Report of Project 2

ZHENG FU
zfu66@gatech.edu

Abstract. In this report, I describe the algorithms underlying a knowledge-based AI agent for solving 3 x 3 puzzles in Raven's Progressive Matrices test using visual representations. A further discussion was made to elucidate the difference between the AI agent behaviors and the ways a human using to solve these problems.

Section I

In this project I intend to use visual representations to solve the 3 x 3 Raven's Progressive Matrices (RPM) test. Since the final project (problem set D and E) does not offer verbal presentation of puzzles and I want to be familiar with how to solve RPM problems in a visual way. I was inspired by Daniel R. Little et al. 2012 paper that summarized eight rules to solve the RPM problem visually:

> "1) constant, 2) increment or 3) decrement, 4) permutation, 5) logical AND, 6) logical OR, 7) logical XOR and 8) a Distribution of 2 rule." (Little 2012)

Based on the rules mentioned above, here are the steps I improved my AI agent over time:

1. On the first round I design the AI agent to check the simplest scenario: all three images in one row are identical, or all three images in one column are identical.
2. On the second round the AI agent is able to check if the three images in one row have the "Addition" relationship. Here "Addition" relationship is defined as the (Number of black pixels of image 1) + (Number of black pixels of image2) = (Number of black pixels of image 3).

3. Then I make it possible for AI agent to determine if the three images in one row have a "Zoom-Out" tendency, such as the Basic Problem C-02 (see Figure 1 for more details).

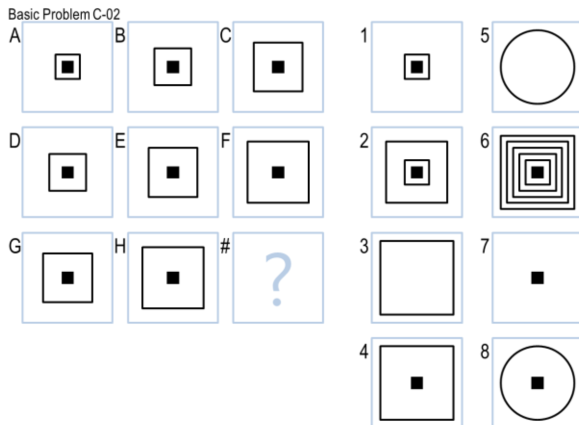

Figure 1. Basic Problem C-02.

4. After that I implement a new function to AI agent to evaluate if the three images in one row have the "Incremental" relationship. Here "Incremental" relationship is defined as (Number of black pixels of image 2 – Number of black pixels of image 1) is proportional to (Number of black pixels of image 3 – Number of black pixels of image 2).
5. In order to break ties of candidates who have equal opportunity to be the correct answer, I introduce the "Overlap-Incremental" rule to the AI agent. Thus it could gauge the pixel differences between "Overlapped image 1 and 2" and image 3 and make a decision.
6. Finally the AI agent could recognize if the whole picture (3 x 3 images) is center-symmetric such as the Basic Problem C-07.

Section II

As far as my final agent was concerned, its process of solving the problem contains two major parts: 1) Pre-processing the images and 2) Determining the correct answer through a series of rules.

1. Pre-processing the images.

All images (question images and potential answer images) were first resized to 184 pixels times 184 pixels and converted into greyscale ("L") as Numpy arrays. Then if a pixel grey scale value is no more than 65, it will be reset to 0 (black), otherwise it will be reset to 255 (white). This step will make the image sharper and reduce the errors while computing the pixel differences. Finally the greyscale Numpy arrays were transformed to python Pillow image objects.

2. Determining the correct answer by rules.

The AI agent employs a series of rules one-by-one until it finds the correct answer. If all rules were applied and no answer was selected, the AI agent will skip such questions.

1). "Identical" rule.

If the number of black pixels of all three images in one row, or that of all three images in one column are identical, the AI agent will select answer image who has the same number of black pixels as that of image G and H, or as that of image C and F.

2). "Addition" rule.

If the absolute value (number of black pixels of image A + number of black pixels of image B - number of black pixels of image C) over total pixels (184 x 184 = 33856 pixels) is no more than 1.05%, and the absolute value (number of black pixels of image

D + number of black pixels of image E - number of black pixels of image F) over total pixels is no more than 1.05% as well, the AI agent will believe the three images in one row follow the "Addition" rule. And it computes the absolute value (number of black pixels of image G + number of black pixels of image H - number of black pixels of candidate answer) over total pixels and find out whose number of black pixels will make such ratio no more than 1.05%.

