**Cognitive Modelling Comparisons**

**COMP 40270**

**Lecturer: Fintan Costello**

**Student: Stephen Kohlmann**

**Student Number: 14204243**

**Date: 25/4/16**

**­Introduction**

The objective of this essay is to compare my cognitive model for categorization with two models submitted by other students in the class. My model outlay was logical in its predictions and utilized some basic computation to yield results. With this in mind the two models I have chosen for comparison are different from my approach.

The first of the models approaches categorization with added mathematical complexities while the second model approaches categorization theory with minimal mathematical complexity. It is important to note that mathematical complexity does not mean that a model is better or worse in it’s predictions. The analysis presented of the models below is to point out the advantages and disadvantages of each approach followed by a comparison with future suggestions for improvement.

The layout of the essay will start with an analysis of my model theory. The advantages and disadvantages will then be discussed in detail. From here I will describe the theory of Diego’s model followed by the advantages and disadvantages of his model. Continuing on from here Fiona’s modeling theory will be discussed and the advantages and disadvantages of her approach. Finally, a comparison of the models will be presented showing similarities and differences that the models produced.

**Model Theory – Stephen Kohlmann**

In my model I proposed that participants used what they had learned from the example items to count the number of times a symptom occurred for each test item. For example category C has 6 lines in the long-term memory model. When a participant was shown a test item e.g. CYB the participant would compare the long-term memory model for each dimension. In Category C symptom C shows up 6 times on dimension 1, symptom Y shows up 2 times on dimension 2 and symptom B shows up 1 on dimension 3.

I believe that the amount of occurrences of each symptom directly affects the participant’s decision to whether a test item is a member of a category or not. If there is no entry on a dimension this also affects the participant’s decision. To model this scenario I gave each dimension entry an even weight that was then multiplied by the number of occurrences to find the degree of membership.

A parameter was also created to model my theory of zero on a dimension affecting a participant’s choice. To describe this in a simple way if each dimension for a test item was populated with a number > 0 than it would be more likely to be chosen as part of the category. If two dimensions were populated with a number > 0 but one dimension had 0 than it would be less likely less likely chosen as part of the category and finally if one dimension was populated with a number > 0 and the other two dimensions had 0 than it would be highly unlikely to be chosen as part of that category.

If a number > 0 was present in a dimension the value of 10% per dimension was assigned. The parameter therefore had a max value of 30% and was added to the degree of membership results. This sum was then compared to the average membership score of the model to identify if the model theory matched the results given for the assignment.

**Advantages of Model\_1**

The model has some distinct advantages that I will describe below. The model takes a very realistic approach to what happens when learned items are used for categorization. Personally if I were given the example items to learn I would do my best to remember them and then compare what I had committed to my long-term memory to the new test item. I would count the symptom occurrences in the long-term memory model for each dimension and the more times the symptom occurred the more likely I would think it would be a category member. This can be broken down into two steps, the first is learning off and the second is comparisons.

The model uses even weights to turn the concept described above into numbers without any bias or added weight attributes. The numbers associated with the degree of membership and parameter ensures calculations are taken only by the amount of occurrences with no added importance for a particular dimension. The model uses a 90% max rate for the degree of membership measurement.

This is not necessarily an advantage but does make the degree of membership numeric value easier to read. The details of the degree of membership are explained in the cognitive model power point presentation in more detail but in summary depending on the category 2.5%, 3% or 5% was assigned to the formula.

The model follows a logical structure in how categorization occurs and in general fits the results provided. The model displays the results of the parameter + degree of membership score compared to the average membership score provided. The results are displayed in scatterplots with a line of best fit that shows a positive correlation for most categories. Another advantage is that a scatterplot is completed for each category as well as combined results. This helps narrow down where the model seems to not quite match a particular test item.

**Disadvantages of Model\_1**

My proposed cognitive model is the same for both single and conjunctive categories. Although it is possible that participants take a similar approach nothing has been added to the model to address the complexity of conjunctive categorization or the conjunction fallacy. The scatterplot results for each conjunctive category show a positive correlation but the results are further away from the line of best fit than the single categories. The degree of membership weights of 2.5%, 3% and 5% are weighted evenly according to the amount of entries for each category. The issue here is that if a new line was added for each category then the figures above would not be even for the degree of membership.

