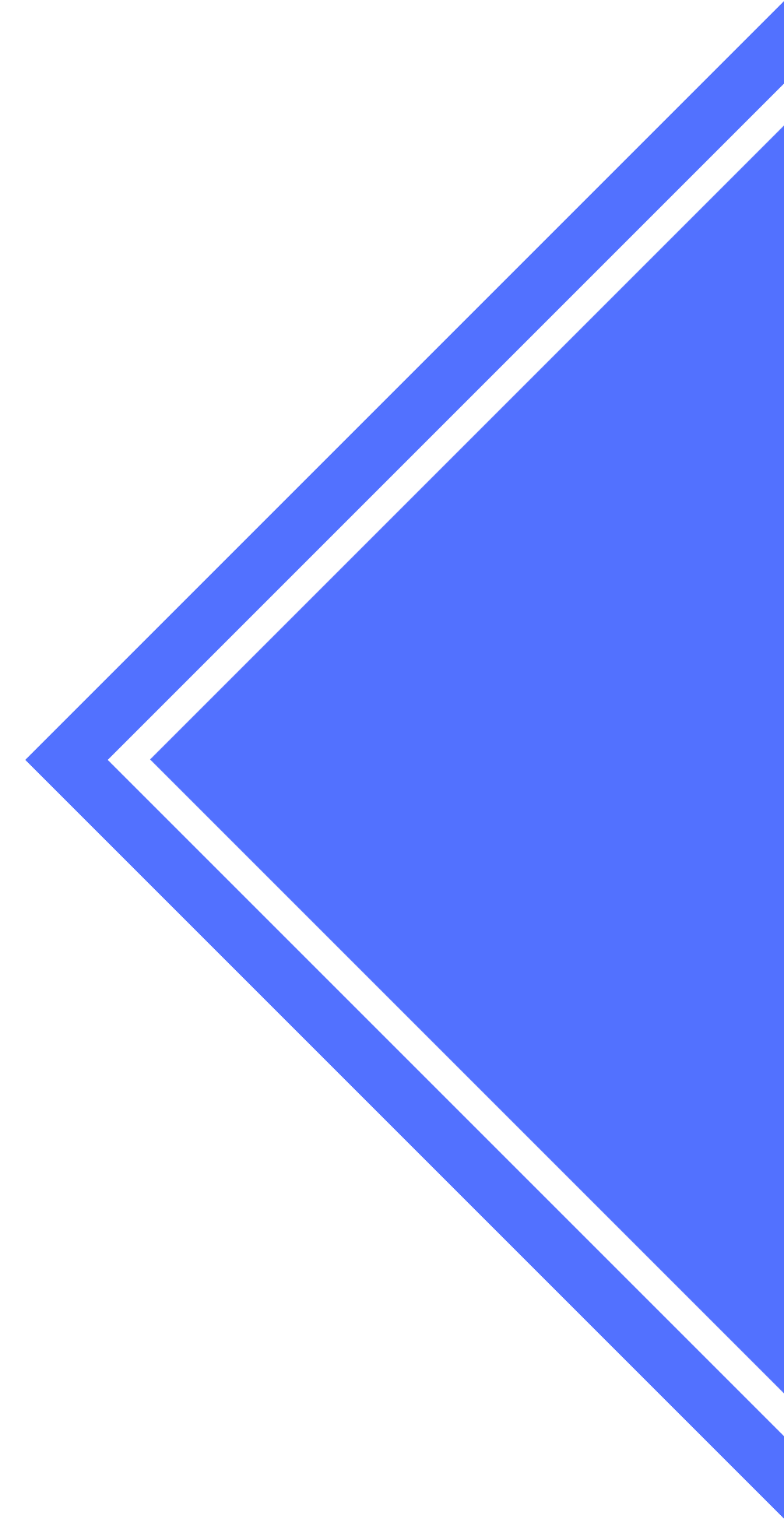


COLX523

PROJECT ARCOT

**Abstraction and Reasoning Corpus - using Chain of
Thoughts**



CURRENT AI LIMITATIONS

AI models require extensive data
for training
&
Limited to narrow tasks specifically
designed for.



