

RPM Milestone 1 Journal

Xinghao Chen

xchen785@gatech.edu

Abstract—This document preliminarily discusses the methods to program an AI agent to solve the Raven’s test. It is suggested how humans, based on their prior knowledge, detect patterns, understand problems and reason through the test. Possible challenges and plans of implementing the AI agent are documented.

1 HUMAN’S SUPERIORITY OF PRIOR KNOWLEDGE AND PARALLEL SEARCHING

To implement a human-like AI agent, we may first analyze how humans may utilize the superiority of biological intelligence. From my perspective, humans, compared with computers, are extremely good at pattern recognition, both visually and conceptually, and the prior knowledge about geometric and logical operations. To implement an AI agent, we should actually teach our knowledge, and enable the agent to mimic humans’ computation efficiently.

Similar to what I have learned in class about the cognition system, I think humans do act with the deliberation system consisting of reasoning, learning and memory. However, these functional processes must be established along with humans’ firm knowledge about the latent features of shapes, transforms and even some common logical operations. Meanwhile, it seems to me that humans are extremely good at pattern recognition. We, by the light of nature, are able to extract and distinguish complex concepts quickly from images, without even being aware of doing so.

The following subsections will discuss the two foremost advantages of human brains, which are also likely to be the biggest challenges for computers.

1.1 Human’s Prior Knowledge

When we start to program our AI, it is natural for us to first parse the visual input to detect shapes and their latent patterns. However, patterns should always accord with human’s common-sense learned from earlier life. Therefore, it is likely that we had better initialize our AI with sufficient prior knowledge. We may also enable the AI to learn efficiently from the correct answers. Using my personal knowledge as a human, I have mined the following knowledge that I

have possible used in solving sample problems:

Geometric patterns:

- absolute position
- relative directions ($\uparrow \downarrow \leftarrow \rightarrow$ and clockwise/anti-clockwise)
- dots
- lines
- angles
- polygons (triangles, diamonds, squares, etc.)
- circles
- fillings

Geometric transforms:

- no change
- rotation
- translation (repositioning)
- symmetry
- scaling (can be non-uniform in different directions)
- filling
- changing to another shape
- deletion

Logic operations:

- AND
- OR
- NOT
- XOR
- complementary set
- difference set

Sample problems using the categories of knowledge stated above are depicted in Fig. 1.

Without a firm basis of knowledge, our AI agent can not even understand the input. However, teaching the agent requires human labor or sophisticated machine learning methods. Hence, giving human's knowledge to a computer will be a laborous challenge.

