“Easy” presentation

The logic route to strong AI

YKY

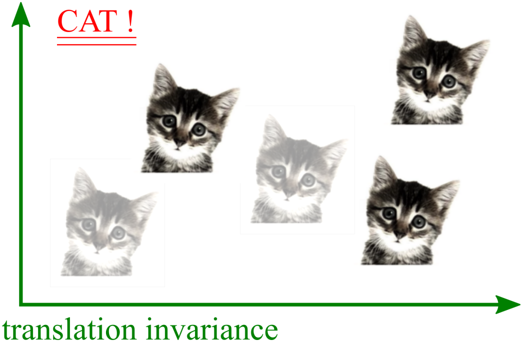
March 21, 2021

Table of contents

1 The success of CNN in computer vision
3 Symmetry and inductive bias
4 Richard Sutton's view
6 Doubts about logicism
7 Reinforcement learning
8 Structure of logic
9 Symmetric neural networks
11 For example: symmetric NN for object recognition
12 Logicalization of BERT
13 Advantages of logical AI (1)
14 Advantages of logical AI (2)
15 AGI

The success of CNN in computer vision

- In geometry, vision is said to possess the property of **translation invariance**:



(1)

- **Convolution** is an operation invariant under translation:

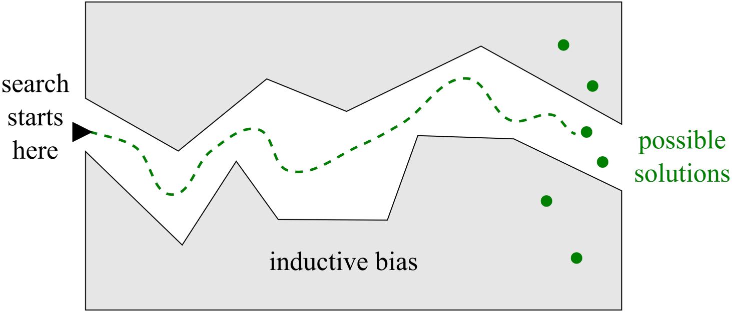
$$(T_x \circ f) * g = T_x \circ (f * g) \tag{2}$$

- Yann LeCun *et al* exploited the symmetry of CNNs to accelerate learning, successfully solved the visual recognition problem

(3)

Symmetry and inductive bias

- In mathematics, symmetry often simplifies computation, which is why mathematicians love to study symmetries
- In machine learning, one introduces **inductive bias** to narrow down the **search space**:



(4)

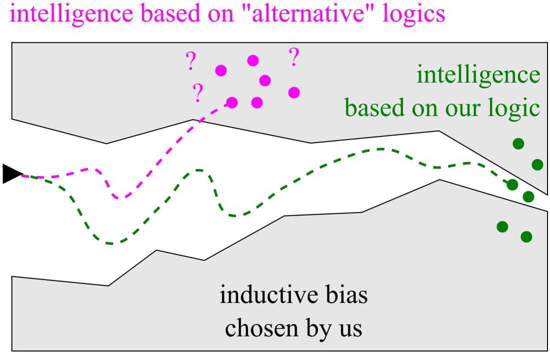
- Oftentimes, if inductive bias is chosen correctly, solution is found quickly, otherwise problem becomes **intractable**

Richard Sutton's view

- In contrast, Sutton expressed the view that AI can be solved merely by **increasing computing power**, under the reinforcement learning framework

(5)

- Our choice is just one out of many possible **forms** of logic:



(6)

- This is not only a theoretical issue;
  Indeed, AI labs around the world had begun the search for AGI with various strategies!

Doubts about logicism

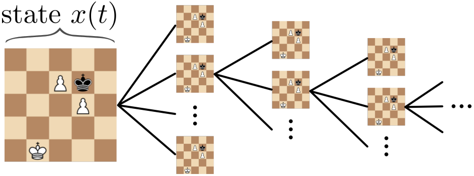
- Admittedly, the human brain does **not** think like symbolic logic
- By symbolic logic, we mean any "linear" (sequential) representation
- In contrast, **pictures** and **music** are *multi*-dimensional
- The brain may use representations that are not sequential or symbolic
- However, logic is a very efficient representation (from a computer perspective) even though it deviates from the biological brain
- Neuroscience is too difficult to crack, thus not a practical route to AGI
- Human cognition may be much closer to logic than we've thought. If we're given a description: "woman finds new lover, murders husband":

(7)

We may not know: Is the woman blonde? What is she wearing? In other words, our "mental model" is devoid of details and it may not possess any more information than just a sequence of symbols.

Reinforcement learning

- Think of the "**state**" in reinforcement learning like a **board positon** in a chess game:



(8)

- Reinforcement learning seeks to maximize the total **rewards** accrued over a (possibly infinite) **time horizon**:

$$\text{maximize } S = \int_0^\infty L \, dt \tag{9}$$

- Such maximization gives the AI **intelligence** because it is often beneficial to **delay** rewards, eg: to plot a clever chess move
- But reinforcement learning is a **brute force** approach; we need to give the model some additional **inductive bias**, eg, in the form of **logical structure**

Structure of logic

- The idea is: introduce symmetries of **logic** into deep learning to solve the AGI problem
- Because human cognition has logical structure, this inductive bias may help us find a solution to AGI faster
- Logic is a complicated structure, but its simplest symmetry is the **commutativity** (or permutation invariance) of **propositions**:

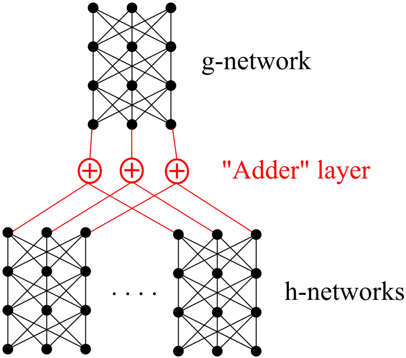
$$\begin{array}{ccc} A \wedge B & \equiv & B \wedge A \\ \text{it's raining} \wedge \text{lovesick} & \equiv & \text{lovesick} \wedge \text{it's raining} \end{array} \tag{10}$$

- Its importance may be analogous to translation invariance in vision
- The significance of commutativity is: it **decomposes** the AI system's **mental state** into individual **propositions**

Symmetric neural networks

- Permutation invariance can be handled by **symmetric** neural networks
- I wasted 2 years trying to solve this problem, and then found out it had been solved 3 years before: [PointNet 2017] and [DeepSets 2017] and their mastery of mathematics is significantly above me!
- Any symmetric function can be represented by the following form (a special case of the Kolmogorov-Arnold representation of functions):

$$f(x, y, ...) = g(h(x) + h(y) + ...) \tag{11}$$



(12)

- Sym NN gives a powerful boost in efficiency $\propto n!$ where $n = \#\text{inputs}$
- The code for Sym NN is just a few lines of Tensorflow:

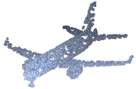
```
h = Dense(3, activation='tanh')
ys = []
for i in range(9):
    ys.append( h(xs[i]) )
y = Keras.stack(ys, axis=1)
Adder = Lambda(lambda x: Keras.sum(x, axis=1))
y = Adder(y)
g = Dense(3)
output = g(y)
```

(13)

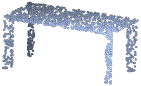
- Very easy to adopt this to existing models such as BERT and reinforcement learning
- I have successfully tested it on the game of TicTacToe: https://github.com/Cybernetic1/policy-gradient

For example: symmetric NN for object recognition

- Imagine objects represented as **point clouds**:



$\xmapsto{f}$ "aeroplane"

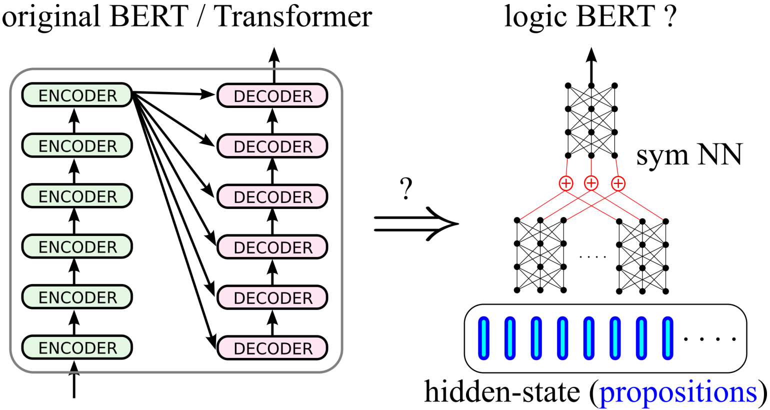


$\xmapsto{f}$ "desk"

- It does not matter **in what order** the points are in a sequence; the function $f(x_1, ..., x_n)$ is symmetric in its arguments (the points)
- Permutation invariance is **essential** for this to work

Logicalization of BERT

- Similarly, we can convert BERT's hidden state into a set of propositions, by replacing the original **decoder** with a sym NN:

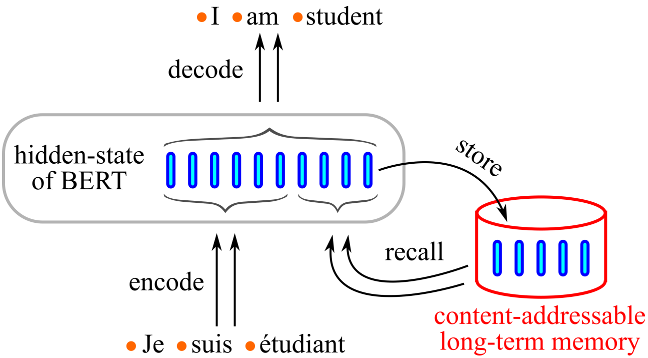


(14)

- The original **encoder** can be retained. As the **decoder** imposes symmetry on the hidden state, error propagation is expected to cause its representation to change
- Of course, this remains to be proven by experiment ☺

Advantages of logical AI (1)

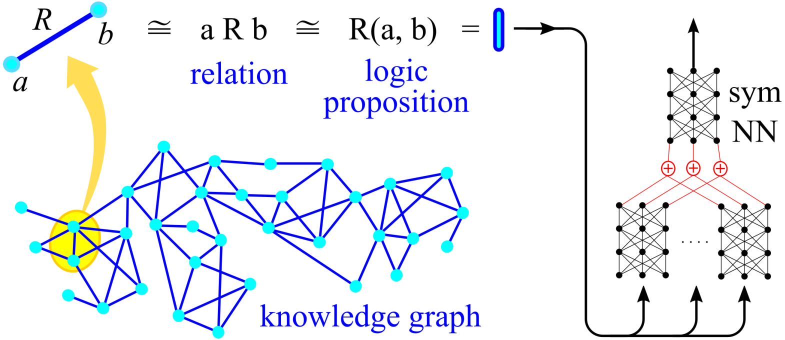
- With logic, it becomes easy to design cognitive architectures, eg: long-term memory module



(15)

Advantages of logical AI (2)

- Integrate seamlessly with **knowledge graphs**
- graphs are made up of edges,
  edges = relations between nodes = propositions:



(16)

References

Thanks for watching ☺️

Illustration credits:

- Translation invariance, from Udacity Course 730, Deep Learning (L3 Convolutional Neural Networks ▷ Convolutional Networks)