An Investigation of Compositional Planning

Alyssa Unell (6.804), Meagan Jens (6.804), Miranda Carson (9.66), and Adrian Meza (9.660)

Introduction

By: Adrian

Human planning can take many shapes and forms. When playing a game like chess, for example, we can imagine the brain forms some type of decision tree, and certain latent heuristics are learned with experience to prune such a tree for more strategic and game winning moves. Humans use planning on a daily basis to figure out what the most efficient way to get to work or school might be; whether to take certain sidewalks, or buses, or other forms of transportation. Sometimes even using tools like digital maps which themselves are planning problems behind the scenes. All the being said however, the consensus on how exactly humans make planning decisions is limited.

There have been a few studies on how people prune decision trees, what types of habits are formed to save the amount of effort one must do, and how we come the decision to stop after a sufficiently good solution is found [5]. Other studies have also shown that real life pedestrians navigate through near-optimal paths [6].

But for our project, however, we examine the planning problem with a focus on maze planning with hidden information. Specifically we want to analyze how humans make planning decisions on a maze given a start node, and some hidden state in the maze, among which a target node is randomly placed. By analyze human data on a few sample mazes, along with other planning techniques on the same mazes, we hope to draw insight on the latent heuristics humans incorporate into their planning approaches.

We want to focus specifically on the hypothesis that people form efficient state space representations [7] of their planning problem. With respect to the maze-search task , we hypothesize that humans actually break up the maze into a hierarchical representation where they plan on each representation separately. In other words, the maze is broken up into pieces, and a path is made for each separate piece, or composition, without much knowledge of the other parts. This is what we call Compositional Planning.

Our goal is to address the following questions under these assumptions:

1. Do humans use composite planning?
2. If several composite map representations are possible, which representation do people use?
3. Do people use a comparison that allows for maximal compression
4. Does humans planning follow realistic goal setting?

By tackling these questions we hope to better understand how humans undertake planning tasks in complex environments. When analyzing and contrasting various planning strategies, including an Expected Utility Model, a Random model, and a Composition model, with real human data on a few custom mazes, we hope to show new insights into human heuristic planning and general human decision making strategies.

Maze Structure

By: Miranda

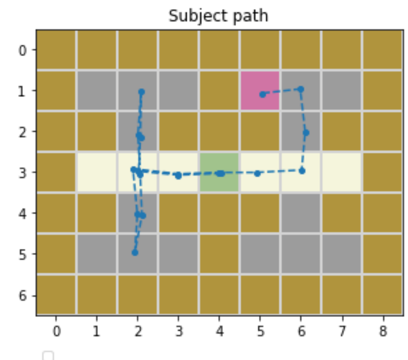
The mazes in this study were generated as an extension of the human planning research performed by Essie Yu and Marta Kryven [4]. Their pre-existing experiment setup for human testing and modeling of planning tasks involves the use of two-dimensional grid mazes, in which the given task is to seek an exit in the least number of steps. Exploration of the maze unveils cells in the line of sight of the player character's location. The human subjects (and the model) are told that the exit tile is equally likely to be placed under any hidden tile.

During prior experimentation with a variety of maze patterns, Yu and Kryven observed that under some circumstances, human paths diverged significantly from the Expected Utility model, indicating the human subjects were navigating in a non-optimal manner. Their key finding was that when the mazes could be deconstructed into smaller sections that each involved deeper exploration, many players would opt to delve deeper into each room to explore them fully, one at a time, rather than search the breadth of rooms to unveil as many tiles as possible as soon as possible.

Our team was presented with one maze from the original experiment's maze set that originally sparked interest into the possibility of a compositional heuristic rather than a utility-based one. We have included this original maze in our investigations as Maze 6, in order to validate our models against prior data. We created five additional mazes, labeled Maze

1 through Maze 5, that we believed to contain features that would evoke this compositional planning behavior. Each maze contained ”rooms” that required the subject to navigate through two nodes in order to fully explore each room: one node at the entrance to the room, and another node to check around an obscured corner inside a room. The mazes varied in path lengths required to fully explore each room, and the rooms were repeated either symmetrically or linearly along a hallway.

