RPM Final Journal

James Sherman
jcsjr0014@gatech.edu

***Abstract*—**This paper summarizes the approach my AI-agent uses to solve Raven's Progressive Matrices, a form of intelligence test. Its performance is then summarized, exploring the types of problems the agent performs correctly and struggles to solve. Finally, the intelligence practiced by this agent is compared to the intelligence a human may practice when solving these problems.

1 AGENT'S APPROACH

My agent is a generate and test agent that uses means-end analysis to select the best answer choice to solve the Raven's Progressive Matrices (RPMs) such as the one depicted in Figure 1 below. My agent first generates a frame representation of the shapes within each figure (A, B, C, D, E, F, G, H, and the answer choices as shown in Figure 1). Then the agent generates a frame representation of the relationships between each figure. The relationships are scored using a large set of heuristics to test which answer choice most accurately matches the relationships of a particular matrix using means-end analysis. The agent works slightly differently depending on if the problem is a 2×2 matrix or a 3×3 matrix. The following subsections describe the agent's approach to 3×3 matrix problems and conclude with how the approach varies to its approach to 2×2 matrix problems.

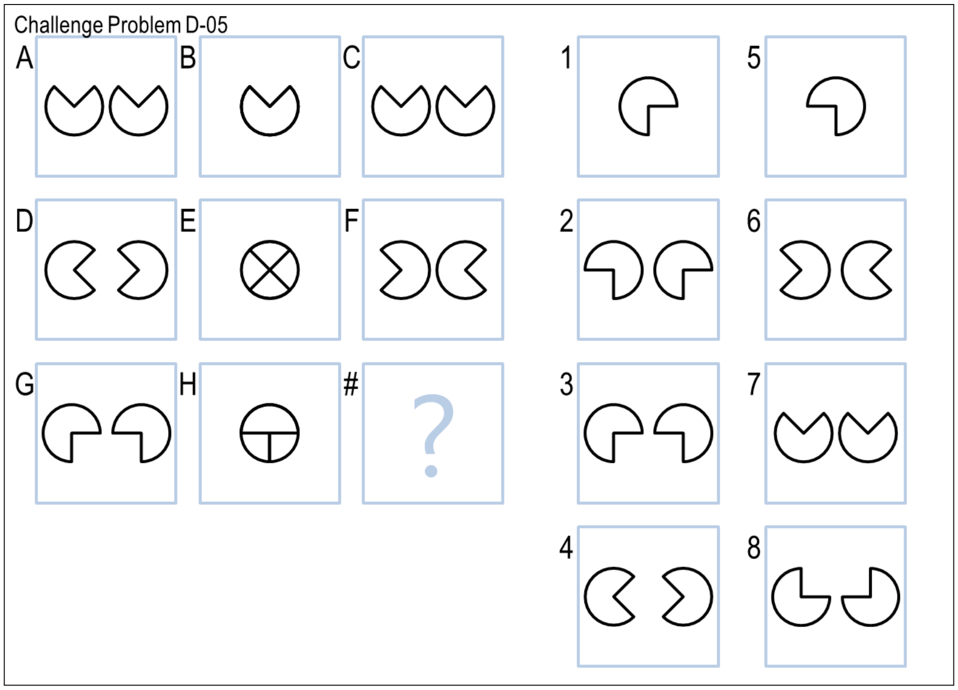

Figure 1—An example of a 3×3 Raven's Progressive Matrix problem

1.1 Generating frame representation of figures

For a given RPM problem, the agent begins by creating a frame representation of all the shapes in each figure. The OpenCV library is used to locate each shape in a particular figure. That shape is then associated with the coordinates of the shape's corners and whether the shape is "hollow" or not (if a shape has black pixels filled inside of it or not). The agent also uses the OpenCV library to capture the grayscale matrix of each figure. This matrix is adjusted such that each black pixel in each figure is a 0 and each white pixel is a 1.