Image generated by DALL.E

A NEW DIRECTION IN AI RESEARCH

The ARC Challenge

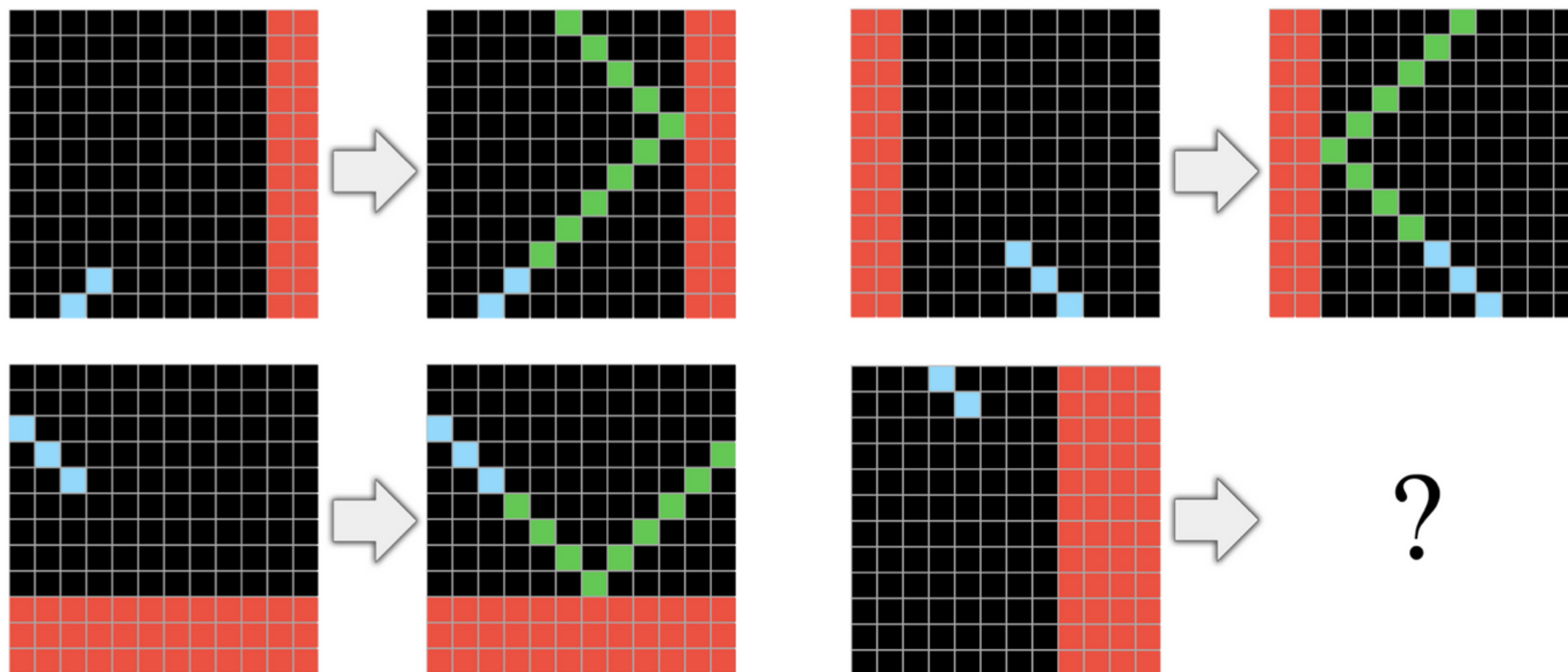
On the Measure of Intelligence

François Chollet *

Google, Inc.

fchollet@google.com

November 5, 2019



GPT4 FAILS AT ARC REASONING

GPT4 SOLVES 86% OF THE TESTS
WHEN GIVEN THE CORRECT THOUGHTS

Table 3. Tasks solved with iterative feedback loop after either in-
correct output or compile error

| Total | Incorrect Output | Compile Error |
|-------|------------------|---------------|
| 50 | 6 | 1 |

HOWEVER ONLY 45% OF THE
THOUGHTS ARE CORRECT

| Total Tasks | Tasks Solved | Tasks Not Solved | Tasks Partially Solved |
|-------------|--------------|------------------|------------------------|
| 111 | 50 | 58 | 3 |

