RPM Milestone 1 Journal

Victor I. Sierra
vsierra3@gatech.edu

1 HUMAN APPROACH

If a human is asked to solve a Raven's Progressive Matrices (RPM) test, the first step that one might take will be to identify what type of category the problem falls under. Not all problems can be solved the same way. Next, once the category has been identified, one will try using a technique appropriate for that category. Finally, they'll compare the generated solution with the possible ones and see if it fits any of the possible answers. If there is not a match, it might be that the person needs to iterate and make use of a combination of categories.

As part of this exercise, a few categories have been identified. This should help identifying what type of problems they are and how they could be solved. Not all problems could use one technique. In some more advanced problems, a combination of two or more of these categories might be fit.

These are the categories identified:

1. Shape Transformation – Ex: Transform a rectangle to a circle
2. Shape Scaling – Ex: Change a small square to a large square
3. Shape Filling – Ex: Fill in a white shape with black
4. Shape Rotation – Ex: Rotate a shape by 90 degrees
5. Shape Relationships and Patterns – Ex: Every row has shapes A, B and C
6. Numeric Relationships – Ex: First row has two, four and six shapes
7. Logical Relationships – Ex: Shapes are transformed using XOR
8. Shape Equality – Ex: Shapes are the same

2 PROBLEM SELECTION

This section will go over four different problems from the Raven's Progressive Matrices. These problems will be explained in detailed using the human approach categories defined in the previous section.

2.1 Basic Problem B-09

This problem falls under the Shape Filling category. As it can be seen on Figure 1, the relationship between A and B is that the same shape is preserved. The only difference between these shapes is that it goes from white to black. Based on the relationship between A and B, it is possible that C will be transformed from a white square to a black square.

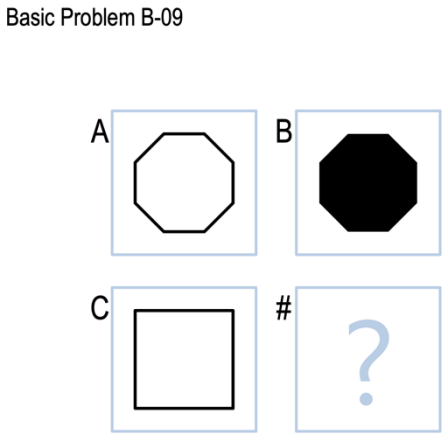

Figure 1 - RPM Basic Problem B-09

2.2 Challenge Problem C-06

This problem combines more than one category. In this case these are Shape Relationships, Numeric Relationships, and Shape Filling.

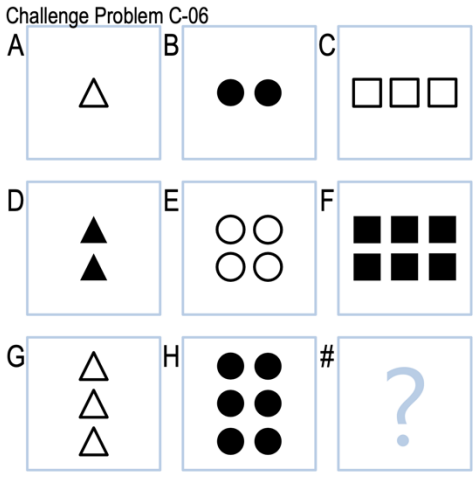

Figure 2 - RPM Challenge Problem C-06

The first category will be Shape Relationships. As it can be seen in Figure 2, the first row (A-C), starts with one triangle, two circles, then followed by three squares. The second row contains the same shapes, triangles, circles, and squares. On the third row, items G and H contain triangles and circles. If the same pattern is followed in this row, the missing item should be squares.

The second category is Numeric Relationships. After reviewing the first row, one will notice that it contains one, two, and three shapes. The second row contains two, four, and six shapes. At this point the numerical pattern seems to be that the first column contains (in number of shapes) the row number, followed by the row number times two, then the row number times three. Given this, the missing item will contain nine shapes.

The third and final category is around Shape Filling. All rows contain a mix of black and white shapes. Row one contains (in this order) white, black, and white. Row two contains black, white, and black. Given this pattern, it seems that it is fair to reach the conclusion that row three will have white, black, and white as its pattern. In this case, the missing shapes will have a white color.

Putting everything together for this more complex problem, the missing item should be nine white squares.

2.3 Basic Problem D-08

This problem can be solved by using the techniques from Shape Patterns/Relationships, and Shape Filling.

As it can be observed on Figure 3, The first row contains a square, a rhombus, and a circle. The second row contains the same shapes. The third row is missing a rhombus.

The second observation is that both rows 1, and 2 contain two black shapes and one white one. Given this, the third row is missing a black shape.

The third and final observation is that every row contains a black shape inside a slightly larger identical white shape. Rows one and two already contain a circle and a rhombus that match this criterion. The third row contains already a square that matches this criterion.