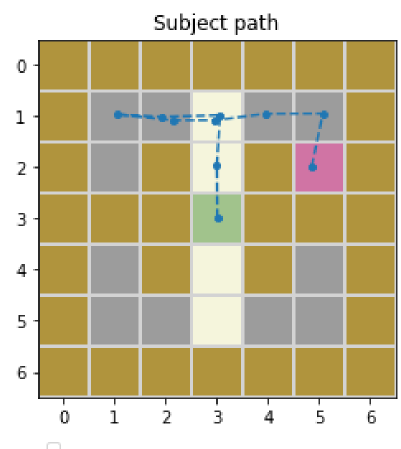
Maze 1



(a) Maze 1, Subject Path

Maze 1 and Maze 2 require the player to move two steps to a decision node, then two more steps to enter the room. Maze 1 rewards the player with two additional unveiled squares for exploring a room, while Maze 2 rewards the player with only one. Maze 3 increases the path length between each room, shifting the utility of the room exploration to only one unveiled cell after 3 moves into a room. Due to this, in Maze 3, our expected utility model matched our compositional model. In fact, the players that were attempting to follow the original heuristic of searching each hallway before searching each room were using so many steps that we had to remove the experiment’s step limit. These players were unable to reach the exit within the original limit due to significant backtracking. Maze 4 is the most complicated maze, allowing for the possibility of players deconstructing the maze into either two or four sub-mazes. Maze 5 is modeled after Maze 6, but with a decision node at an entrance to two rooms at one location, rather than one room at a time.

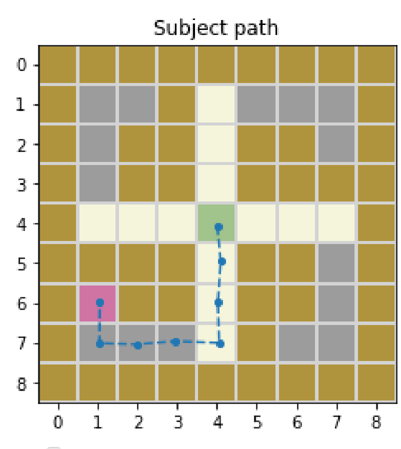
Maze 2



(a) Maze 2, Subject Path

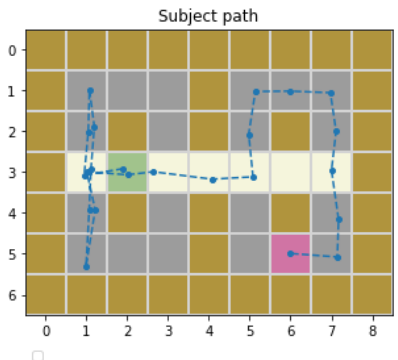
By hosting the experiment on MIT XVM, we were able to collect the trajectory of tested human subjects who were given a task to complete the mazes in as few steps as possible. For each maze, the step-wise trajectory was collected as a list of tuples. Each tuple recorded the x,y coordinates of the player character on the map at each step, and the length of the list recorded the number of steps taken in each maze. The mazes were presented in random order, and additional information on reasoning, logic skills, and simple demographic information was collected after the completion of the six mazes. The experiment was presented to friends and acquaintances of the authors of this paper, returning 50 logged experiment results, of which 23 completed the full experiment. Incomplete results were discarded. The six mazes are presented here along with a single experimental subject’s trajectory plotted across each maze who consistently uses the compositional planning techniques we were looking to observe.