THOUGHT CLONING DATA

Thought Cloning: Learning to Think while Acting by Imitating Human Thinking

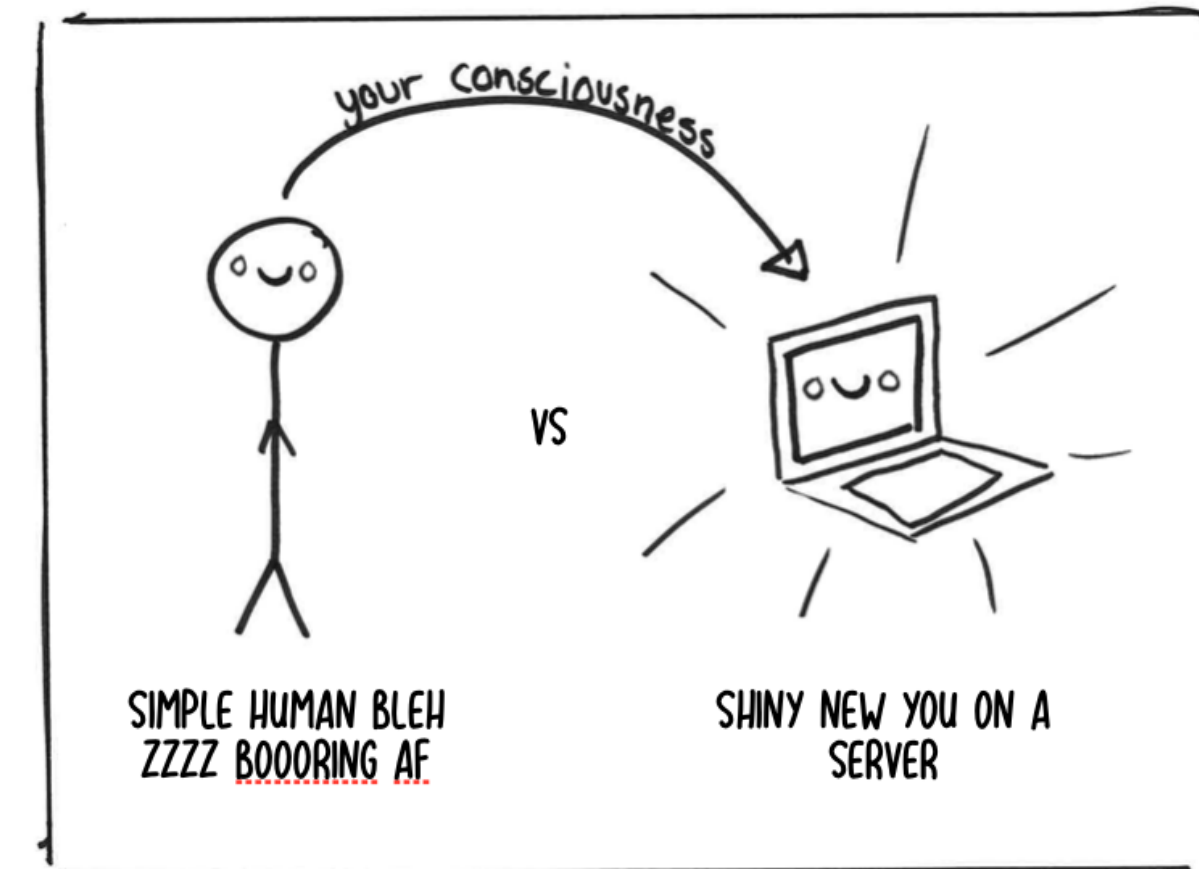
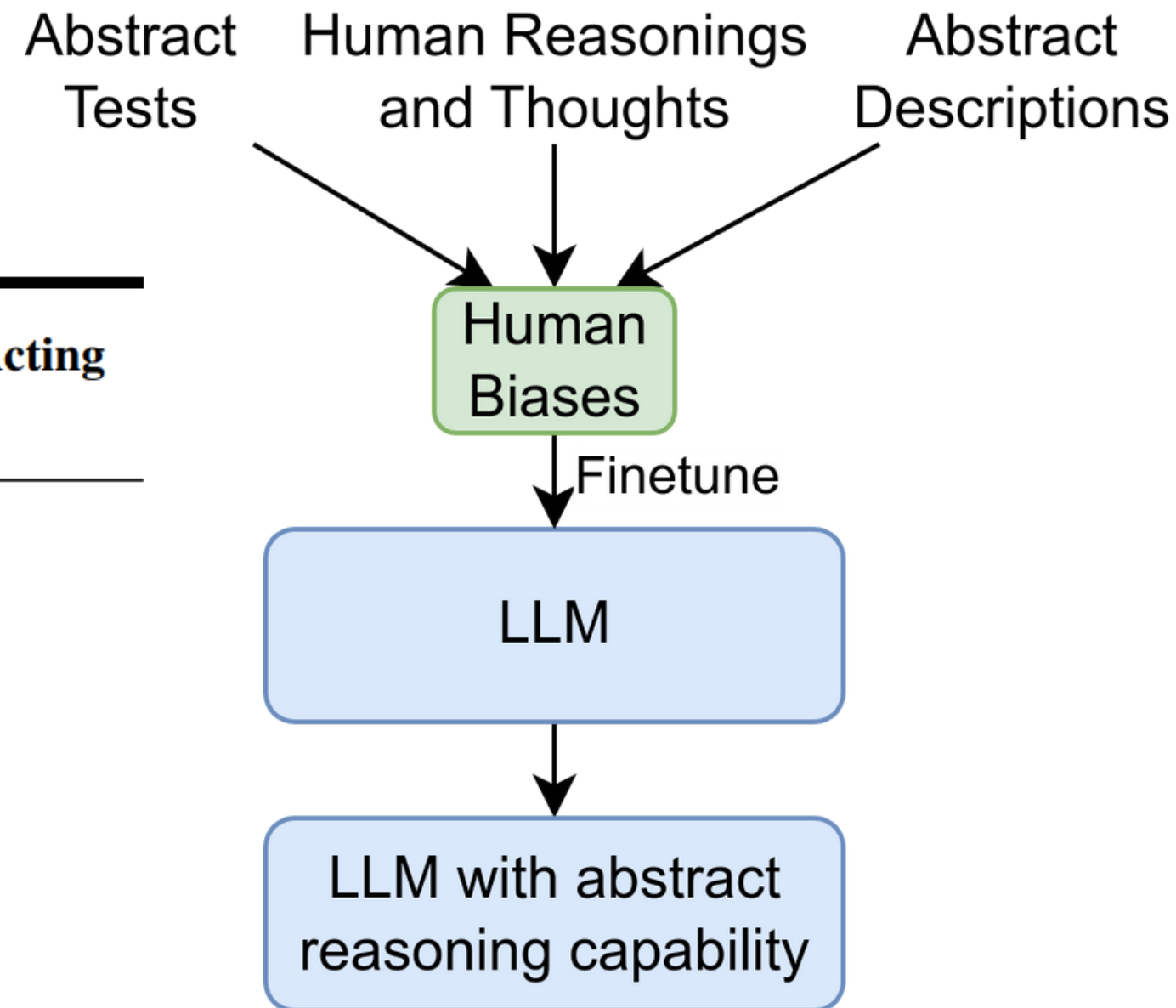
Shengran Hu^{1,2}
srhu@cs.ubc.ca