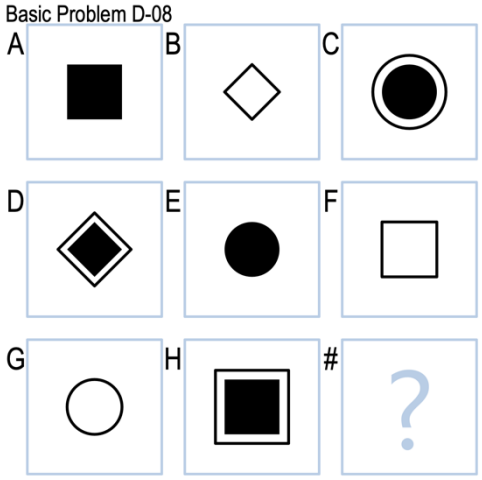

Figure 3 - RPM Basic Problem D-08

Given all these observations, the missing figure is a black rhombus.

2.4 Basic B-04

For this problem, the simplest approach will be to solve to use Shape Rotation. In this case, the same can be rotated 90 degrees between each figure and it will achieve the desired outcome.

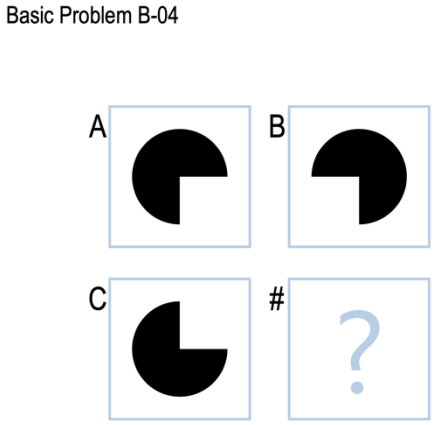

Figure 4 - RPM Basic Problem B-04

3 AGENT APPROACH

The most likely scenario for designing an agent that solves the RPM test will be a combination of multiple approaches.

3.1 Simple Problems

The first approach will be a pixel-based representation of the images. In this case, problems these are images where they are the same or are rotated. This could easily be solved by leveraging the number of black and white pixels in an image.

Once an agent has identified the number of black and white pixels in an image, it could use this information to determine if two images are the same or if they are rotated. No matter how many times an image is rotated, the image will have the same number of black and white pixels. If the images are the same, the number of black and white pixels will also be the same.

3.2 Complex Problems

For more advanced problems, the agent could use a combination of approaches.

First, it could start with building a verbal representation of the images. For each image, it will identify what are the shapes present and their position. The same way semantic networks are generated, a representation that is abstracted from the visuals could be generated.

While using semantic networks could still be the base for identifying these patterns to provide an abstraction, it is possible that some type of pixel-based interpretation could take place too. An example is, for filled shapes, it might make sense to use this type of approach, so that when the semantic network is built, this information can be incorporated to it.

Once the shape and their "filling" have been identified, the next step will be to iterate over the different techniques to produce potential solutions. Once this has happened, a similarity score should be generated and compared with the possible answers. Whatever is the highest score, that should be the winning one.

If the similarity score doesn't reach a specific threshold, a combination of techniques from the categories listed on Section 1 should be used. At this point the

similarity score will be generated, and the process will be repeated until a threshold, or a specific number of tries have been met.

3.3 More on Complex Problems

There is a possibility that the previous approach doesn't work, as an alternate solution, decomposing the images into smaller images might be an approach to consider. As an example, an image could be split into smaller images and these images could be analyzed taking into consideration the approaches previously discussed.

4 EXPECTED CHALLENGES

As part of this agent, there are a series of challenges that one can expect.

Unseen Categories

The categories defined need to be cover most types of problems that will be provided. If a new category of problem is provided, the agent might be weaker with this type of problems as it will need to learn how to tackle this new type of problem.

Overfitting

This area of concern might bring some challenges. An example is problem Basic B-04. If the rotation criteria are made only based on existing problems (ex: only 90 degrees rotations), it might fail to handle other type of problems where new degrees of rotation are encountered (ex: 45 degrees).

Non-Standard Figures

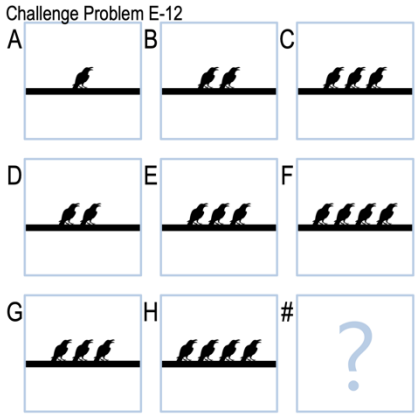

Figure 5 - RPM Challenge Problem E-12

Some of the shapes are not basic geometrical shapes. An example, as seen on Figure 5, is Challenge Problem E-12. This problem has ravens instead of simple geometrical shapes. This could be a challenge to identify as they are filled in and might be difficult to identify if any type of image pixel analysis happens.

Not good enough answer

There is a possibility that the agent is not able to generate an answer that is similar enough to the potential ones provided by the RPM test. Even though the agent will iterate on this and will then use combinations of techniques, what if it never finds one? The agent will need some sort of threshold to know that it is time to stop trying and move on to the next problem.

Computational Cost

If the proposed implementation for the agent is not optimal from a performance perspective, it could take a long time to generate the potential solutions.