

AI DATING

Development of a Novel Dating Application
with Fuzzy Inferencing

Roderik Frederik Emmerink

MSc Computational Intelligence and Robotics / De Montfort University / Computing Sciences and Engineering

AI Dating:
Development of a Novel Dating Application with Fuzzy
Inferencing

Roderik Frederik Emmerink
MSc Computational Intelligence and Robotics
Faculty of Computing Sciences and Engineering
De Montfort University

September 12, 2008

Abstract

This report describes the development of an online dating application with the use of AJAX technology. Within the dating application a fuzzy expert system is developed that is able to make inferences based on the OCEAN model. Based on the outcome of the OCEAN matching and some classical inferences the system is able to present matches. The work results in a prototype that is able to match users and that can be adjusted and extended to result in a novel dating application with sophisticated matching by means of computer mediated communication.

Contents

1	Introduction	2
2	Background	4
2.1	Predictors for a Good Match	4
2.2	Fuzzy Logic	6
2.2.1	Fuzzy Sets	6
2.2.2	Fuzzy Operators	6
2.2.3	Fuzzy Inferences	7
2.3	Related Work	8
2.3.1	Online Dating	8
2.3.2	Fuzzy Inferencing	8
3	Analysis and Design	10
3.1	Fuzzy Matchmaking	10
3.2	User Interaction	11
3.3	Privacy and Security	12
3.4	Storing User Data and Inference Knowledge	13
4	Implementation	14
4.1	The Fuzzy Inferencing System	14
4.2	The Graphical User Interface	16
4.2.1	The Dating Assistant	17
4.2.2	Authentication	18
4.2.3	Photo Upload	18
5	Testing	20
6	Conclusions, Evaluation and Further Work	21
6.1	Reflection and Conclusions	21
6.2	Further Work	21
A	The Database Model	26
B	The Fuzzy OCEAN Model	27
C	The Project Planning	29

Chapter 1

Introduction

Online dating is highly popular. With many people being single and looking for a partner online dating can be an outcome before meeting someone. In (Scott, Mottarella, and Lavooy, 2006) it is stated that online dating provides more revenue than any other type of online sales.

Despite its popularity most dating websites are still limited in comparison to dating in real life or dating through dating agencies. This project aims to take online dating a step further and to overcome some issues with current dating sites.

The main problem with online dating sites is that they lack intelligence. Where dating agencies have specialists responsible for coupling people, at most online dating sites the user is responsible to pick a partner himself. The user at a dating site does not have the knowledge that a matchmaker has from a dating agency and might therefore not make an optimal choice. In this project it is suggested to partly overcome this by finding out what contributes to successful matches from the literature and by modeling this into an expert system.

Another weakness of dating sites is the user experience. Browsing hundreds of profiles with characteristics like length, hair colour, the preference to get children, etc. might bore the user. By using the expert system as suggested a selection can be made for the user reducing the profiles that are not interesting for him/her and providing him/her with matches of his/her interest. With the ongoing developments and the gradual move toward applications on the internet, more and more techniques become available to provide the user with a more rich experience. The project aims to use these new elements in a functional and enjoyable manner.

Since many dating sites make use of profile pages that can be filled in by the user, the information available is also highly dependent on the user. Some properties as age and gender might be available in almost every profile but other information like interests or hobbies might be very limited since it is up to the user to write this in his self-description. By creating a different more direct interaction with the user the project aims to acquire more knowledge on his personality or preferences than at the classical dating sites.

Where expert knowledge is usually very specific and when users are able to express their selves with linguistic terms there might be a possibility to infer relationships. Initially it was thought of to match users based on similarity or emergent relationships between characteristics like hobbies or preferences. Whether this is a fruitful approach or not can be investigated with the use of neural networks to do unsupervised clustering to derive different types of people that match well.

Expert knowledge is usually expressed qualitative with linguistic terms. Users are also used to express themselves linguistically and therefore it makes sense to make use of fuzzy logic to model expert knowledge and draw inferences based on that. Initially, the idea was to make a hybrid implementation

for matchmaking with artificial intelligence based on classical reasoning, fuzzy logic and the use of artificial neural networks (ANNs). During the project this appeared to be too much work for a three month period and it was decided to have the main focus on the implementation of a prototype for a dating site where fuzzy logic is applied in the matchmaking.

The report discusses the analysis and design of the dating application and the fuzzy inferencing in particular. Thereafter the implementation of the design into an online dating application is discussed in the implementation chapter. A small part is dedicated to testing the functionality and thereafter the result and possibilities for ongoing developments are discussed. Firstly, the information from literature on matchmaking, fuzzy logic and current dating sites will be discussed in the next chapter.

Chapter 2

Background

This chapter discusses some findings from social psychology according to what can be good predictors for a successful enduring relationship. Furthermore the chapter introduces the reader to fuzzy logic and its use within expert systems. In section 2.3 related work is discussed. Both current dating sites are examined and fuzzy inferencing is discussed.

2.1 Predictors for a Good Match

No dating site will be able to achieve good matchmaking without a good understanding of what constitutes a good match. Especially since the author is no insider in this area of research, some time has been spent on finding out which characteristics are desired to acquire from the user as a basis for matchmaking or which will be worthy to acquire in a later stage. Besides that, it has to be known how they logically combine in order to implement this knowledge in the form of fuzzy inferences, mathematics or classical logic.