Jeff Clune^{1,2,3}
jclune@gmail.com

¹Department of Computer Science, University of British Columbia

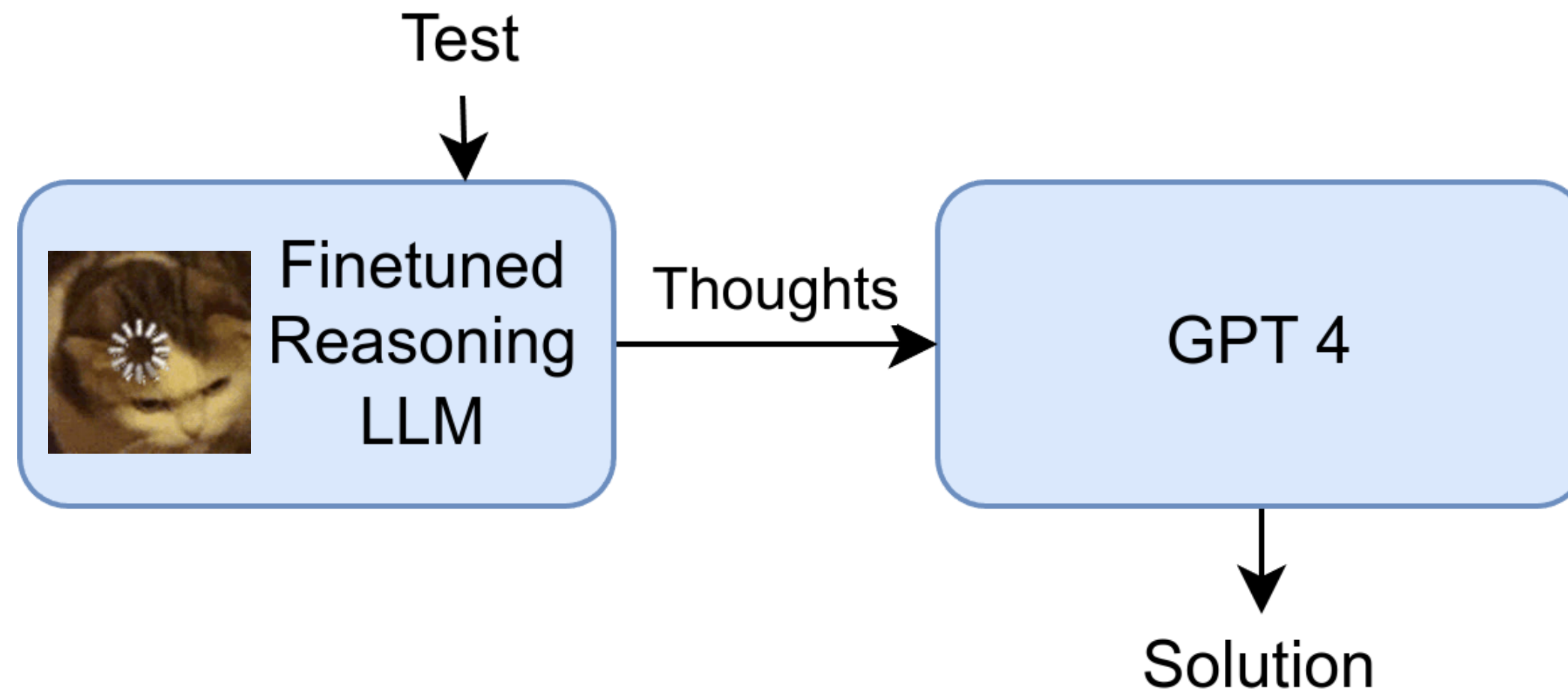
²Vector Institute

³Canada CIFAR AI Chair

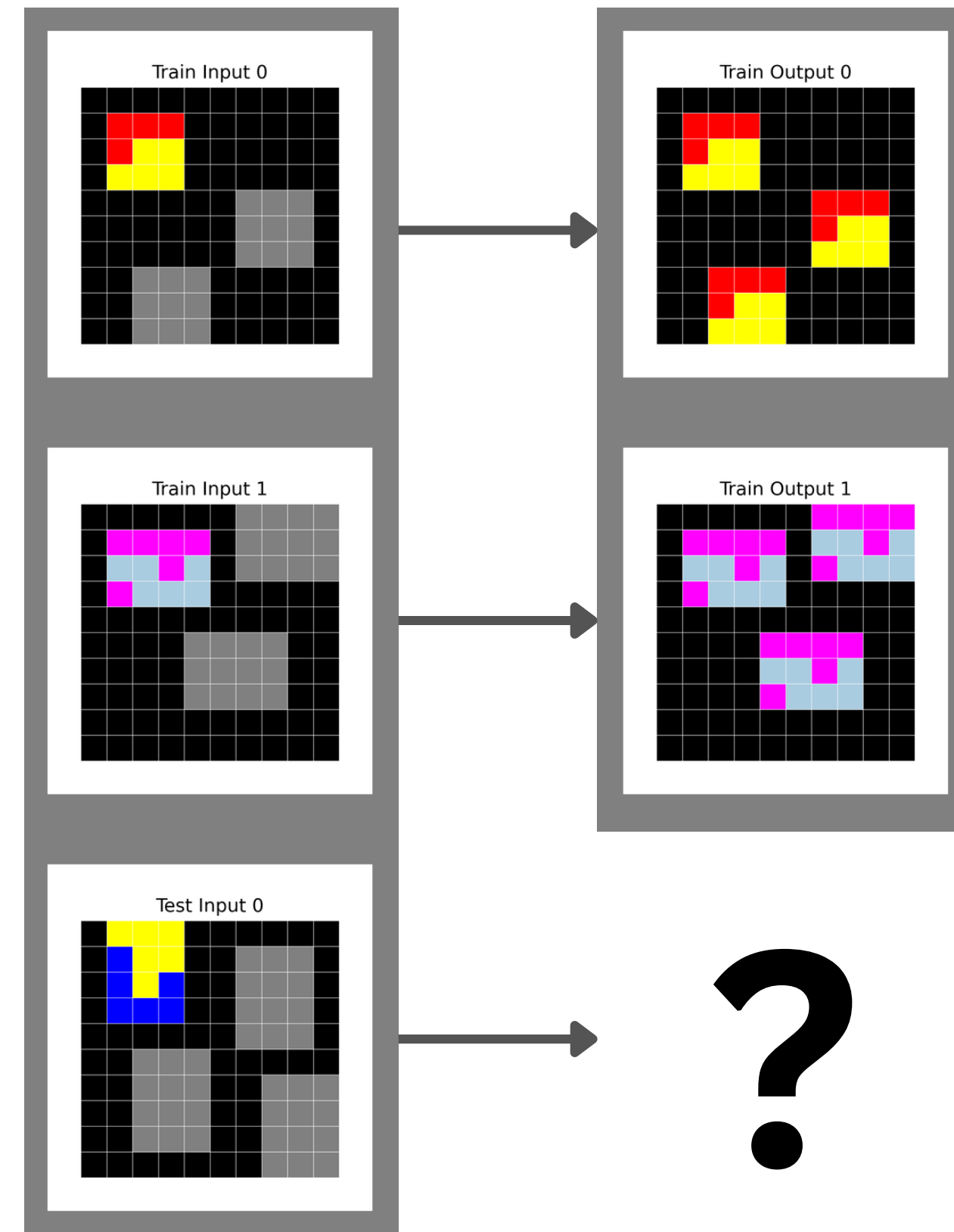