There is away around this using Excel or any programming language to calculate the amount of lines per category and then add the correct number to multiply to the formula. For example category c has 6 lines and 3 dimensions, if all dimensions were populated with 1 for dim 1 this would result in 30% (6X5% = 30%). If the long term model is changed for another test by adding an extra four lines the hard coded 5% would need to be adjusted to 3% to gain the same 30% result (10X3% = 30%). In the current model this is not taken into account but would be a good future improvement.

The parameter of 10% is not adjusted according to the degree of membership percentages of 2.5%, 3% and 5%. This can be looked at in two ways, firstly that the parameter is completely separate and should have no bearing on the amount of records in the degree of membership. The alternative is the parameter should change accordingly to reflect the amount of records and degree of membership changes. For example the parameter of 10% is associated with category c, which has a degree of membership of 5% and is calculated accordingly. When we move to conjunctive category A&C, which has 12 records the degree of membership is divided in half resulting in 2.5% but the parameter of 10% is not adjusted accordingly. The question here is should the parameter be adjusted accordingly to 2.5% degree of membership?

The model does not use Z-scores. The use of Z-scores would have yielded a clearer result between the raw score and the distribution of scores. It may have been easier to see how good or bad the scores were compared to the entire group. The parameter + degree of membership is on a scale of 0 – 90% while the average membership score is on a scale of -10 to 10. Z-scores would have allowed the comparison of two different scales with more accuracy.

**Model Theory – Diego**

Diego looked at single instances first before looking at conjunctions. Diego investigated the long-term memory model and found where certain probabilities of symptoms were occurring. The model attempted to reflect the diagnosis of diseases that occur in the real world. The theory of the model is based on probability theory utilising the Naïve Bayes classifier. The model has some adanced mathamatical concepts but they are easy to follow which I hope to show in the advantages section below.

**Advantages of Model\_2**

The single category and conjunctive categories are treated separately at the start of the model explanation. This is an advantage as it allows for issues such as the conjunction fallacy to be addressed later on in the model. The model is built on a common machine learning technique called the Naïve Bayes classifier. This is a popular baseline method for text categorization and is an advantage as it is a clear way of converting the model theory into a cognitive model with numeric data.

The numeric output of the model is converted to percentages that make the model easier to follow and understand. Although the Naïve Bayes classifier is not overly complex it can be difficult to understand and the conversion makes it possible for the majority of people to follow the models output. The model also utilizes Laplacian correction removing null values from the model calculation. This is important, as any null values present would have made the multiplication of the dimensions values return zero. Finally the probability calculations are converted to match the -10 to 10 scale of the average membership score from the test results. Matching the scale is a great advantage for delivering a clear result set and comparison.

Another advantage of this model is that the scatterplot shows a positive correlation to the test results of the participants. Although not a perfect match it does show a general pattern that the participants were approaching the task closely to the model theory description. The model is slightly adjusted for the conjunctive categories to take into account for each category showing up twice in the conjunctions.

**Disadvantages of Model\_2**

Single categories and conjunctive categories are treated separately in the explanation of the model but the model is very similar for both categories. The model calculation was adjusted for conjunctive categories but the actual model is the same. The issue here is that conjunctive categories are more complex and the model does not take this into account.

There are two scatterplots one to show single category comparisons and another to show conjunctive category comparisons. To make the result set clearer having a scatterplot for each individual category would have been beneficial. This is simply on an aesthetic level for the presentation whereby it would have been easier to identify the particular test items that varied from the line of best fit.

The main disadvantage of the model is that it does not take into account possible crossovers for each category. For example a participant may have used the conjunctive category A&B when counting symptoms for single category A or single category B. I believe not addressing this is a disadvantage within the model.

