Introduction to Artificial Intelligence ENGR-3720U

Project Part B

Created by

Daniel Smullen

Joseph Heron

Jonathan Gillett

09/04/2013

Document Revision 1.0

Table of Contents

Fuzzy Expert Marriage Decision Maker System

In this project phase, we use the MATLAB fuzzy toolkit to design an expert system which makes decisions for evaluating whether to marry or not marry a potential mate.

Introduction

Mamdani type fuzzy inference was used with a bank of 317 expert rules along 5 main factors. In combination, these factors and expert rules evaluate to a binary decision – that is, two different prospects – to break up, or to marry.

Defining Linguistic Variables and Determining/Encoding Fuzzy Sets

All linguistic variables used in the decision making process map to a spectrum from 0 to 1, as qualitative attributes (especially those such as beauty, friendship, and personal characteristics) cannot be quantified numerically. This method of encoding was chosen in order to create a simplistic model, where each attribute has a roughly qualitative value in relation to the others when quantified along the spectrum.

Each attribute was divided into several membership functions with a qualitative descriptor that maps to a logical spectrum of fuzzy ranges for each attribute. All membership functions for the fuzzy linguistic variables used triangular membership functions to greatly enhance the performance and simplicity of the system. The breakdown of membership functions per linguistic variable is detailed below.

Education Level

This attribute was divided up into 6 different membership functions, corresponding to common education brackets in Canada. Regions for each membership functions were decided based on subjective factors, with large overlapping areas. Brackets representative of higher education, such as College, University, and Post-Graduate were highest on the spectrum, with membership functions skewed toward 1. This is visible in Figure 1.

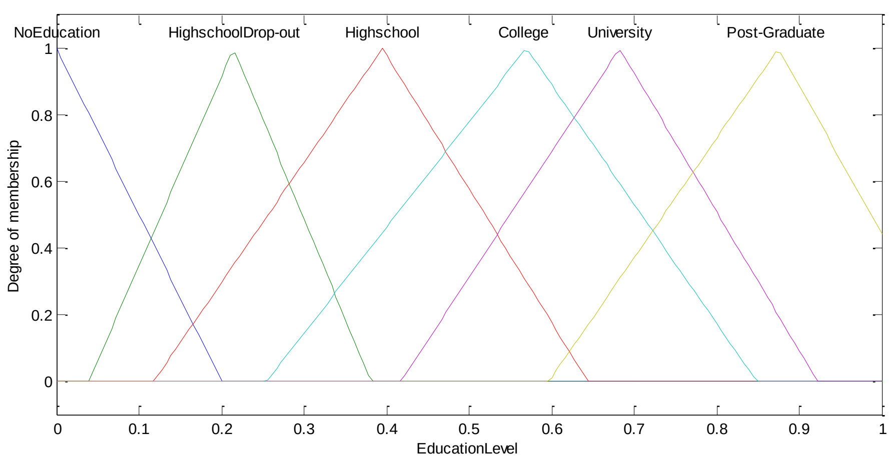

Figure 1 - Education Level Membership Functions Plot

Table 1 lists the membership functions that break down the spectrum of fuzzy values.

Membership Functions

The leftmost column shows the name of the membership function, which translates to a human-discernible qualitative value. The middle column shows the membership function characterization in terms of shape. On the right, the parameters for the function are listed. In the case of triangular functions, the function parameters correspond to a 1×3 vector containing: the left x-intercept for the first value, $x \mid f(x) = 1$ for the second value, and the right x-intercept for the third value, respectively.

All membership functions in the Education Level attribute are triangular. This was chosen to minimize computational complexity.

Table 1 - Education Level Membership Functions

Education Level	Membership Function Type	Membership Function Parameters
No Education	Triangular	$[-0.4\ 0\ 0.2]$
High School Drop-Out	Triangular	$[0.0416\ 0.202\ 0.384]$
High School	Triangular	$[0.1178\ 0.3948\ 0.6438]$
College	Triangular	$[0.2543\ 0.5693\ 0.8483]$
University	Triangular	$[0.4179\ 0.682\ 0.922]$
Post-Graduate	Triangular	$[0.5969\ 0.8749\ 1.099]$

Financial Level

This attribute was divided up into 5 different membership functions, corresponding to what we selected subjectively as 5 representative financial brackets. This group of membership functions is also right-skewed, with Middle Class, Rich, and Filthy Rich making up the top three brackets. This is visible in Figure 2.