1.2 Generating frame representation of relationships between figures

The frames generated for each figure are used to create frames for relationships between figures to represent the pattern change between each figure. The relationships found are listed below.

- **Shape relationships**: the agent matches the shapes between two figures and marks if the shape is unchanged or has been reflected across an axis. These relationships are determined using the coordinates of the corners. The agent also determines the total number of shapes and corners added or lost between figures.

- **Grayscale matrix MSE**: the agent uses means-squared error (MSE) to determine how similar two figures' normalized grayscale matrices are. The MSE is modified to create a scale between 0 and 1 such that two images that are exactly the same would receive a score of 1, and images that are increasingly different would have lower scores. In 3×3 matrix problems, each row (i.e figures ABC in Figure 1) and column (i.e figures ADG in Figure 1) would receive a ratio of their MSE scores by dividing the MSE of the first two figures within a row or column by the MSE of the last two figures within the same row or column.
- **Corner coordinate ranges**: the agent records the maximum and minimum x and y values listed in all the corner coordinates of the shapes within a figure. The difference between the minimum x and y values between each figure is recorded. A ratio is then determined for each row and column similar to the ratio described above but using the differences in the x and y values.

As hinted above, the relationships between rows and columns are recorded. Using the row ABC for example, figures A and B were compared, and figures B and C were compared. Whole row/column relationship is determined using the ratios described above. The agent also generates a frame that represents the relationship along the diagonal (AE answer choice).

1.3 Testing for best answer choice using means-end analysis

We now have frame representations of how each row and column relate to one another, including for the row GH and each answer choice and the column CF and each answer choice. The relationships of rows ABC and DEF are separately compared to the relationships of rows GH and each answer choice. The agent uses similar logic to examine the relationship of the columns. Similar MSE ratios were weighted the heaviest, followed by similar maximum and minimum x and y value ratios, followed by matching shape relationships. The scoring method described in the subsection below was applied to the diagonal to score if its relationships were relevant. These heuristics give each answer choice a score that determines how likely an answer choice solves the matrix. The answer choice with the highest score is selected.

1.4 Difference in 2×2 matrix problems' approach

The agent follows a similar approach as the approach described above with a few differences. First, only the row AB is compared to C answer choice, and only the column AC is compared to B answer choice. Furthermore, the agent weighs shape relationships the greatest. The agent does not use an MSE ratio and just uses the adjusted MSE value to determine the pixel similarity between a row or column. Finally, the ranges of the x and y values are not considered in scoring these problems.

2 AGENT'S PERFORMANCE

Performance in this scenario is defined as how often the agent returns the correct answer choice. The agent's performance has room for improvement, as it answers 62 of the 96 ($\approx$ %65) Basic set questions correctly. Below is a table that represents the agents' performance on the different sets of problems. Note that this table excludes the Test set that was part of the score above.

Table 1—Agent's performance on different groups of problems.

Problem set	Problems correct	Total problems
Basic B	11	12
Basic C	10	12
Basic D	3	12
Basic E	8	12
Challenge B	4	12
Challenge C	7	12
Challenge D	2	12
Challenge E	5	12

2.1 Agent performance on particular problems

The following examples will show why the agent performs well on certain problems. Consider the 2×2 matrix depicted in Figure 2 below.