**Model Theory – Fiona**

Fiona’s method for coming up with her cognitive model was interesting. The approach was to physically cut up the categories from the long-term memory model and then assess the categories from here. The idea behind this was to organize the categories into relevant heaps and then compare the dimension and symptom entries. From here the most common and least common symptoms for each category were calculated. Each category was then assigned the most common occurrence of a symptom and the dimension of that symptom. This categorization was then compared to the participants’ answers from the test items.

**Advantages of Model\_3**

The model starts with physically cutting up and heaping the data into relevant piles for categorization. This is an advantage as it helps to breakdown and understand the complexity of the assignment. Another advantage is the model looks for patterns of repetition which models what the participants would have done in the learning stage.

Each category is broken down to show the most common symptoms. This helps with representing commonalities that can be later used for comparisons on the test results from the participants. This also has the benefit of modeling what a participant would do when starting to assess the test items.

**Disadvantages of Model\_3**

There are some disadvantages with this model that I will list below. The model theory is not translated into a mathematical representation. This is a disadvantage as it is a pre-requisite for a cognitive model. The theory of the model needs to be converted into numbers to become a cognitive model. Mathematical equations appear later in the model but there is no explanation of the calculations or formulas.

The model results do not include a scatterplot or line of best fit. The disadvantage with this is that there is no clear representation of the comparison of the model to the participant results. The scatterplot would show if there was a positive or negative correlation and without this or a clear listing of comparisons it is impossible to know the accuracy of the model. There is no parameter or weight associated with the dimensions or symptoms again making the model results unclear.

Finally the same model is used for analyzing the single categories and conjunctive categories, which does not take into account the more complex decision process for participants in conjunctive categories. There is no clear answer to whether or not the model fits the participants’ results.

**Comparisons**

Fiona, Diego and I start out the model theories in a similar way by counting each symptom occurrence for each category and dimension. The three models attempt to replicate a real world scenario of how a participant would attempt this experiment. My model is slightly different as I also use the conjunctive categories as possibilities in the single category classification.

The models differ in mathematical complexity with Diego’s model been the most complex. Interestingly comparing Diego’s model results to my model results show similarities and both show a positive correlation. Although both our methods of turning our model theory into a cognitive model differ they do show similarities. This is interesting because it suggests that there is many ways to come up with similar results. Fiona’s model has a numeric output in the slides but the numbers are not explained clearly.

Both Diego and Fiona do not have an additional parameter in there cognitive models. My model differs as I do have an additional parameter added of 10% for each dimension where the value is > than 0.

Diego’s model makes a minor adjustment for the conjunctive category analysis but apart from this all three models do not change to take into account the difficulties of conjunctive classification. I believe with further research the current models could be built upon to take conjunctive classification issues into account.

Both Diego’s model and my model compare with the average membership score showing the results in scatterplot diagrams. The scatterplot diagrams show the line of best fit and in both cases show a positive correlation. Diego’s model differs as he uses probability theory for the mathematical component of his model. Due to this Diego needed to remove all possible zero entries and he did this by using Laplacian correction. Diego also made sure the scale of his model result and the average membership score matched before building his scatterplot diagrams.

My model does not change the scale but I do have a scatter plot for each single and conjunctive category. Also included is a combined single category scatterplot, combined conjunctive category scatterplot and one scatterplot including all categories. This is mainly beneficial for the presentation and assessing the results to see the test item results that are further away from the line of best fit. Fiona’s model does not have any scatterplot or diagrammatic comparisons of her results to the average membership scores provided. This is a problem as it is very difficult to see any clear comparison of the proposed model to the average membership score of each participant.

**Conclusion/Future Work**

I believe that every cognitive model has a weakness when it is compared to a group of participant results for categorization. It is difficult to pinpoint an individuals thought process and using one cognitive model for all participant results is presuming that every participant has the same approach. In my opinion getting a positive correlation to a result set is a numeric way of saying “the participants might be using this method”.

To pursue this project further I would firstly address what test items are furthest from the line of best fit. I would take these test items and look for a pattern or multiple patterns trying new models to find a match for these anomalies. I found this assignment interesting and the challenge of taking the abstract question of “how are people doing this?” and converting it to numbers to create a cognitive model. I found it helpful to see other student’s approaches and the variation of ideas presented.