Project 2

Sergiy Palguyev

sergiy.palguyev@gatech.edu

Introduction

The second project of this class involves producing an Artificial Intelligence agent that is capable of solving Raven's Progressive Matrices problem in 3x3 type problems. The first project solved 2x2 problems which were simple enough to evaluate based on symmetry, rotation and equivalencies. This project, however, is much more complicated since there are many more checks that are possible in order to evaluate for a possible answer. In order to facilitate this, the solution begins by evaluating algorithms similar to those from the 2x2 project. Afterwards, the differences between the items in the 3x3 matrix are calculated and evaluated for similarities, the basis of which will provide for possible solutions.

Submission: Mar 12, 07:51:46 UTC

For this submission, the solution from the 2x2 matrix is rewritten for 3x3 evaluation. The algorithms for checking rotations are hardcoded for the 2x2 matrix and do not make it in to the first submission. With the limited amount of evaluation for similarity across rows and columns, very few of the problems are found to be correct. This agent is very loosely related to human cognition as it checks simply if the images are similar across rows and columns. This very basic level of evaluation is barely at a child's level of understanding shapes and their relationships. The performance of this submission is quite poor as evidenced by the results as "correct:skipped:wrong" answers for Set C are:
Basic-1:9:2, Challenge-0:12:0, Test-2:9:1 and Raven's-0:9:3.

Submission: Mar 12, 07:54:16 UTC

This submission was accidental and occurred only a few minutes after the previous one as a mistake when executing the local project. Answers for Set C were
Basic-1:9:2, Challenge-0:12:0, Test-2:9:1 and Raven's-0:9:3.

Submission: Mar 17, 05:00:25 UTC

For this submission, every image in the 3x3 matrix is translated to a binary matrix. Afterwards, a ratio is computed for each image, specifying the Black to White pixel count ratio. Furthermore, an algorithm evaluating the differences across rows, columns, and diagonally is implemented in order to determine a possible answer set. This evaluation is much closer to more advanced human cognition as the evaluation covers similarity across all dimensions of the 3x3 image matrix. On the other hand, there is still no evaluation of actual shapes within the images yet, as that would present a much higher level of cognition and understanding by the AI agent. The performance at this point is fairly good with both Basic and Test sets achieving a 5/12 correct answers.
Answers for Set C are:
Basic-5:0:7, Challenge-1:0:11, Test-5:0:7 and Raven's-2:0:10.

Submission: Mar 17, 08:51:26 UTC

For this submission, the algorithm that evaluates the possible answers is expanded. If there are less than 3 answers suggested, then the lowest differential value is picked as the answer value. For any other set of answers above 3, a random value is chosen. This evaluation gets is even closer to human level understanding as it moves ever closer to an "educated guess". As a form of background knowledge, when there are very few answers, the one with the lowest difference would tend to be the correct one and is picked as such. The performance of the agent at this point gets even better with Basic, Challenge, and Raven's data sets all getting an increase in correct answers.
Answers for Set C are:
Basic-8:0:4, Challenge-4:0:8, Test-5:0:7 and Raven's-6:0:6.

Submission: Mar 17, 09:21:33 UTC

For this submission, in the answer evaluation algorithm if the returned form the initial 3x3 basic evaluation is -1 then and the differential check also failed to provide an answer, the potential answer set is further evaluated. If the answer set contains two images with the same black/white ratio, then the answer is most likely one of those two images. The images are then compared against the problem set and evaluated for best answer. This behavior is further development

of background knowledge present in human cognition. If there are two answers that are, for lack of argument, the same, then it is most likely that one of these two choices will be the correct answer. With this logic, the ratio of correct answers further increases for the Basic set to have 10/12 problems correct. None of the other categories increased in correct answers.
Answers for Set C are
Basic-10:0:2, Challenge-3:0:9, Test-5:0:7 and Raven's-6:0:6.

Submission: Mar 17, 09:30:30 UTC

For this submission, the algorithm for a returned value of -1 is expanded again from the previous submission. If there are no two or more answer images with similar ratio, then the answer with the highest black/white ratio is picked as the answer. This expansion of the background knowledge and slight modification of the answer algorithm allows more correct answers to be picked for the Challenge and Raven's problem sets, with the Raven's set achieving 7/12 correct answers.
Answers for Set C are
Basic:10:0:2, Challenge:4:0:8, Test:5:0:7 and Raven's:7:0:5.

Submission: Mar 17, 09:59:29 UTC

For this submission, the algorithm for a returned value of -1 is expanded again from the previous submission. If the previously described subroutines also return -1, then a number is randomly picked between 1 and 8 to serve as the answer. This is a very minor change to allow for random picks at the end of an algorithm and reflects human cognition in simply reflecting the humans would prefer to guess rather than skip over any answers if guessing wrong is not penalized. The performance of the agent increases in accuracy with the challenger results gaining 6/12 correct answers, while the other sets stay the same.
Answers for Set C are
Basic:10:0:2, Challenge:6:0:6, Test:5:0:7 and Raven's:7:0:5.

Submission: Mar 18, 04:57:34 UTC

For this submission , the final "else" statement, after all possible other options are exhausted, is modified from returning the value initially calculated by the 3x3 matrix evaluation to return a random value. This again is a small change to see if

returning a random value works better thant he 3x3 matrix evaluation performed at the beginning. This change has no similarity to human cognition other than the same behavior of random guessing when no other logic works. The performance of the Agent takes a fairly significant hit, with correctness from Challenge, Test, and Raven's set going down to 5/12, 3/12 and 6/12 respectively.
Answers for Set C were
Basic:10:0:2, Challenge:5:0:7, Test:3:0:9 and Raven's:6:0:6.

Submission: Mar 18, 04:59:24 UTC

The algorithm from the previous submission is set back to return the value from the original 3x3 matrix evaluation. The performance is somewhat lackluster and does not return to original correct numbers.
Answers for Set C are
Basic:10:0:2, Challenge:5:0:7, Test:4:0:8 and Raven's:5:0:7.

Submission: Mar 18, 05:02:30 UTC

No change to the algorithm, simply another run of the entire Agent. The performance of the AI agent returns back to original values set in the submission from Mar 18, 04:57: UTC
Answers for Set C are
Basic:10:0:2, Challenge:4:0:8, Test:5:0:7 and Raven's:5:0:7.

Conclusion

In conclusion, the AI agent works fairly well for evaluating 3x3 Raven's Progressive Matrices problems. The best performance reached included Basic set with 10/12 correct answers, Challenge set with 6/12 correct answers, Test set with 5/12 correct answers and Raven's set with 7/12 correct answers. The Ai agent closely models human cognition to the extent of evaluating images in black/white ratios and comparing the images across all patterns of a 3x3 matrix. Additionally background knowledge is provided I the answer algorithm to account for "educated guassing" when multiple answers are still left over in the end of the evaluation. For future development, the most likely set of advanced algorithms would involve evaluations for image rotations, further development of 3x3

problem evaluations as well as shape recognition, if possible. These feature will be attempted for the next project.