3). "Zoom-out" rule.

The AI agent first calculates the number of black pixels ratio between B and A, C and B, E and D, and F and E. If all four values are more than 1, the AI agent will believe the first image in each row is "zoomed-out". Then it will estimate the average "zoom-out" ratio of the first row and the second row as:

{[(Number of black pixels of B over that of A) over (Number of black pixels of C over that of B)] plus [(Number of black pixels of E over that of D) over (Number of black pixels of F over that of E)]} over two

And the AI agent computes (Number of black pixels of H over that of G) over (Number of black pixels of candidate answer over that of H) and use this value dividing the average "zoom-out" ratio of the first row and the second row as "zoom-out ratio difference". If the "zoom-out ratio difference" is no more than 1.05 and no less than 0.95, the candidate answer will be selected as a correct answer.

4). "Incremental" rule.

The AI agent first calculates the number of black pixels differences between B and A, C and B, E and D, and F and E. If all four values are more than 0, the AI agent will believe the first image in each row is continuously added new objects. Then it will estimate the average "Incremental" ratio of the first row and the second row as:

{[(Number of black pixels of B minus that of A) over (Number of black pixels of C minus that of B)] plus [(Number of black pixels of E minus that of D) over (Number of black pixels of F minus that of E)]} over two

And the AI agent computes (Number of black pixels of H minus that of G) over (Number of black pixels of candidate answer minus that of H) and uses this value dividing the average "Incremental" ratio of the first row and the second row as "Incremental ratio difference". If the "Incremental ratio difference" is no more than 1.05 and no less than 0.94, the candidate answer will be selected as correct answer.

5). "Overlap-Incremental" rule.

If more than two correct answers were selected after applying "Incremental" rule, the AI agent will employ "Overlap-Incremental" rule for further judgements. It overlaps image B with image A and computes the number of black pixels differences between "Overlapped image A and B" and image C, and do the same thing for image D, E and F. Then it estimates the average "Overlap-Incremental" ratio of the first row and the second row as:

[(Number of black pixels of "Overlapped image A and B" minus that of image C) plus (Number of black pixels of "Overlapped image D and E" minus that of image F)] over two

And the AI agent computes (Number of black pixels of "Overlapped image G and H" minus that of candidate image) and uses this value dividing the average "Overlap-Incremental" ratio of the first row and the second row as "Overlap-Incremental ratio difference". If the "Overlap-Incremental ratio difference" is no more than 1.01 and no less than 0.99, the candidate answer will be selected as correct answer. If there still are more than one correct answer after applying "Overlap-Incremental" rule, the AI agent will automatically choose the smallest answer key from the potential correct answers.

6). "Center-Symmetric" rule.

The AI agent will overlap 180-degree rotated image B and image H, and compute the number of black pixels difference between these two images. The same action will take on image C and G, D and F as well. If all three number of black pixels differences are no more than 1%, the AI agent will believe the whole picture (3 x 3 images) is center-symmetric. Then it will overlap 180-degree rotated image A and candidate answer image and compute the number of black pixels difference between these two images. If such difference is no more than 1%, the candidate answer image will be selected as correct answer.

Section III

How many problems does it answer correctly?

The following table summarizes the accuracy of the final AI agent.

Table 1. The accuracy of the final AI agent for Problem Set C.

Problem Type	Correct	Skipped	Incorrect
Basic	11	1	0
Test	11	1	0
Challenge	3	7	2
Ravens	4	4	4

How efficient is it?

The AI agent uses 6.04123 seconds to solve the problems, so I believe it is quite efficient.

How general is it?

The AI agent did well on Basic and Test problems yet for Challenge and Ravens questions, its performance was not very good.

Different performance on the Basic and Test sets?

The AI agent has the same good performance on Test set as Basic set. On Basic set it has 92% accuracy and on the Test set it has 92% accuracy as well.

Section IV

What types of problems does it currently answer incorrectly?

The final AI agent skips one questions in Basic Set and one in Test set. Due to the lack of information about Test Set, I do not know what types of problems in Test Set it skips. And the details of skipped one question in Basic Set will be discussed following.

What kinds of problems would it struggle with?