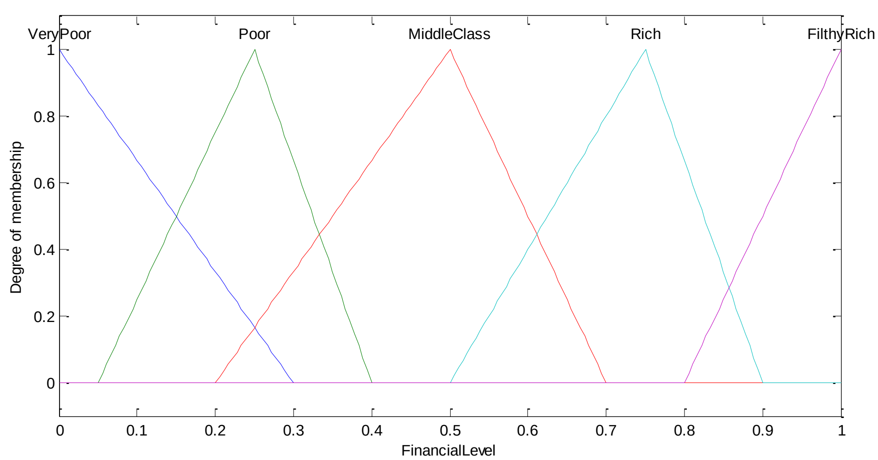

Figure 2 - Financial Level Membership Functions Plot

It should be noted that the skew actually favors Middle Class and Rich in terms of the widest fuzzy brackets, making Filthy Rich a more exclusive category, relatively speaking. Similar to Education Level, the Financial Level attribute uses only triangular membership functions and can be seen in Table 2.

Table 2 - Financial Level Membership Functions

Financial Level	Membership Function Type	Membership Function Parameters
Very Poor	Triangular	[0 0 0.3]
Poor	Triangular	[0.05 0.25 0.4]
Middle Class	Triangular	[0.2 0.5 0.7]
Rich	Triangular	[0.5 0.75 0.9]
Filthy Rich	Triangular	[0.8 1 1]

Beauty

This attribute was also divided up into 5 different membership functions, corresponding to 5 qualitative descriptors for the spectrum of beauty. This group of membership functions is roughly equally distributed, with only a slight right skew for the top three brackets. This is visible in Figure 3.

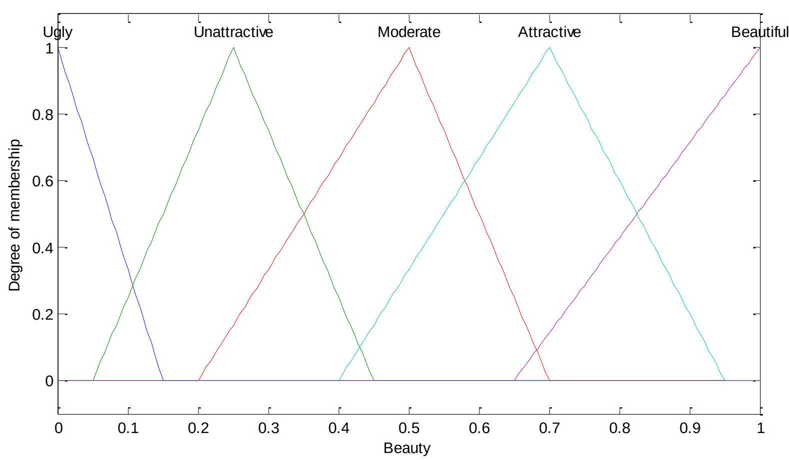

Figure 3 - Beauty Membership Functions Plot

Beauty also uses only triangular membership functions, and can be seen in Table 3.