Maze 3



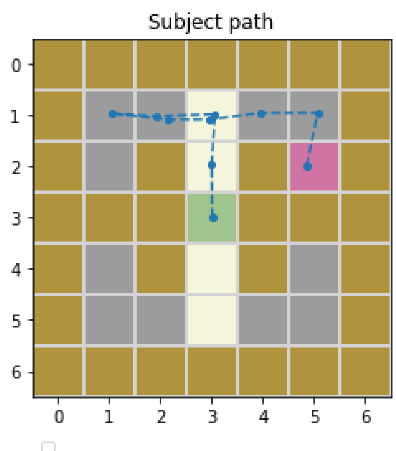
(a) Maze 3, Subject Path

Maze 4



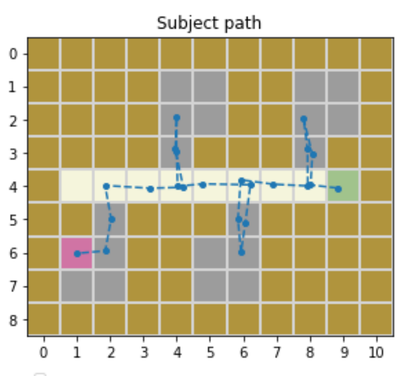
(a) Maze 4, Subject Path

Maze 2



(a) Maze 2, Subject Path

Maze 6



(a) Maze 6, Subject Path

Baseline Models

By: Alyssa and Meagan

To model the navigation of the maze, we broke down the planning process into states. To move from one state to another, the agent must make an observation. An observation is made when an agent lands on a cell in the maze that reveals one or more hidden cells. A tree structure is used to represent the planning process as the agent moves from one state to another. The root node represents the initial state, which is when the agent is in the starting position. The children of a given node each represent a possible next state that can be reached from that node. Figure 1 is an example of the tree structure used in the models.

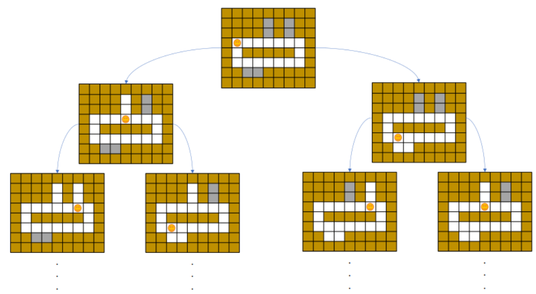


Figure 7: Tree Structure

Using this tree structure, we implemented two baseline planning models: Random and Expected Utility. The value of a node is defined as the number of steps required in order to exit the maze from that state. Both models use a different heuristic to assign node values and, therefore, decide on different optimal paths. An initial maze is passed into a function that identifies the best path on the maze given the desired heuristic. In order for the heuristic to function, it converts this tuple of tuples (the initial format of the maze) into the above tree structure. This tree structure utilizes a dictionary datatype so each node and its value can be tracked with regards to its root and parents. In this way, many different aspects of path finding are able to be explored and leveraged given the different visualization functions. The goal of the these baseline models is to compare them to human data and compositional planning models to gain a deeper understanding of how humans navigate mazes.

Random Model

The random model evaluates each path equally, meaning that for a given node in the tree, the node's children are all assigned the same value. For example, if a node had three children, each of those children would be assigned the value of $\frac{1}{3}$. The purpose of this model is to verify that humans do in fact use a planning mechanism as they navigate the maze.

Expected Utility Model

The expected utility model assigns the value of node N_i by calculating the expected number of steps needed to be taken to reach the exit. This value can be computed according to the following formula:

$$Q(N_i) = p_i(s_i + e_i) + (1 - p_i)min_{c_j \epsilon C(N_i)} Q(c_j)$$

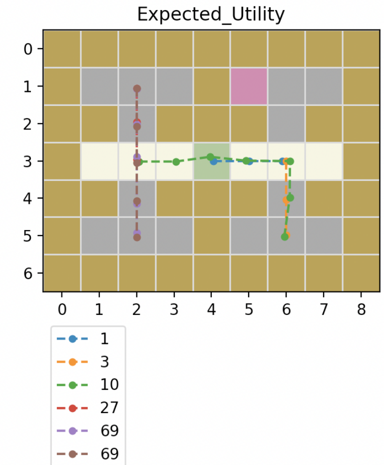
In the formula, $Q(N_i)$ is the value of node N_i, p_i is the probability of finding the exit at node N_i, s_i is the number of steps taken to reach node N_i from the root node, e_i is the expected number of steps that need to be taken to reach the exit when the exit was observed, $c_j \varepsilon C(N_i)$ is the set of all children nodes of node N_i, and $Q(c_j)$ is the value of child node C_j.

Overall, expected utility value of going to node N_i is defined by the sum of the expected number of steps to the exit (given that the exit was found on that node) and the expected number of steps to reach the exit by exploring more of the maze beyond N_i (given that the exit isn't found on that step) where these expected number of steps are weighted by the probability of finding the exit at N_i. The model evaluates the possible paths by using the expected utilities of the nodes as a heuristic to identify the shortest path.