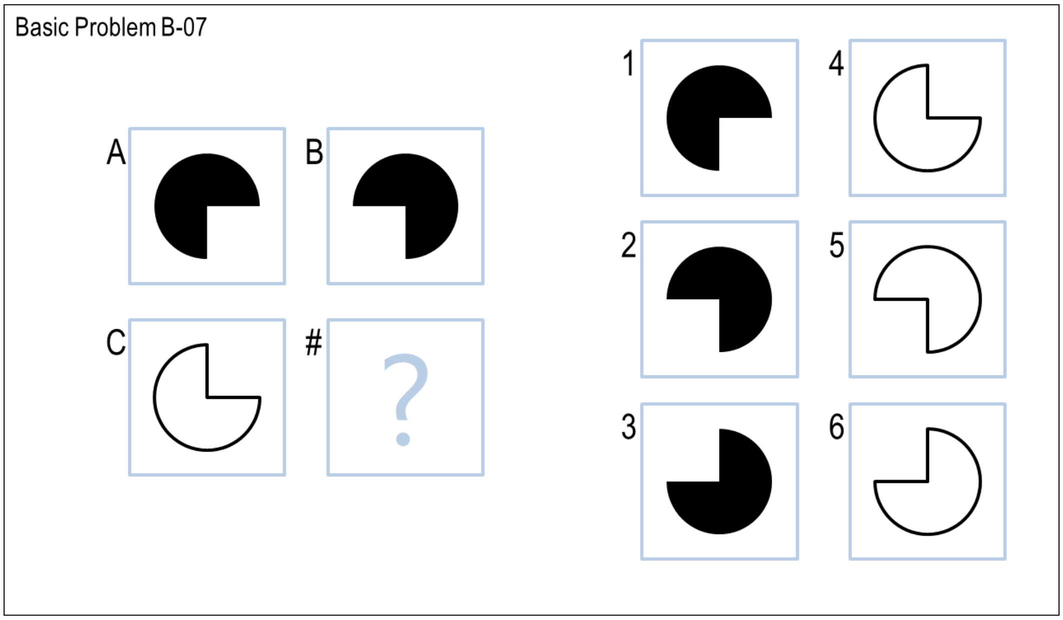

Figure 2—Basic Problem B-07

The agent is able to determine that the shape relationship between AB is a reflection. The MSE between AB implies that the shape does not change in hollowness. Answer choice 6 most accurately describes this relationship between C and an answer choice. The agent uses similar logic on the columns AC and B answer choice to further support that 6 is the correct answer. The agent easily handles problems like this where there is only one shape present in the figures. Let's now look at a more difficult 3×3 problem, shown in Figure 3 below.