Table 3 - Beauty Membership Functions

Beauty	Membership Function Type	Membership Function Parameters
Ugly	Triangular	$[-0.4\ 0\ 0.15]$
Unattractive	Triangular	[0.05 0.25 0.45]
Moderate	Triangular	[0.2 0.5 0.7]
Attractive	Triangular	[0.4 0.7 0.95]
Beautiful	Triangular	[0.65 1 1]

Friendship

This attribute was divided up into 5 different membership functions, corresponding to 5 different qualitative descriptors for level of friendship, that is to say, how well you know the person.

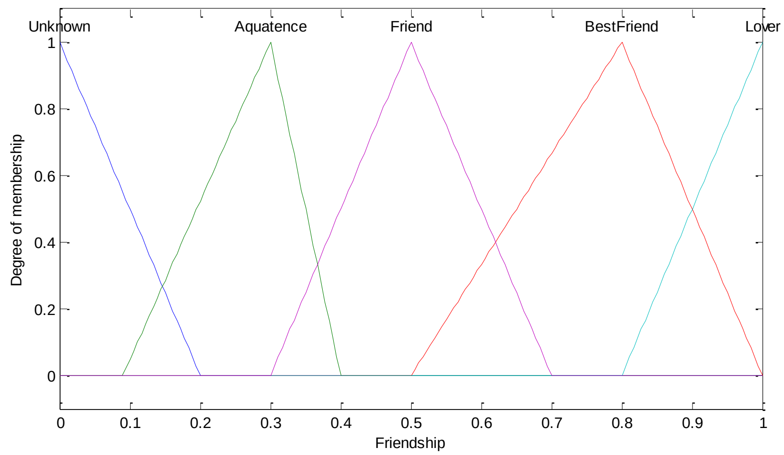

Figure 4 - Friendship Membership Functions Plot

This group of membership functions is right skewed all the way down to the leftmost membership function, which is directly visible in Figure 4. Friendship uses only triangular membership functions, and can be seen in Table 4.

Table 4 - Friendship Membership Functions

Friendship	Membership Function Type	Membership Function Parameters
Unknown	Triangular	[0 0 0.2]
Acquaintance	Triangular	[0.09 0.3 0.4]
Friend	Triangular	[0.3 0.5 0.7]
Best Friend	Triangular	[0.5 0.8 1]
Lover	Triangular	[0.8 1 1]

Personal Characteristics

This attribute was taken to represent the sociability of the prospective mate, and was divided up into 5 different membership functions, corresponding to 5 different levels of sociability.

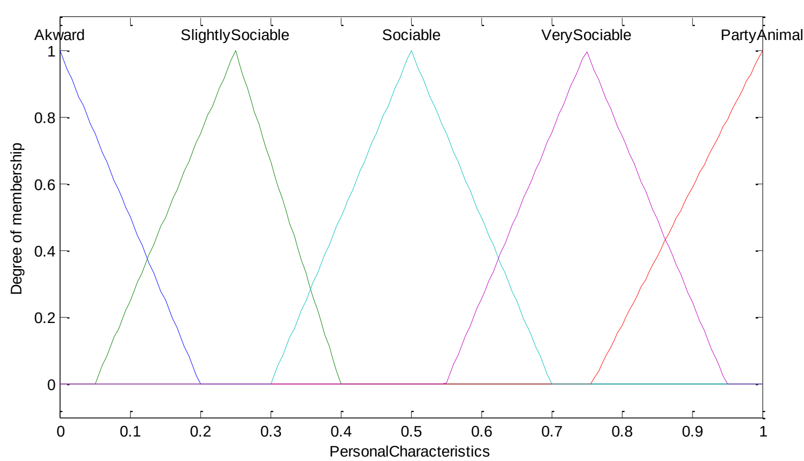

Figure 5 - Personal Characteristics Membership Functions Plot

This group of membership functions is evenly distributed, and uses only triangular membership functions, visible in Figure 5. The parameters can be seen in Table 5.

Table 5 - Personal Characteristics Membership Functions

Personal Characteristics	Membership Function Type	Membership Function Parameters
Awkward	Triangular	[0 0 0.2003]
Slightly Sociable	Triangular	[0.05 0.25 0.4]
Sociable	Triangular	[0.3 0.5 0.7]
Very Sociable	Triangular	[0.5491 0.7491 0.9491]
Party Animal	Triangular	[0.7572 1 1.4]

Overall Prospects

The resultant prospect is a binary decision – to marry, or to break up with your potential mate. This decision is divided up into two equal regions using triangular membership functions, seen in Figure 6.