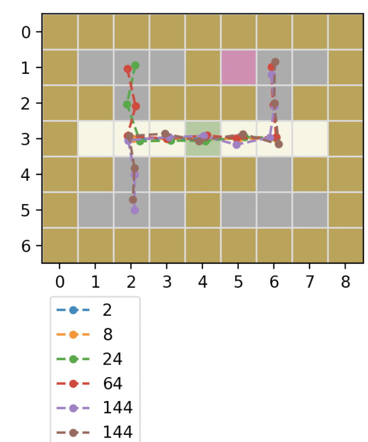
Random and Expected Utility Model Results

We look here at how the expected utility model performs on different graphs with the objective of comparing this model to human path planning. The green space in the maze is the start node, the yellow blocks are walls, the white space is seen pathways, and the gray areas are areas that can be explored by the agent. These mazes are further discussed in the above section, Maze Structure, where the objective of the maze's construction is elucidated. Each maze is shown below along with the path the agent took when following the expected utility model as well as a randomly generated path.

Maze 1

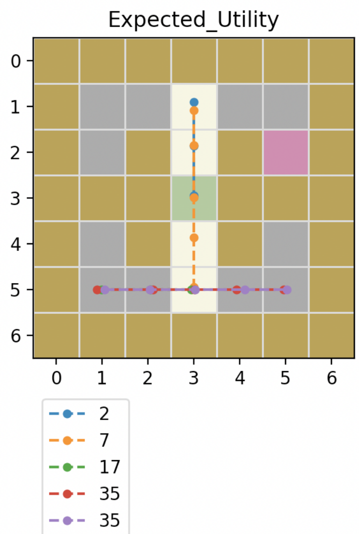


(a) Maze 1, Expected Utility

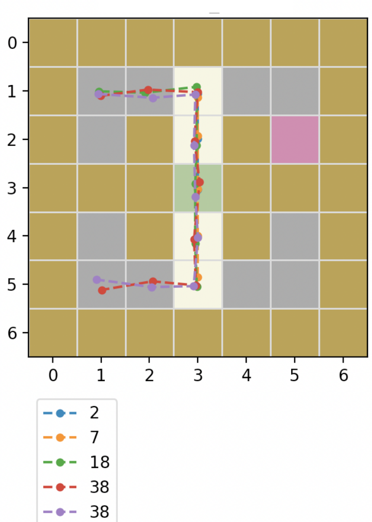


(b) Maze 1, Random

Maze 2

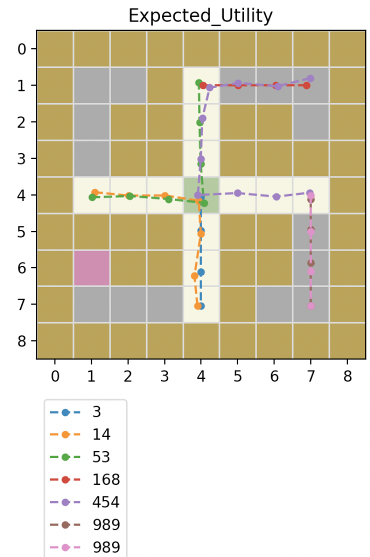


(a) Maze 2, Expected Utility

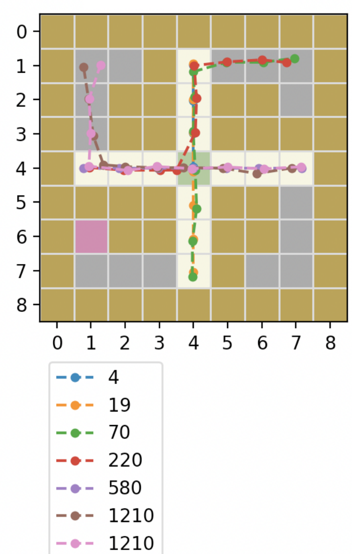


(b) Maze 2, Random

Maze 3

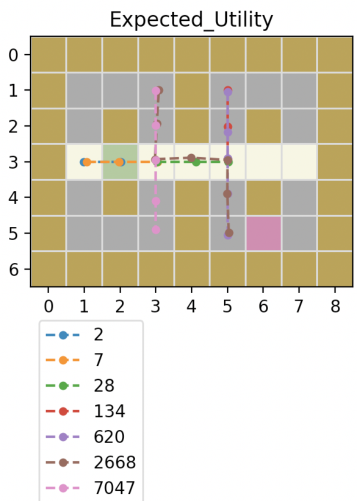


(a) Maze 3, Expected Utility

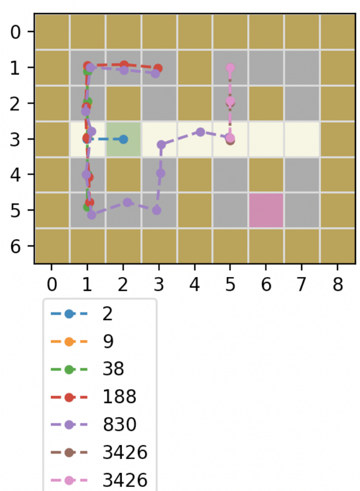


(b) Maze 3, Random

Maze 4

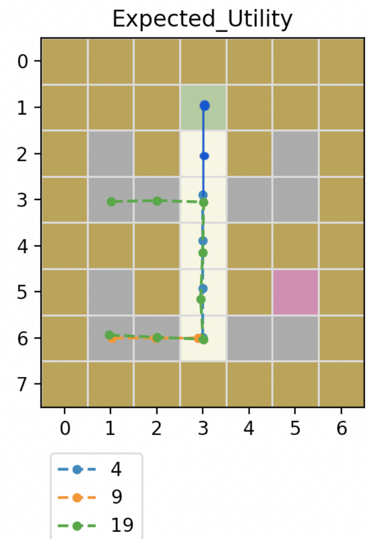


(a) Maze 4, Expected Utility

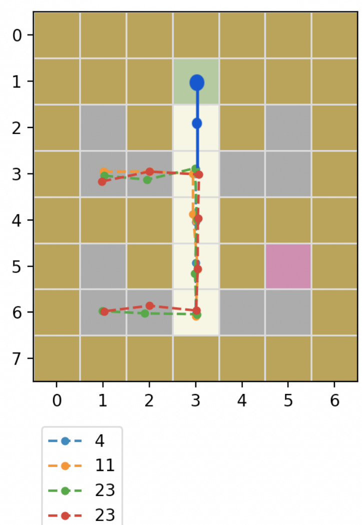


(b) Maze 4, Random

Maze 5

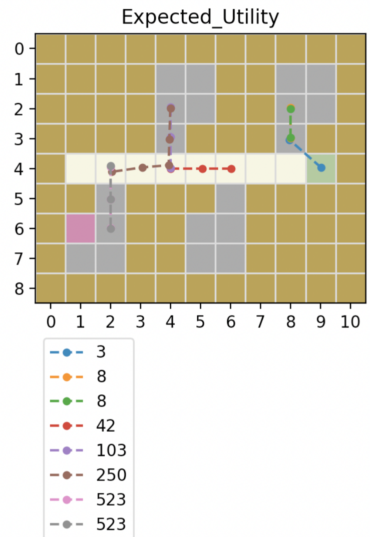


(a) Maze 5, Expected Utility

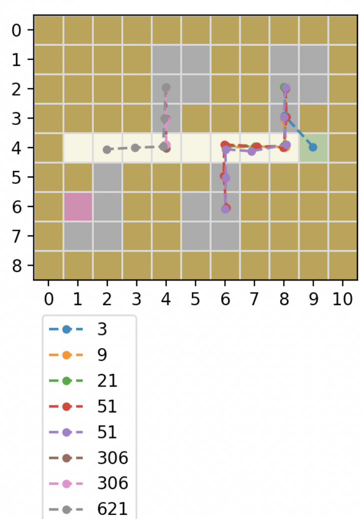


(b) Maze 5, Random

Maze 6



(a) Maze 6, Expected Utility



(b) Maze 6, Random

Compositional Planning Model

By: Adrian

For the Compositional Planning Model, we began by building a script that allowed us to split a graph into our desired compositions. Graphs were decomposed into sub-graphs that best fit the natural symmetries of the graph. We expect this resembles how a human might break apart a 2D maze were they looking at it from the top-down. For many of the test mazes, a decomposition into 4 even blocks was generally the best way to decompose the maze and capture all the symmetries.

Once the compositions were established, we ran our planner on each sub-maze, using the hidden states in each maze as our guiding points. For this project we also wanted to include the flexibility to swap out different planners to use on the compositions, with the eventual goal of studying which combination of composition and heuristic best resemble human planning. For the results presented here we ran an A* planner on each sub-maze with a euclidean distance heuristic, as well as dynamic end-points that are set to be the next closest hidden block on the map.

Compositional Model Results

The Mazes below are the same as the previous mazes show, just in grayscale. The red arrows represent the collective path of each compositional block added together, and the dotted green lines help delineate the separate compositions.

For clarity, let's work through an example on example of the steps to get the path shown in Maze 3.