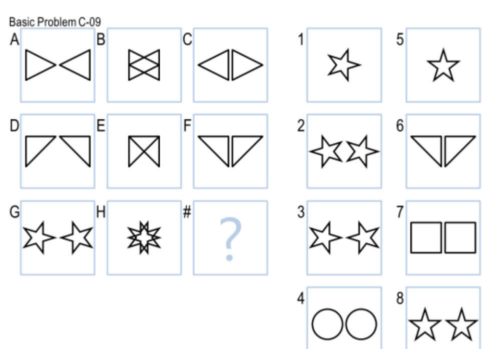

Figure 2. Basic Problem C-09.

According to the design of my final AI agent, there are two kinds of problems it will struggle with:

1. Problems like Basic Problem C-09. After applying "Incremental" rule followed by "Overlap-Incremental" rule, there are still three answers (image 2, 3 and 8) in potential correct answer pool (see Figure 2). So the AI agent just choses the smallest key value that by chance it is the correct key.
2. Problems like Basic Problem C-12. After applying all rules mentioned in "Section II", the AI agent still could not determine which one is the correct answer thus it skips this question.

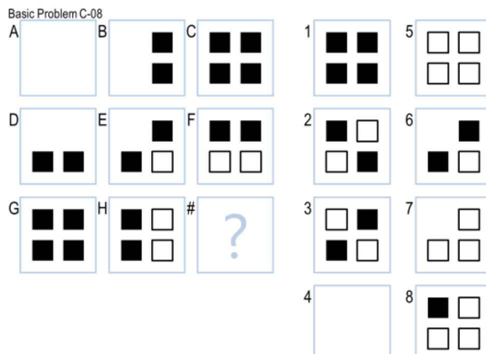

Figure 3. Basic Problem C-08.

Section V

For simple questions like Basic Problem C-05, the AI agent could "perceive" that objects were added to the image by computing the number of black pixels changing, as we human-beings directly see that by eyes. However for the following problems, I neither feel that the nature of my revisions reflecting the way a human learns from experiences, nor feel my final agent solves the problems similar to how a human would do so.

1. Basic Problem C-09: We could see in each row, the two objects pass through each other and stop at each other's initial position (see Figure 2). Such process

could be an easy and natural imagination in our brains, yet it is quite difficult to teach AI agent how to learn such "dynamic" process by using static images. As described in Section IV, my AI agent just selects the correct answer by chance.

2. Basic Problem C-08: We could recognize the correct answer will make the whole picture symmetric by diagonal, and the colors were decayed from G to # and from C to # as well (see Figure 3). Such two recognitions occur almost simultaneously even at the first glance. Yet the AI agent does not have the concept of "whole picture" nor "color". It must compare images one by one and compute their number of pixels changing to make the final decision.

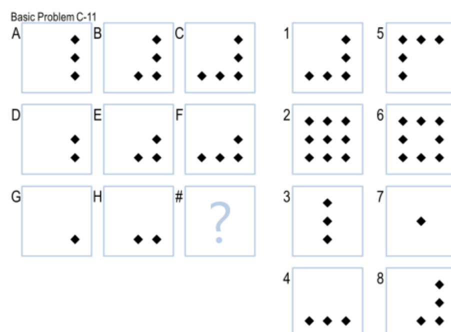

Figure 4. Basic Problem C-11.

3. Basic Problem C-11: We could see on each row, a diamond object is continuously added at the bottom line (see Figure 4). Yet the AI agent could only understand there is one object was added. So we have to overlap the first two images on each row and compare it with the third image, so as to tell the AI agent where this object is added, otherwise both 3 and 4 will be the correct answer. Such additional step is not necessary for human beings to solve this problem.

References

1. Little, D., Lewandowsky, S., & Griffiths, T. (2012). A Bayesian model of rule induction in Raven's progressive matrices. *In Proceedings of the 34th Annual Conference of the Cognitive Science Society.* Sapporo, Japan.

2. Evans, T. G. (1964). A heuristic program to solve geometric-analogy problems. In *AFIPS '64 (Spring) Proceedings of the April 21-23, 1964, spring joint computer conference.* Washington, D.C., USA.
3. Joyner, D., Bedwell, D., Graham, C., Lemmon, W., Martinez, O., & Goel, A. (2015). Using Human Computation to Acquire Novel Methods for Addressing Visual Analogy Problems on Intelligence Tests. In *Proceedings of the Sixth International Conference on Computational Creativity*. Provo, Utah.