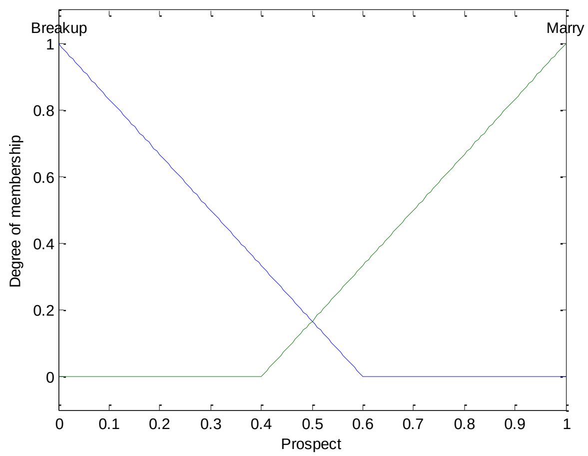

Figure 6 - Overall Prospects Membership Functions Plot

The parameters can be seen in Table 6.

Table 6 - Overall Prospects Membership Functions

Prospect	Membership Function Type	Membership Function Parameters
Break Up	Triangular	[0 0 0.6]
Marry	Triangular	[0.405 1.01 1.01]

Eliciting and Constructing Fuzzy Rules

This section describes the process undertaken to create the 317 rules that were required to cover all decision-making processes within the fuzzy expert system.

Team Roles
Since there were 3 individuals working on the project, all members of the project team were required to fulfill each of the usual team roles for an expert system. That is to say, each member at one time had to be project manager, programmer, knowledge engineer, and domain expert. These roles were switched in an ad-hoc arrangement until all the rules were created.

Multi-Tier Expert Rule Elicitation
By using 'expert' input, and going through selected combinations of the fuzzy sets that span each selection suitable for marriage, rules were generated to determine which attributes' fuzzy sets are suitable for marriage overall. These were connected using AND connections. AND connections were chosen because these rules were designed as exclusive rules.

The remaining combinations that contained attributes that were selected as undesirable were connected using OR connections, and result in breaking up. OR connections were chosen because the highest value would be factored in most, since OR connections are inclusive.

The result of this organization of connections was that important characteristics (notably, attributes with many rules such as Personal Characteristics) take major precedence, so if your selected mate has poor characteristics overall in one important field, you won't marry them.

The process of eliciting these rules, or 'expert' input, primarily consisted of asking directed questions toward our group members attempting to determine their preferences. In doing so, a multi-tier approach was used. The first tier of questions addressed the overall importance of individual attributes. This data was written down and used as a basic evaluation of the suitability of the fuzzy membership functions, and saved for later. Then, each combination of rules was evaluated, creating a pairwise evaluation toward each attribute. This was done by asking directed questions toward each combination - for example, are both sociability and beauty equally important? Would you marry someone who was poor if they were well educated? What if they were filthy rich, without any education? Finally, after all combinations were covered, the rules were aggregated and entered into the system.

Performing Fuzzy Inference
Using the fuzzy toolbox, fuzzy inference is automatically calculated given a sample input vector, where each value represents a sample attribute. Since the spectrum is from 0 – 1, all values within this range are accepted. Figure 7 shows an example where 30 rules are evaluated and fuzzy inference is conducted for each attribute. Regions highlighted in yellow show the fuzzy set that captures the values selected which correspond to the input vector. This is automatically performed. Crisp input values are used, fuzzified (by numerical integration and evaluation of the membership functions for each attribute) and the rules are evaluated. The aggregation of the rule outputs is then performed across all rules, and the resultant values are defuzzified and displayed across the top row of the diagram. It should be noted that Figure 7 also demonstrates that the Mamdani-style rule evaluation in this instance uses clipping for faster performance.

Discussion

Given the number of rules required for the system to complete evaluation and fuzzy inference, as many performance-enhancing choices as possible were chosen. Even in spite of this, performance when plotting out the inference across all of the rules is slow, and takes several seconds to a minute to plot all of the clipped membership functions.