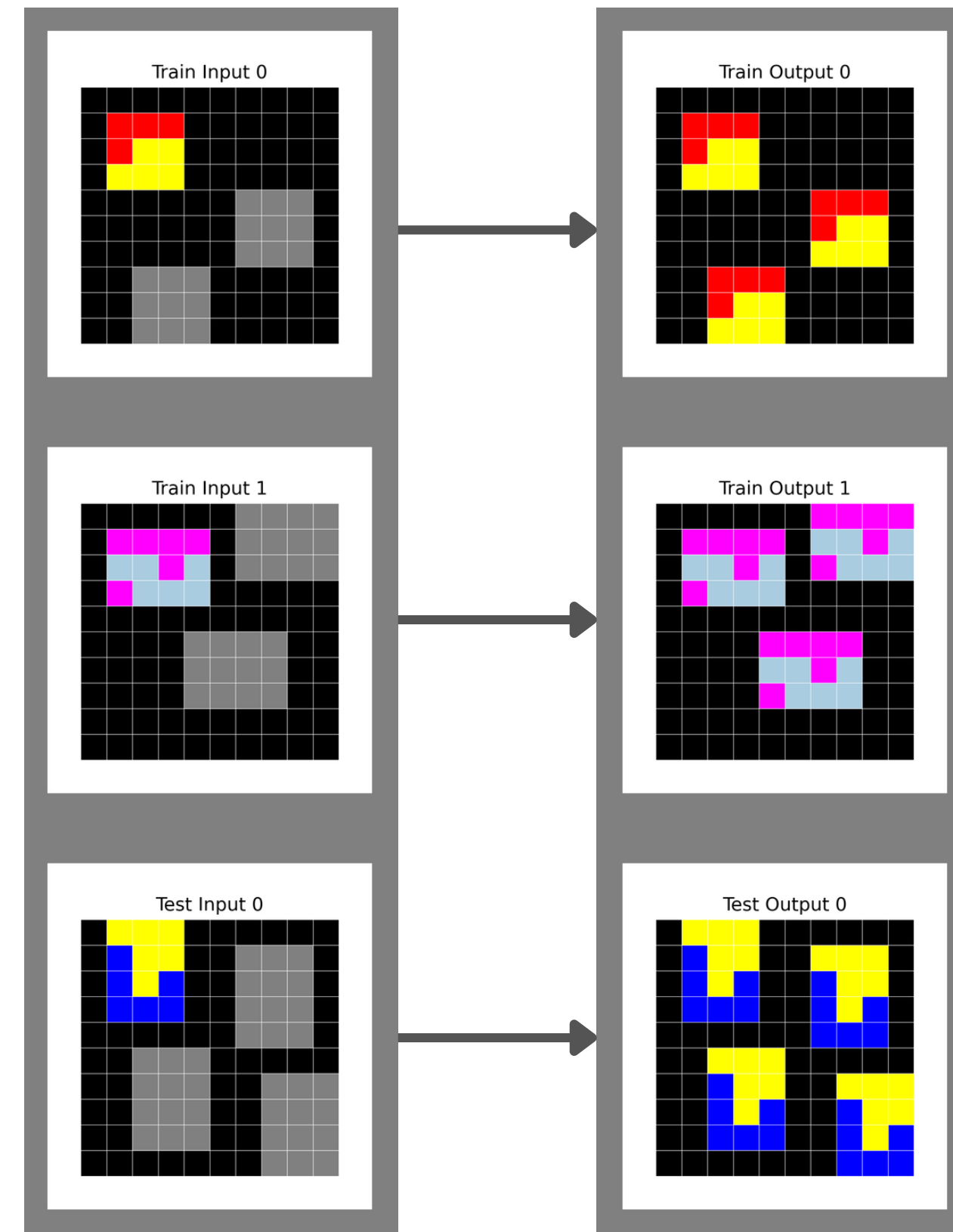
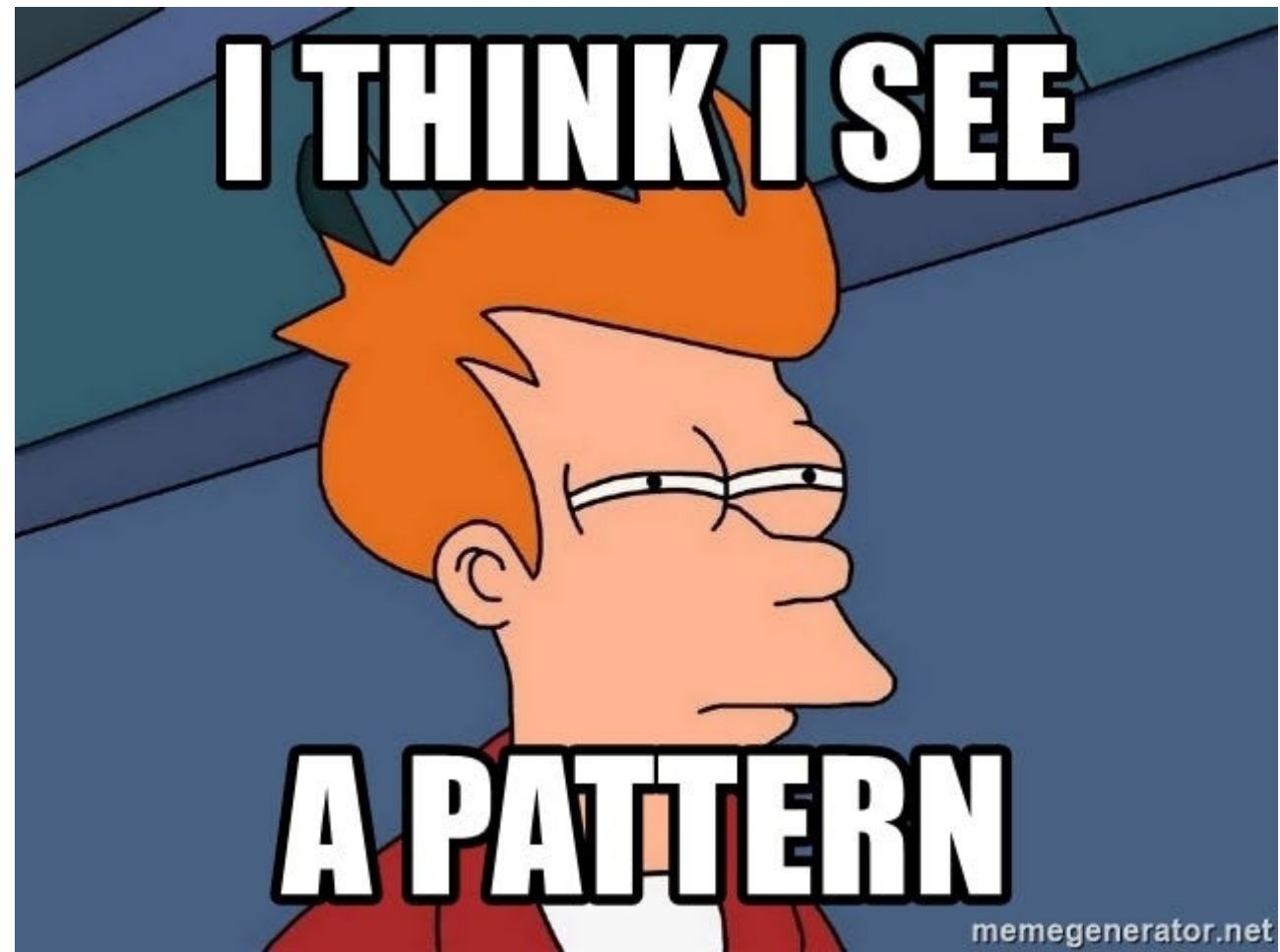
REASONING EXPERT AGENT

