Solving the Raven Progressive Matrices

Casey Long (clong80@stanford.edu)

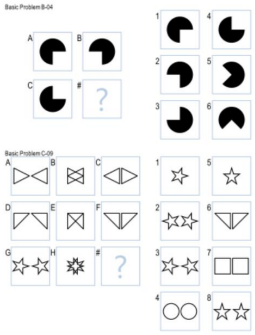
Introduction

The Raven Progressive Matrices are common tests that measure "fluid intelligence"—the capacity for one to solve problems independent of prior knowledge. Over time, it has been shown that the Raven Progressive Matrices correlate well with other intelligence tests, making it one of the best single psychometric tests for predicating human general intelligence, *g*.

Understanding these tests provide unique insight for the development of artificial intelligence programs that model human intelligence

Dataset

120 questions (1,944 images in 32-bit RGBA PNG format) partitioned into difficulty sets which resemble the original Raven Progressive Matrix intelligence test book. Each question is a 2x2 or 3x3 matrix where the artificial intelligence agent must select the correct solution image that satisfies the matrix constraints.



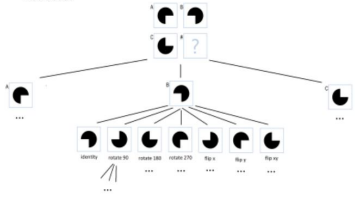
Background

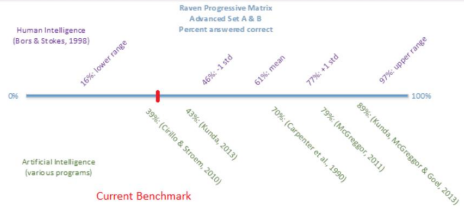
Hunt proposed two types of algorithms for solving the RPM problems; Analytic and Gestalt. Analytic algorithms involve decomposing RPM problems into a propositional representation network and performing logical operations on that data extracted. Gestalt refers to Gestalt psychology, and deals with the generation of some new image out of constituent images or patterns. If Analytic algorithms are said to be *deductive* in nature, one could say Gestalt algorithms are *inductive* or *abductive* in nature.

With the exception of the work Kunda, McGreggor, & Goel, all RPM algorithms have been analytic in nature. Analytic type algorithms have problems with more complex transformations that can not be found in the domain of multiple permutated basic transformations (e.g., rotations, translations, pixel AND, OR, XOR). (Problem B-04 to the left is fairly easy with simple transformations, but Problem C-09 a bit harder without overfitting algorithm to work for that specific problem)

The algorithm I have been trying to implement has been based off the work of Kunda, McGreggor, & Goel, but with two main deviations:

- Kunda et al "hardcode" (for lack of better words) their transformations. (i.e., 180 rotate + flip x is considered 1 transformation, and the search tree would only have depth of 1). My implementation does a greedy BFS in Steps I and II (not optimal, but needed to get it to work) to iteratively add transformation 1 at a time. This, hopefully, increases the domain to find a better fitted image
- Kunda et al worked primarily on pixel differences. Feature data, such as shapes, centroids, connected components, were not used.





Algorithmic Model

General idea based off Kunda, McGreggor, & Goel :

I. For each image, $\mathbf{I_{base}}$, in matrix, apply some transformation, $\mathbf{t}$, to create new image $\mathbf{I_{new}} = \mathbf{t}(\mathbf{I_{base}})$

II. For each newly generate image, compare image similarity via function $\mathbf{s}(\mathbf{I_{base}}, \mathbf{I_{new}})$
- if similarity is above threshold (i.e., 0.9) or reach depth limit go to III
- else, goto I and apply transformations again to form next image $\mathbf{I_{new+1}}$

III. Determine image composition operation, $\oplus$, and psuedoimage, $\mathbf{X}$
- Calculate feature differences between $\mathbf{I_{base}}, \mathbf{I_{new}}$. If $\Sigma(\mathbf{I_{base}} - \mathbf{I_{new}})$ is positive, $\mathbf{X} = \mathbf{I_{new}} - \mathbf{t}(\mathbf{I_{base}})$ Else, $\mathbf{X} = \mathbf{t}(\mathbf{I_{base}}) - \mathbf{I_{new}}$
- Composite transformation $\mathbf{T} = \mathbf{t}(\mathbf{I_{new}}) \oplus \mathbf{X}$

IV. Solve for solution image $\mathbf{Z}$ where $\mathbf{T}(\mathbf{Z}) = \mathbf{t}(\mathbf{Z}) \oplus \mathbf{X}$

*$\mathbf{X}$ represents a set of compensation variables to scale/infer what the next image is. You can think of it like the portion of a step-size scalar that increments a function to the next step. In the case of just comparing pixel differences, it would be the noise difference between two otherwise equal images. The intent here was to use more than just pixels however, i.e., sets of features.

* $\oplus$ represents the operation by which to concatenate two images (e.g., for pixel differences, it could be pixel-by-pixel AND, OR, XOR), but for sets it can be more defined logical operations.

*$\mathbf{s}(\mathbf{I_{base}}, \mathbf{I_{new}})$ comparing images for similarity can be done many ways. Tversky's similarity algorithm used in many RPM examples is defined as: $s(A,B) = \frac{f(A \cap B)}{f(A \cap B) + \alpha f(A-B) + \beta f(B-A)}$

Where α and β represent weights for exclusive features in A and B.

Results

Current benchmark results in a correct score 40% of the time. The chart above shows comparisons of human intelligence (purple) to other RPM programs (green). More indepth analysis can be done by comparing Item Response graphs of human sampled test scores to agent scores to see how agent performance compares to human performance.

Future Work

A few thoughts on working through this project:

- An interesting work I recently found solves RPM problems via image segmentation Markov Networks. It does not use propositional representations or image transformations. It involves only the statistical interaction between pixels.
- I hope to work on this project a bit more after this class and hopefully top the scores of Kunda, McGreggor, & Goel and publish something novel!

References

Bilker, W. B., Hansen, J. A., Brensinger, C. M., Richard, J., Gur, R. E., & Gur, R. C. (2012). Development of abbreviated nine-item forms of the Raven's standard progressive matrices test. Assessment, 19(3), 354–369.

Carpenter, P., Just, M., and Shell, P. (1990). What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test. Psychological Review, 97(3): 404-431.

Hunt, E. (1974). Quote the raven? Nevermore! In L. W. Gregg ed., Knowledge and Cognition (pp. 129–158). Hillsdale, NJ: Erlbaum.

Kunda, M., McGreggor, K., & Goel, A. K. (2012). Reasoning on the Raven's advanced progressive matrices test with iconic visual representations. In 34th Annual Conference of the Cognitive Science Society. Sapporo, Japan.

Kunda, M., McGreggor, K., & Goel, A. K. (2013). A computational model for solving problems from the Raven's progressive matrices intelligence test using iconic visual representations. Cognitive Systems Research, 22:47–66.

Wang, K., Su, Z. (2015). Automatic Generation of Raven's Progressive Matrices. International Joint Conference on Artificial Intelligence. 24:903-909

Winston, P. (2011). The Strong Story Hypothesis and the Directed Perception Hypothesis. In Advances in Cognitive Systems: Papers from the 2011 AAAI Fall Symposium.