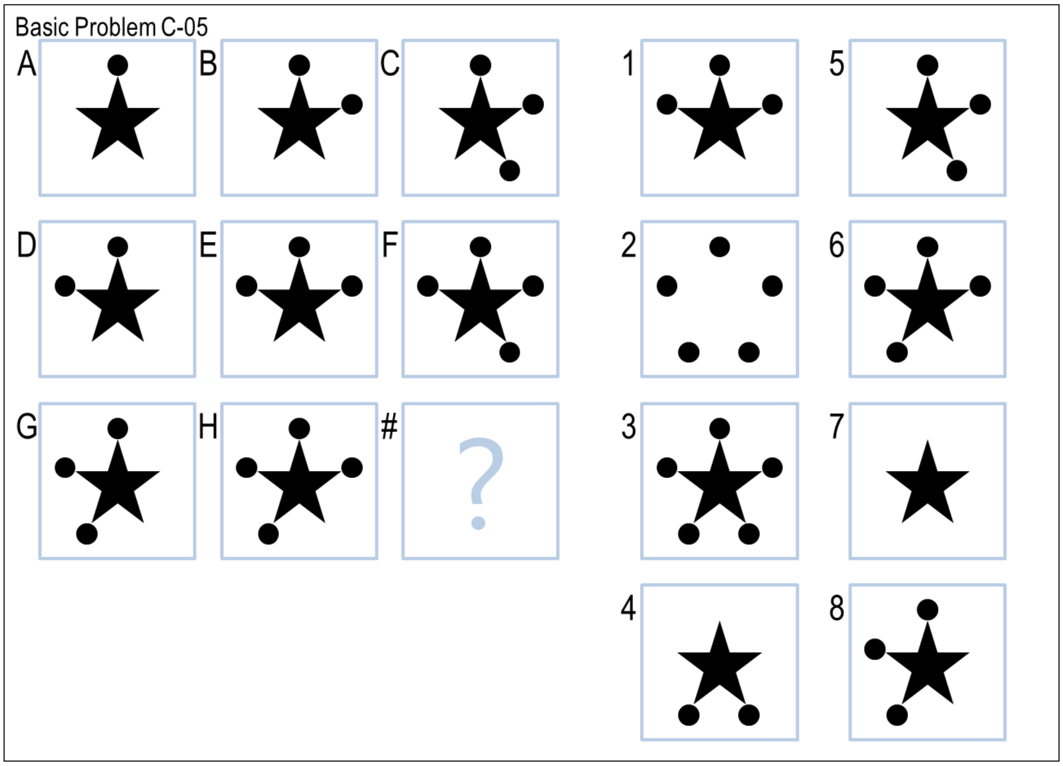

Figure 3—Basic Problem C-05

The agent is able to determine that one shape is being added from figure to figure within a row and column. Answer choice 3 is the only option that satisfies that relationship. Furthermore, the MSE ratio between a row or column are most similar when looking at answer choice 3. The agent is able to successfully select the correct answer. The agent can successfully solve such problems where shapes are being added/deleted at a consistent rate. Let's look at another 3×3 problem, shown in Figure 4 below.

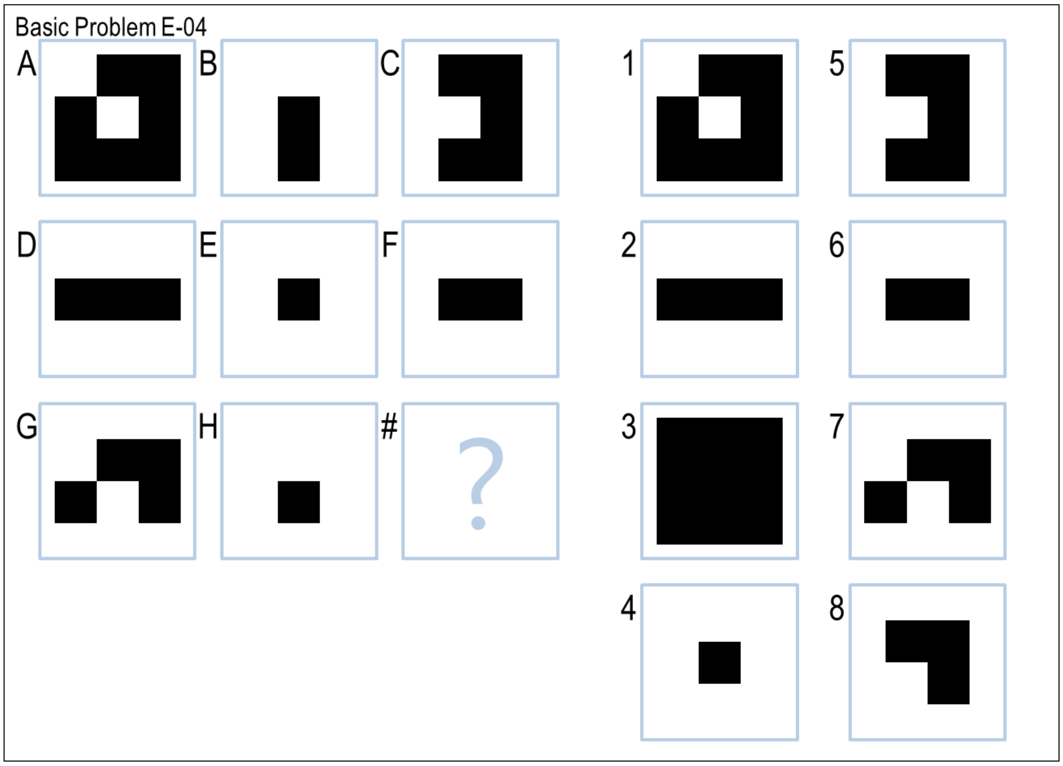

Figure 4—Basic Problem E-04

The agent struggles to determine a shape relationship for the rows and columns in this problem. However, the MSE scores imply that the shape goes from very large, to very small, to somewhere in the middle. The ratio of the ranges of the minimum and maximum x and y values further validate this argument. This leads to answer choice 8 being correctly selected. This problems shows how the MSE ratio and the minimum and maximum x and y value ratios can be used to solve problems where the figures in each row and column need to be considered as a whole and not just a transition from one figure to the next.