1. The maze is broken into four even blocks, denoted by the dotted green lines.

2. We iterate through the sub-mazes from left to right, top to bottom.

3. In this case the planner would start with the top left block, move from the starting point (center of the white cross) to the first black square (hidden state).

4. From this viewpoint the agent can see up through the little corridor (clearing the hidden states), but cannot see the hidden corner, which is then set as the new end goal.

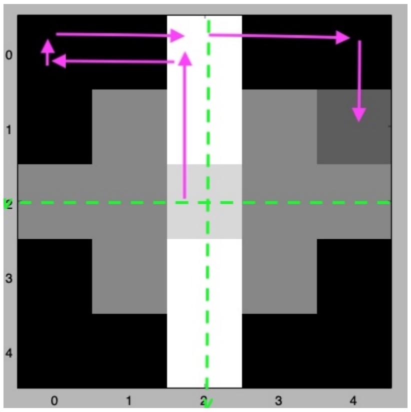
5. Once there, there are no more hidden states in the sub-maze, but the next hidden state is known to be out of bounds, so the planner travels as close as it can, which ends up being the starting point in this example.

6. Move onto the top right and repeat #1 – 5

7. Repeat for each sub-maze or until goal node is found, which is what happens in this example for the bottom left block.

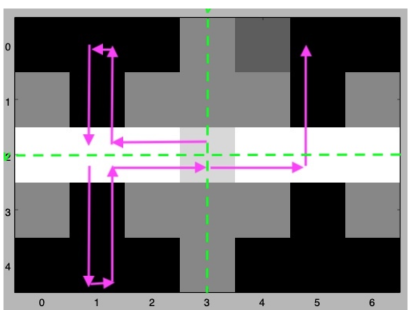
The paths produced by this process (shown in red) seem to closely resemble what a human might do. With the exception of Maze 4, which does slightly more work than one might expect due to the fact that agents' "sight" is restricted to each composition, the paths are efficient in checking the hidden states for the goal.

Maze 1



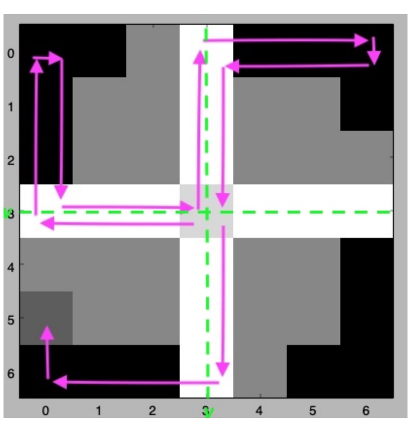
(a) Maze 1, Planner Path

Maze 2



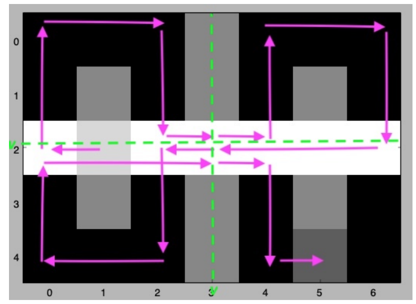
(a) Maze 2, Planner Path

Maze 3



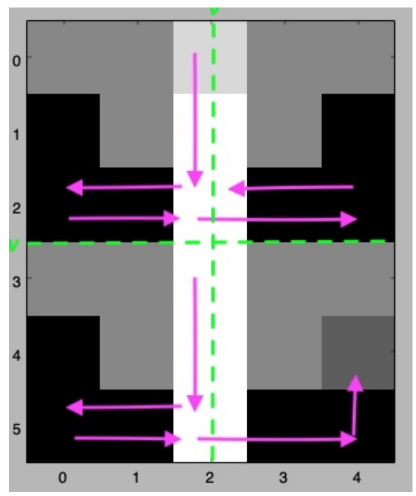
(a) Maze 3, Planner Path

Maze 4



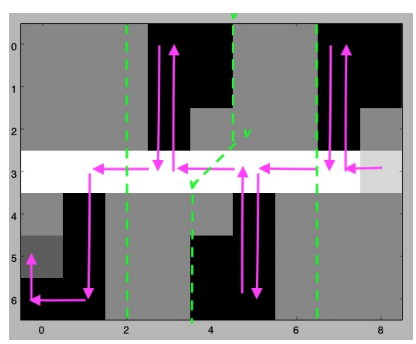
(a) Maze 4, Planner Path

Maze 5



(a) Maze 2, Planner Path

Maze 6



(a) Maze 6, Planner Path

Results and Analysis

By: Alyssa, Meagan, and Adrian

The primary objective of this experiment was to explore human planning and, in particular, to identify whether or not humans use composite planning. We developed a composite planning model as well as two baseline models (expected utility and random) and compared the paths taken by the models and human path taken.