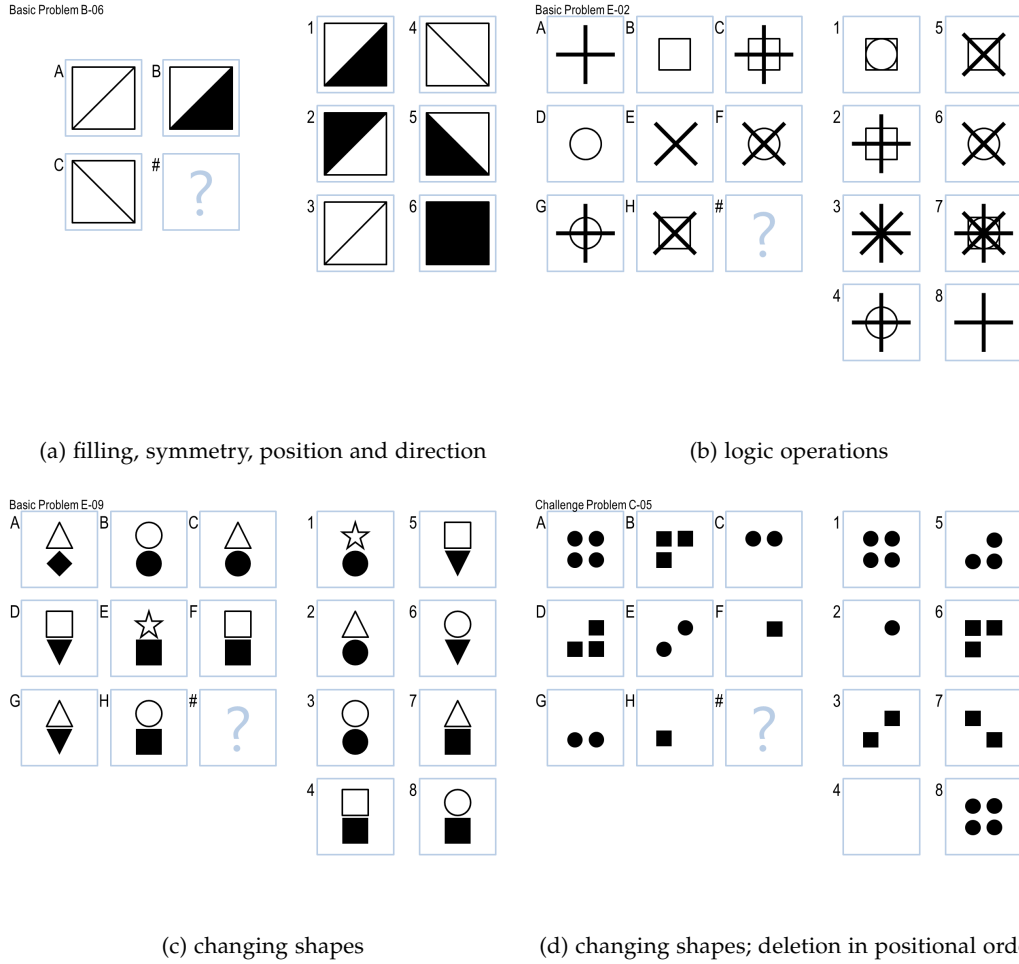


Figure 1—Samples of problems involving different knowledge background

1.2 Human's Parallel Searching and Reasoning

It seems humans are extremely good at launching many threads of parallel searching to recognize patterns. Besides, it is worth noting that human may also learn by matching patterns. Given the amazing parallelization power of human's brain, it is relatively harder for a computer to learn or search for patterns as efficiently as humans do if our AI agent computes with merely a CPU. Given the enormous number of possible patterns, we have to firstly distinguish the most possible ones.

Furtherly, our AI agent should return her answer in reasonable time. Thus we should stop searching when we find an answer that is good enough, which,

requires the ability of reasoning. Reasoning, in my opinion, is to estimate the possibility of the answer to be correct. Based on the pattern we have found and the prior knowledge of the probability of patterns, we should judge how plausible or how ridiculous our intermediate steps are, and decide whether our current answer is good enough. With a firm reasoning module, we can probably rule out the wrong answers, and establish the correct one immediately when we find it.

CPU-based computers are too slow in searching and not that reliable in reasoning (that is, judging whether the reasons and the answers are plausible). This is the second challenge for us to implement the AI agent.

2 EXPECTED MODULES OF THE AI AGENT

The design of my agent is expected to be of several layers of functions. Human's prior knowledge is latently distributed inside each layer of functions. Machine-learning methods will not be inside the plan.

2.1 Geometric Pattern Detection and Recognition

Our AI must be equipped with the ability to find shapes and confirm their attributes, because without being able to detect and describe the shapes in a single image, we cannot even begin to find relationships between images. I do not really have a clear idea about this module, and I may refer to some specialized codes online. I will probably store each shape as a node, with its attributes attached.

2.2 Relationship Searching and Reasoning

This module is to test possible transforms between images. I may start from the most possible transforms, and continue with the less-likely hypotheses. Heuristic methods might be used to deeply search the most possible case. The reasoning module is also necessary to decide whether the input well accords with a pattern, and whether a pattern is too ridiculous to hold.

2.3 Strategies

I may first ask the agent to raise possible answers by itself. If the answer is not available as a choice, or if I find it difficult to form the answer or compare the agent's answer with existing choices, I may have to test each choice.