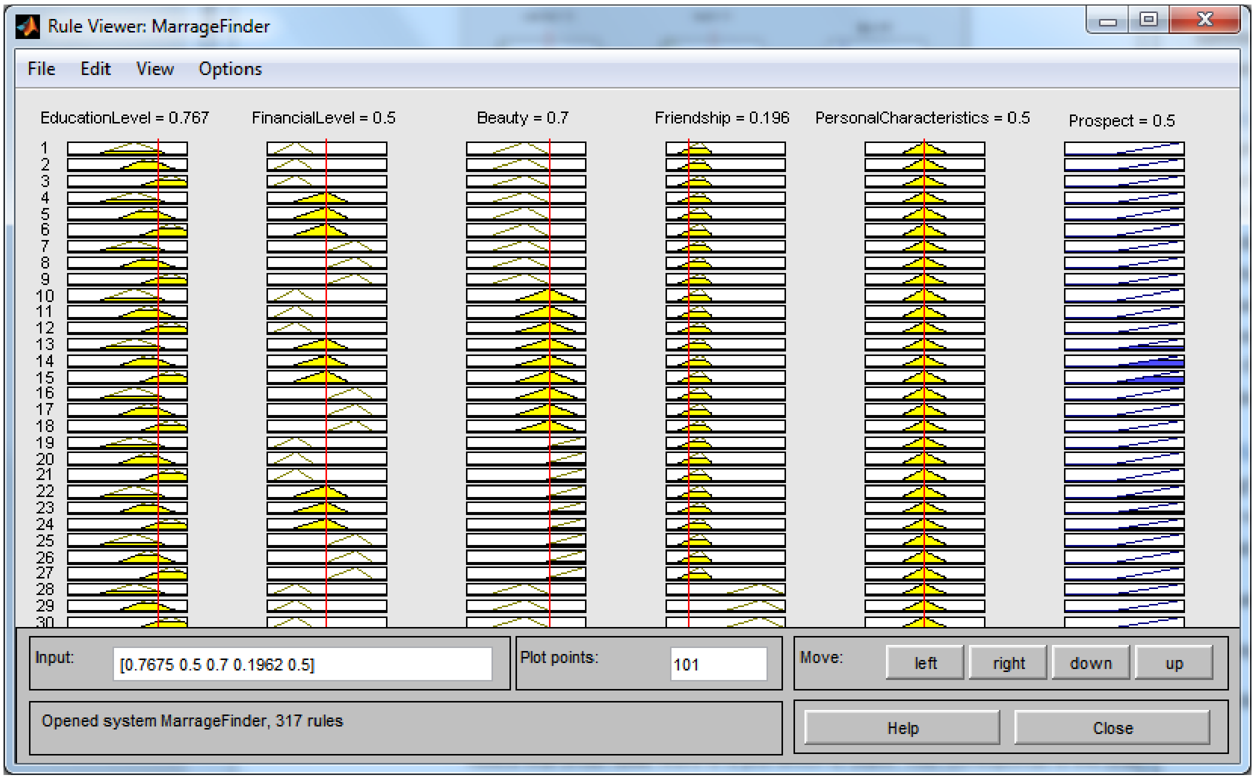

Figure 7 - Performing Fuzzy Inference

Evaluating and Tuning

In order to evaluate each rule and tune the system overall, individual rules had to be evaluated and examined, and compared to the original 'expert' input. The first-stage data from the multi-tier rule elicitation process was used to make fine adjustments to the membership functions for each attribute, and the resultant contour was evaluated and plotted in order to visualize the outcome of the fine-tuning. An example can be seen in Figure 8. In this example, the membership functions for Financial Level and Education Level are compared pairwise and the resultant prospect evaluated. It is evident in this resultant contour that by shifting the membership functions for both functions to the right, the Education Level membership function also shifts the resultant prospect optima to the right. The optimum that is present in this diagram occurs because of the rule relationship with each of these attributes' membership functions. Adjusting the rules to place greater emphasis on Financial Level would distort the optima plateau in Figure 8 such that the contour extends further toward the left on the Financial Level axis, removing the fall-off on that part of the plot.