First, we can qualitatively compare the human path and the random model path for each of the six mazes. Given that the paths significantly differ, we can conclude that humans do not randomly navigate the maze. Instead, humans use a heuristic to develop a plan. While this insight might seem trivial, it is worth noting because it emphasizes the purpose of this study.

The difference in the expected utility output from the random model is encouraging. When compared, both planners combined show us that the expected utility is accurately accounting for node values and attempting to optimize the pathway. We can see that the expected utility pathways tend to be shorter and more efficient at exploring the maze than its random counterpart, which is as expected given the complexity of the expected utility function.

Similarly we can examine the paths between one of the sample human subjects and the compositional planner, and from a glance see that they're are promisingly similar. We observe the compositional planning model, even with the ability to "see" into the distance, ends up stepping into each room to fully explore the hidden states that are difficult to reach, before moving on to the next sub-maze. With the correct maze setup, the compositional planner will actually end up exploring most of the maze out of necessity. Although is does so as a human might, without walking into parts of the maze it doesn't need to (e.g. the corners in Maze 2).

Due to the symmetry of the mazes, some decision points (such as leaving the root node) are random on some mazes, with an equal weight for the expected utility model in any possible direction, preventing us from performing true comparisons between the actual trajectories of the model and human data, but allowing us to make qualitative observations on the path structures. For example with Maze 2 we observe an almost identical path structure in each of the composition, expected utility, and human generated models. For a more realistic comparison, however, care must be taken to randomize the goal nodes as we state to the human subjects, and the order in which the sub-mazes are explored, since the starting points are dependent on that. Maze 1 is a good example of that, as the lower mazes never get explored with the current planner, whereas human subjects do sometimes make their way down there.

We hypothesize that humans are using an internal heuristic such as the expected utility model or the compositional planning model to identify optimal pathways in a maze. The expected utility model's valuation of the future and identification of node values by distance to the exit is one such method that we assert humans could be subconsciously employing while planning paths. Alternatively, we can also gather from this data that searching individual rooms as much as needed to assert the goal node is or isn't there, might be the default planing strategy. A more careful quantitative analysis remains to be done, however, and more importantly a lot more human data on a wider variety of maze shapes and sizes is needed to assert a stronger conclusion.

Conclusion and Future Scope

By: Alyssa, Meagan, and Miranda

While the results we see thus far are exciting in their insight for human planning processes, future steps must be taken to fully leverage the impact of these findings. Given the expected utility model and the random model alongside the human maze data, we can use log likelihood comparisons as well as cross validation to assess the relationship between human path planning and these computational models. Cross validation can be employed to identify the optimal parameters for each model that will lead to the closest possible match to the human approach at path finding. This step should then be followed by a log likelihood analysis that compares these optimized models with the human results in order to assert which model best matches human approaches. Currently, these steps are not utilized in our research as a fully quantitative comparison between the human results and the computational results has not been performed. However, future work would benefit from the running of these analyses to identify the relationship between these models and their human counterparts.

While we were able to capture human data from the experiment, further analysis would be required to determine how common compositional planning is chosen as an exploration technique. A sampling of human data shows that some people would select compositional planning over another technique, but this paper makes no claims as to the frequency of this occurrence. In addition, where some decision points yielded multiple paths with equal utility, some human subjects reported selecting a random path. By returning some asymmetry to the next generation of mazes, we will be able to remove this confound of players switching between their chosen planning heuristic and random decisions.

Players were able to see the entire maze layout in the 2D, and thus were able to spend the time to count squares in an attempt to perform rudimentary calculations on the expected utility on a decision. Much insight could be gained by performing this experiment again in a 3D maze environment, with a first person view, in order to observe players' behaviors towards entering rooms and checking around corners in situations where they are not able to simply count the utility of their next steps. Our expectation in this situation would be a shift away from the expected utility model, and towards the compositional planning model, as players are given less quantitative feedback over their actions in novel environments.