2.2 Agent struggles on particular problems

The agent tends to struggle on particular cases of problems. Consider the following problem depicted in Figure 5.

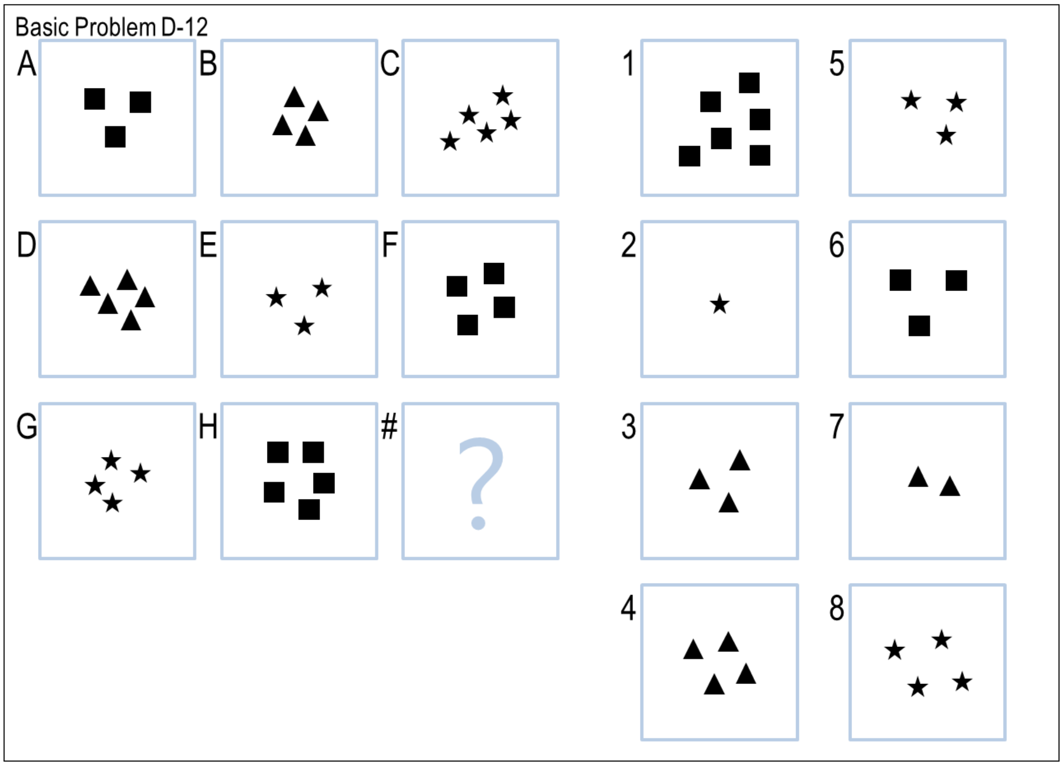

Figure 5—Basic Problem D-12

The MSE ratios and minimum and maximum x and y value ratios do not accurately describe the relationship in these types of problems. These problems would require the agent to consider all the rows and columns together. For example, each of the first two rows and columns has a figure with 3 shapes, a figure with 4 shapes, and a figure with 5 shapes. Using that logic, the answer choice should have 3 shapes. Similar logic would imply that the answer choice needs to be 3 triangles. However, the agent only considers rows and columns separately, which causes the agent to fail problems like these. Let's consider the problem in Figure 6 below.