Discussion

In summary, the conclusion drawn towards this example are that adjusting the rules determines whether the emphasis is placed on one attribute over another coarsely, and how well that attribute 'fits' the membership function improves the finer accuracy of the evaluation.

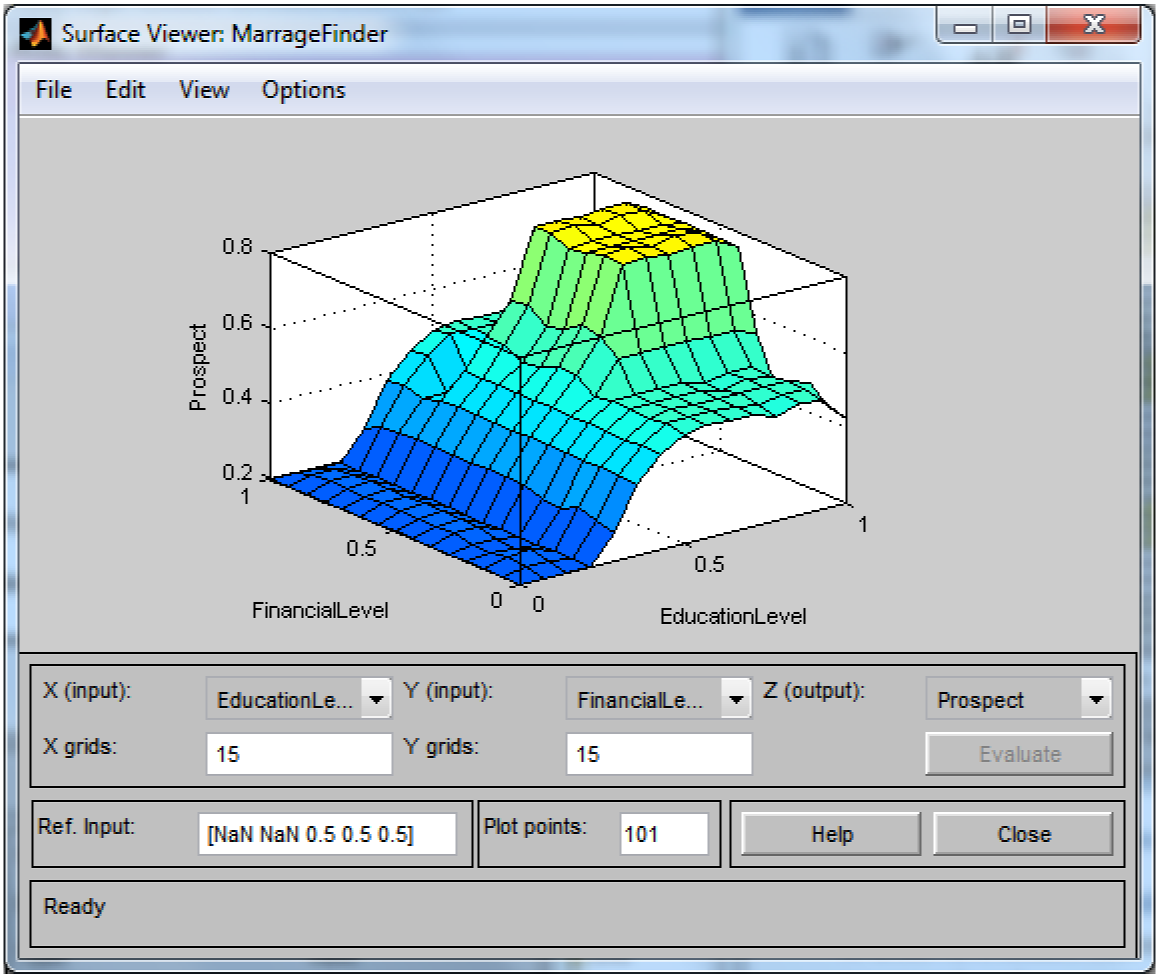

Figure 8 - Plot of Membership Function Contour

Testing

Testing was conducted by generating a random input dataset that roughly corresponds to a mixed population of potential mates. Each candidate for testing was separately evaluated manually to determine whether the candidate is a realistic portrayal of a prospective mate. Hand calculations were conducted for a few rules picked at random to evaluate whether the rules were accurate and that MATLAB was producing the output we expected.