Additional future work would be to continue fine-tuning the models, as there are small bugs that are present within the outputs that can be fixed to provide more clarity in regard to the heuristic's relationship to human planning. For example, Maze 5's first exploration pathway was inverted and thus did not start at the original start node (the output graphs have been augmented to illustrate what the expected pathway would be without this error). Additional shortcomings that can be further modified are the heuristic objective's, as it appears that the pathway doesn't consistently explore until the exit is discovered. Finally, Maze 6 is disjoint for the random and expected utility models, and this error would have to be addressed before passing into the cross validation and log likelihood analyses with the human data.

There remains work to be done with the compositional planner as well. When we initially tackled how to split the mazes, we found that creating a script to read in any maze and split it autonomously was a research problem in and of itself. Instead we limited ourselves to three options: horizontal, vertical, and block compositions, all around the same size. With a few hundred mazes though, it wouldn't be too difficult to train a simple CNN for example, to take in any maze of similar structure and generate compositions. The current A* planner also had the issue of not supporting dynamic end points, so the generated graphs were done by running the planner on a certain start and endpoint, and manually modifying the end point and composition by hand, until we reached our goal. With the A* planner as well, we acknowledge that the assumption that humans guide themselves on euclidean distance, is a big one, and would need be verified against other planner and composition combinations. The infrastructure for these modifications is nearly complete for the most part, and should be done on various settings, along with a log-likelihood analysis, to determine the closest fit to human planning.

Log likelihood analysis in conjunction with cross-validation is imperative in order to synthesize the preliminary results that we have obtained. In cross validation, the models need to be trained with parameters to best fit human data. Our code has the capability for three parameter entries to fine tune its performance, but the plots produced used a stock parameter of value 1 in order to simplify the implementation and ensure a proof of concept was achievable. We can pick parameters like what should we scale the penalty by vs. the reward, etc. and we can fit the model to human data to find these values. Cross validation splits up the training data set list into equal subsets, leaves out one subset on each iteration of fitting the model. It calculates which model does the best with the majority of the data and uses holdout data as validation data (for example, in k-fold cross validation, there are k equal subsets, and the validation holds out each subset as it iterates through the training data). K-fold validation is one cross-validation method that we would employ to explore optimal parameters, but we could also take a randomized approach. We could take out a random section of the training samples (10 or 20 percent) and use these samples for testing. In this way, if human path finding is able to be accurately modeled by a quantitative heuristic, we will be able to find the breakdown of the algorithm that best represents human thought processes.

Once the model is established with optimal parameters, we would want to find the model that matches human planning the best. Max log likelihood estimation answers this question and as such would be suitable for employment in the pursuit of human planning algorithms. The model with the highest probability given the human data would indicate that it is the best model to represent human planning. As such, we are attempting here to maximize P(M—D), where M is the model and d is the human data. By Bayes' rule, $P(M|D) = \frac{P(D|M)}{(M)}$ and so we can assert that P(M—D) is proportional to P(D—M). The probability of the data given the model is equal to the product of decisions made by humans in the training dataset which can be rewritten as P(di—M) for all di in D. This is proportional to the sum of log(di—M). This is where the code for finding maximum log likelihood will reveal the optimal model, since this code will find the maximum sum given the results from all models tested.

From a high level, we can say that our limited data certainly suggests humans might use some form of composition, but the exact representations are parameters that remain to be tuned.

References

1. Z. Bnaya, G. Galbiati, Y. Li, W. Ma, B. Opheusden, "A Computational Model for Decsion Tree Search."
2. F. Callaway, P. Das, T. Griffiths, S. Gul, P. Kruger, F. Lieder,"A Resource-Rational Analysis of Human Planning."
3. A. Curtis, I. Fiete, M. Kryven, S. Sharma, J. Tenebaum, "Map Induction: Compositional Spatial Submap Learning for Efficient Exploration in Novel Environment."
4. M. Kleiman-Weiner, M. Kryven, J. Tenebaum, S. Yu, "Adventures of Human Planners in Maze Search Task."
5. A. Rizzo, C. Bongiorno, C. Ratti, D. Theurel, M. Kryven, J. Tenenbaum, P. Santi, Y. Zhou. "Vector-based pedestrian navigation in cities."
6. Huys, Q. J., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S. J., ... Roiser, J. P. (2015). "Interplay of approximate planning strategies."
7. Tomov, M. S., Yagati, S., Kumar, A., Yang, W., Gershman, S. J. (2020). "Discovery of hierarchical representations for efficient planning."