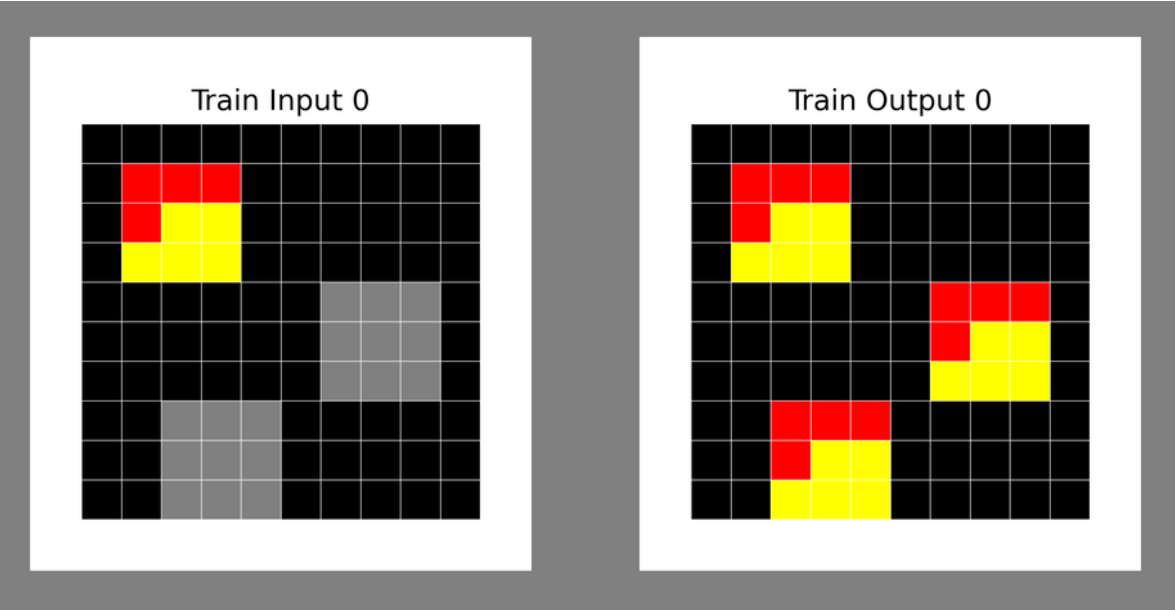
WE CAN COMBINE GPT4 WITH AN EXPERT THAT FOCUSES ON REASONING.



Example!







| Reflection | Pixel/Object Changes |
|--|---|
| Identify colored box and fill the grey ones with it. | Grey boxes are replaced with the colored box in each input grid, transforming the grid into a pattern of colored squares where grey squares previously existed. |

WHEN YOU LOSE YOUR TRAIN OF THOUGHT MIDSSENTENCE



Standard Prompting

Model Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The answer is 27. ❌

Chain-of-Thought Prompting

Model Input

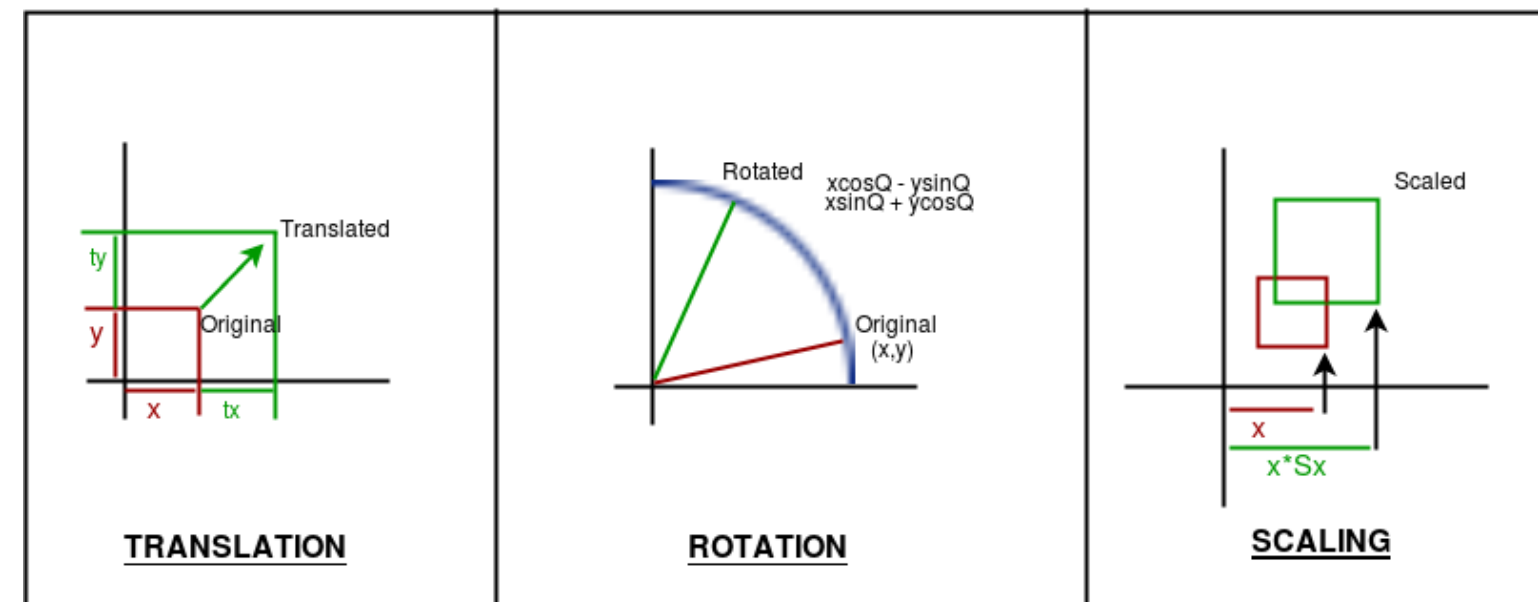
Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