After the test candidates were evaluated for correctness, they were independently evaluated by group members to determine what their overall suitability was; that is to say, to determine whether the expert system coincided with the 'expert' opinion of our knowledge. The sample candidates that were used for testing can be found in Table 8, located in Appendix A. Corresponding output data was generated by programmatically evaluating the rules using the `evalfis(A,MarriageFinder)` command in MATLAB. The resultant output dataset is located in Table 9, in Appendix B.

Conclusions

It is interesting to note that even in a wide variety of tests, it seems that our 'expert' knowledge was very selective towards potential mates, as overall only 1 in 20 or 5% of the acceptable candidates were suitable for marriage, and most candidates scored very poorly. The mean overall was 0.29385 with standard deviation 0.15195 which falls quite short of the membership function for marriage.

Our resultant analysis deemed this successful, as the purpose of the marriage selection process is to be choosy, after all! Even though the actual evaluation of the potential mates uses quantifiable (and fuzzy) evaluations of their attributes, overall the optimal candidate can be seen in Table 7.

Table 7 - Optimal Candidate Attributes

0.649115	0.731722	0.647746	0.450924	0.547009

This particular candidate scored particularly high in the education, financial level, and beauty categories, average in the social characteristics and poorly in the friendship category. This implies that our optimal marriage partner would be highly educated (University), filthy rich, beautiful, a friend (but not a close friend), and sociable (but not very sociable).

Further analysis suggests that given the latter two attributes, an arranged marriage with a beautiful rich bride may be in the cards. Perhaps this speaks to an unrealistic expectation. By increasing our own Education Level attribute beyond University (and our Financial Level to filthy rich) we may be lucrative enough partners to find someone with these attributes!

Appendix A – Sample Dataset

The following dataset was used to represent a sample group of 20 selections for potential mates, with 'realistic' values.

Table 8 - Sample Input Data

0.538342	0.996135	0.078176	0.442678	0.106653
0.961898	0.004634	0.77491	0.817303	0.868695
0.084436	0.399783	0.25987	0.800068	0.431414
0.910648	0.181847	0.263803	0.145539	0.136069
0.869292	0.579705	0.54986	0.144955	0.853031
0.622055	0.350952	0.51325	0.401808	0.075967
0.239916	0.123319	0.183908	0.239953	0.417267
0.049654	0.902716	0.944787	0.490864	0.489253
0.337719	0.900054	0.369247	0.111203	0.780252
0.389739	0.241691	0.403912	0.096455	0.131973
0.942051	0.956135	0.575209	0.05978	0.23478
0.353159	0.821194	0.015403	0.043024	0.16899
0.649115	0.731722	0.647746	0.450924	0.547009
0.296321	0.744693	0.188955	0.686775	0.183511
0.368485	0.625619	0.780227	0.081126	0.929386
0.775713	0.486792	0.435859	0.446784	0.306349
0.508509	0.510772	0.817628	0.794831	0.644318
0.378609	0.81158	0.532826	0.350727	0.939002
0.875943	0.550156	0.622475	0.587045	0.207742
0.301246	0.470923	0.230488	0.844309	0.194764

Appendix B - Output and Result

Table 9 - Sample Output Data shows the actual crisp output data from the expert system on the left hand side, and the interpretation of that data into a decision on the right. This data is gleaned from the membership functions illustrated in Figure 6 - Overall Prospects Membership Functions Plot.

Table 9 - Sample Output Data

0.196808	Break Up
0.4125	Break Up
0.19728	Break Up
0.197761	Break Up
0.447925	Break Up
0.320606	Break Up
0.201085	Break Up
0.207068	Break Up
0.204186	Break Up
0.196796	Break Up
0.197969	Break Up
0.198862	Break Up
0.79314	Marry
0.211533	Break Up
0.198587	Break Up
0.313038	Break Up
0.483894	Break Up
0.236413	Break Up
0.465577	Break Up
0.198672	Break Up