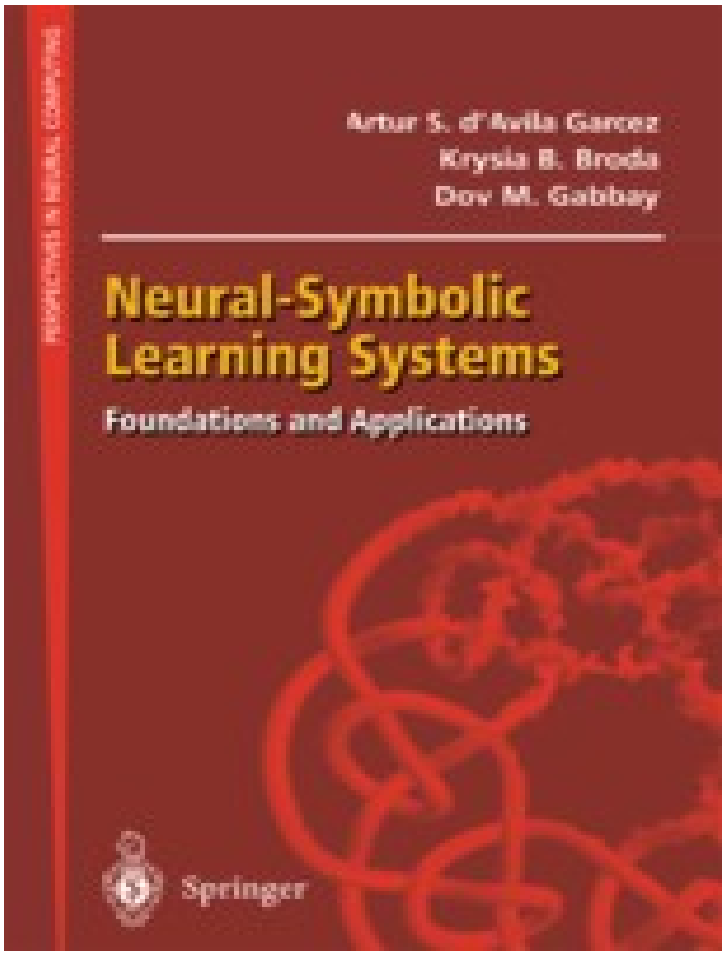
$$
\begin{aligned}
0! &= 1\\
1! &= 1\\
2! &= 1 \cdot 2 = 2\\
3! &= 1 \cdot 2 \cdot 3 = 6\\
4! &= 1 \cdot 2 \cdot 3 \cdot 4 = 24\\
5! &= 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120\\
6! &= 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 = 720
\end{aligned}
$$

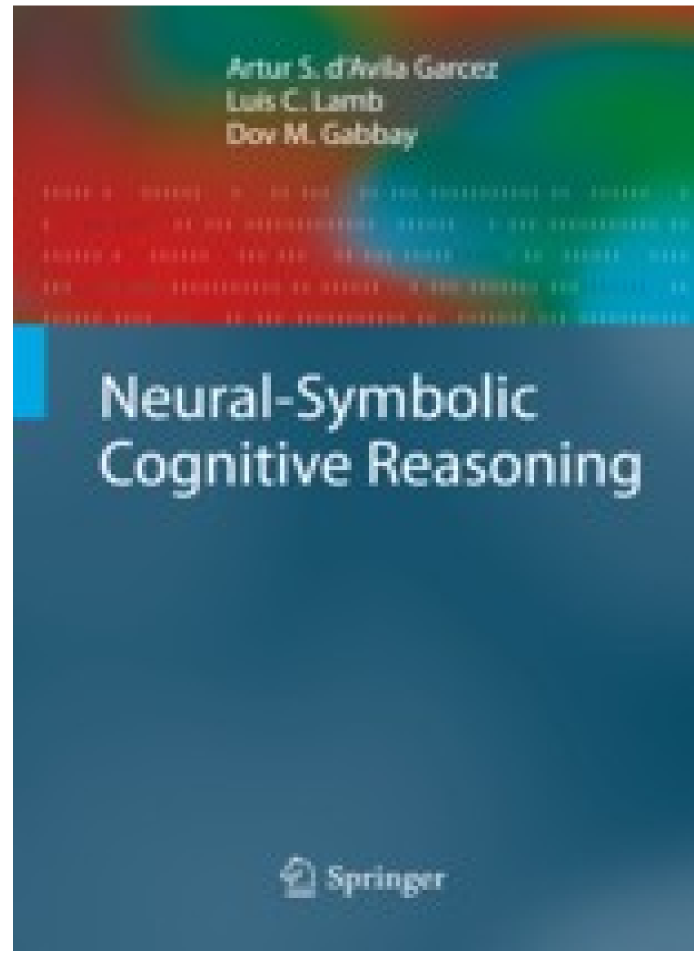





```
factorial(0,1).
factorial(N,F) :- N>0, N1 is N-1,
                  factorial(N1,F1),
                  F is N * F1.
```

For more information…





Conclusion: Why Neurons and Symbols

To study the statistical nature of learning and the logical nature of reasoning.

To provide a unifying foundation for robust learning and efficient reasoning.

To develop effective computational systems for AI applications.

Thank you!

Throw away your paradigm...

neurons

symbols

The future is neural-symbolic

Murray Shanahan