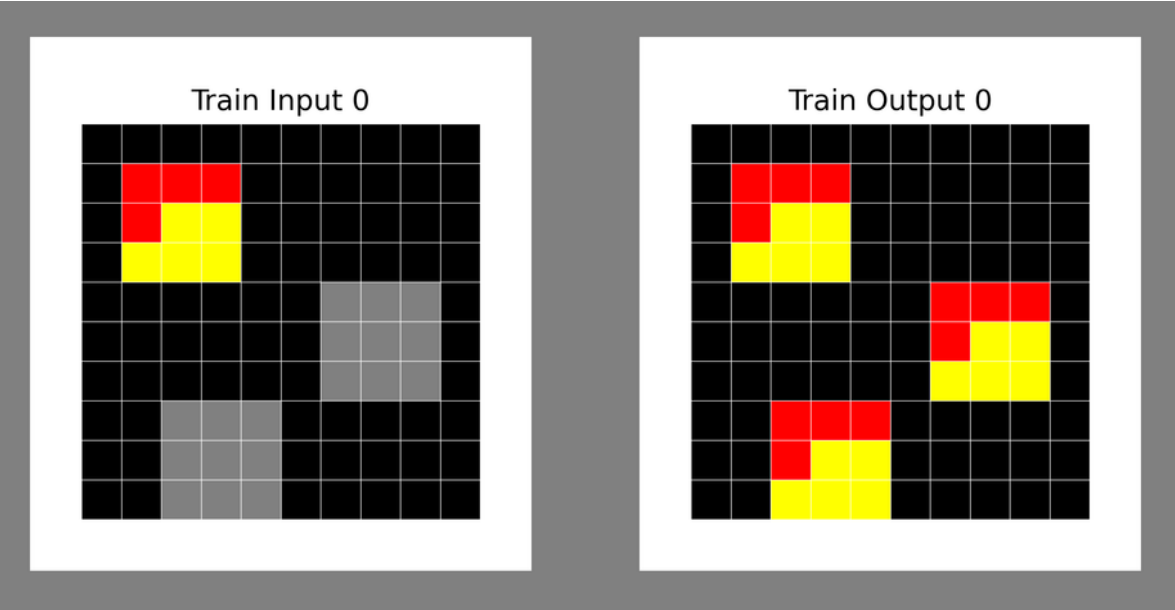
A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The cafeteria had 23 apples originally. They used 20 to make lunch. So they had $23 - 20 = 3$. They bought 6 more apples, so they have $3 + 6 = 9$. The answer is 9. ✅





| Reflection | Pixel/Object Changes | Helper Functions | Program Instructions |
|--|---|--|---|
| Identify colored box and fill the grey ones with it. | Grey boxes are replaced with the colored box in each input grid, transforming the grid into a pattern of colored squares where grey squares previously existed. | <i>get_pixel_coords</i> : To identify the coordinates of the colored and grey pixels. <i>get_object_color</i> : To obtain the color of the colored square. <i>change_object_color</i> : To change the color of grey squares. | <ol style="list-style-type: none">1. Use <code>`get_pixel_coords`</code> to locate the coordinates of the colored square(s).2. Identify the color of the square(s) using <code>`get_object_color`</code>.3. Use <code>`get_pixel_coords`</code> again to locate all grey squares.4. Fill all grey squares with the identified color from step 2 using <code>`change_object_color`</code>, and return the grid. |

WEB DEVELOPMENT USING STREAMLIT

- RAPID PROTOTYPING AND DEVELOPMENT
- USER-CENTRIC DESIGN AND EASE OF USE
- PYTHON ECOSYSTEM SUPPORT



USING THE APPLICATION

- HOME
- ABOUT
- CORPUS

Home

About

Corpus

ARCOT

Welcome to the ARC corpus integrated with the chain-of-thoughts paradigm!

Select search type:

☒ Search the context ☐ Search the helper function

E.g. 'Tasks that involve changing the size of the objects' or 'function1, function2' (both included)

Search the corpus

Search

Thank you