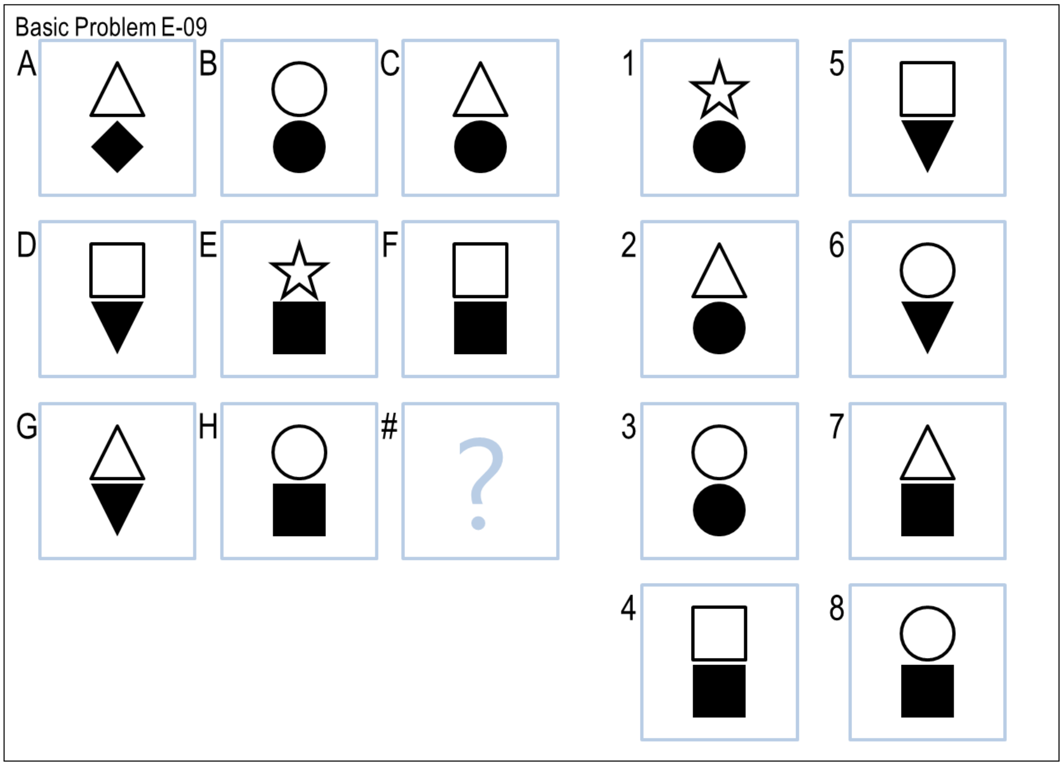

Figure 6—Basic Problem E-09

These problems require one to consider the shape relationship between the two ends of the rows and columns. For example, A is related to C in that the top shape of A is the top shape of C. This logic is true for the other row as well as the columns. This agent fails to make that connection, causing the agent to fail this type of problem.

3 AGENT'S APPROACH IN COMPARISON WITH HUMAN APPROACHES

3.1 Similarities

The agent is similar to a human in that it attempts to determine the relationship on each row and column and compares all the answer choices and weighs which choice most accurately describes the pattern in the matrix. This means-end analysis is often used by humans when trying to narrow down their answer choice to a smaller set of potential answers. The agent can also determine several basic relationships between figures that humans can determine, such as reflections, hollowings/fillings, and stretches in the x and y direction.

3.2 Differences

The agent goes about determining patterns very differently than humans. Obviously, humans are not recording corner coordinates or calculating the MSE of two grayscale matrices. Also, humans have a higher level of case-based reasoning to determine how to solve a problem. For example, consider the problem in Figure 5 above. A human would reason that pattern within a row is hard to determine. However, the pattern when considering all the rows at once is evident. My agent, however, attempts to score each row and column the same way for every problem. This leads to heurstics that should not be applied to certain problems getting applied to those problems. Humans are able to use reason to determine what combination of figures should be used to identify a pattern for a particular matrix.