Many people of which the author himself at the start of the project have the idea that the characteristics of a partner should not only have a degree of similarity but that people in a relationship should also complement each other in a substantial degree. Maybe this thought comes from the idea that one wants to feel unique and also appreciates the uniqueness in the other. However, the literature shows little or no evidence that personalities that complement each other will constitute a good or better relationship. Buss (1985) already expressed his doubts about complementarity: “The tendency of opposites to marry, or mate [...] has never been reliably demonstrated, with the single exception of sex”.

In (Myers, 2002) it is expressed that dissimilar attitudes turns us off more than that similar attitudes turns us on. Also he expresses influences upon liking and friendship which are: proximity for creating the right circumstances; physical attractiveness and similarity of attitudes, beliefs and values. When mating is considered it is noted that people usually choose a mate with similar attractiveness to their own attractiveness or someone that is less attractive but has other qualities as a compensation.

In “Human mate choice as the psychologist views it: Findings, issues, problems and perspectives” Wallisch (2002) describes determinants and predictors of human mate choice. The author presents five factors that are important determinants of interpersonal attraction and the quality and stability of close interpersonal relationships according to the literature about interpersonal attraction. These five factors are: physical attractiveness; personality characteristics; proximity; familiarity; and similarity.

Taylor, Peplau, and Sears (2000) findings in their research related to partner choice with similar patterns are described as: “In dating and marriage, the tendency to choose similar patterns is called

the matching principle. [...] Dating partners and spouses tend to be relatively matched not only in their values and attitudes, but also in their physical appearance, social background, and personality. For example, one study of dating couples found that partners tended to resemble each other in age, intelligence, educational plans, religion, physical attractiveness, and even height. They were also matched in their attitudes about sexual behavior and sex roles. Furthermore, couples who were most similar in background in the beginning of the study were most likely to be together 1 year and 15 years later.”

In (Fiore, 2004), Fiore discusses one of the biggest dating sites in the United States. His findings are that “For men, age, educational level, and self-rated physical attractiveness are the next most important qualities. For women, they are not being overweight, self-rated physical attractiveness, and having a photo. Finally, a discussion of the design implications of these findings and other design issues follow the results.”

According to Myers (2002): “Interpersonal attraction is increased by proximity, familiarity (the mere exposure effect), and similarity. We tend to like people who are similar to us in attitudes, values, interests, background, and personality”. Myers (2002) also discusses Sternberg’s three basic components of love, which are: intimacy (liking), passion (infatuation) and commitment (empty love). Depending on the combination of two of these they result in type of love. Romantic love is build up from intimacy and passion, companionate love is build up from intimacy and commitment and fatuous love is build up of passion and commitment. Sternberg uses a triangle to express this spectrum of love. “Research on heterosexual mate selection finds consistent male-female differences: Men place greater emphasis on a partner’s youth and beauty; women place greater emphasis on a partner’s economic and social resources. Both cultural and evolutionary explanations for these patterns have been offered.” (Buss, 1994)

According to Hill (2008): “Partner characteristics related to experiencing emotional closeness, psychological comfort, power, and physical and psychological pleasure are therefore likely to be factors that are attractive and lead to sexual arousal. In fact, all of these factors have been shown to relate to increased sexual attraction”. However these are merely factors that count in face-to-face dating and are difficult to use in online applications.

Hill (2008) also presents a table summarizing a research (Buss, Abbott, Angleitner, Asherian, Biaggio, Blanco-Villasenor, Bruchon-Schweitzer, Ch’U, Czapinski, Deraad, et al., 1990) about characteristics that are considered important in mates based on a research of 37 cultures around the world. There is a high similarity in what men and women look for in a mate. The four most important items are the same for men and women and are as follows: mutual attraction (love); dependable character; emotional stability and maturity; pleasing disposition. Then women rate education and intelligence as slightly more important than the other way around. Good health and sociability are also high rated. Many more factors are discussed and will be considered to be used in the matching algorithm.

The OCEAN model or FFM (Five Factor Model) Wikipedia (2008a) is one method to describe someones personality. Another method is the MBTI developed by Myers-Briggs. These two methods are among the most used by psychologists.

“The Five-Factor Model (FFM) is nowadays the personality structure which has reached the highest consensus among researchers” (Marco Perugini, 1997). It describes personality by means of five personality factors which are Openness, Conscientiousness, Extraversion, Agreeableness, and Negative Emotionality. (Michael C. Ashton, 2001) suggest a model very similar to the FFM that makes use of six factors of which four are also used in the OCEAN model.

2.2 Fuzzy Logic

Fuzzy Logic was proposed by Lotfi Zadeh who published about it in his article “Fuzzy Sets” (Zadeh, 1965). Since then it has been applied in many fields and has proven to be an efficient way to model linguistic knowledge. In a recent interview (Dumitras and Moschytz, 2007) in IEEE Signal Processing Magazine Zadeh attributes this big impact of fuzzy logic to its wide area of applications and its necessity for modeling reality in human-centric fields.

Fuzzy logic is an alternative to classical logic and is able to deal with linguistic or qualitative expressions. With the use of fuzzy logic it is possible to reason with vague terms. In classical logic elements either belong to a set or they are not part of it. Fuzzy sets contain all elements but till a certain extent. This is called the degree of membership.

2.2.1 Fuzzy Sets

Engelbrecht (2002) describes fuzzy sets: If X is the universe of discourse and x is an element in X then the membership function of a fuzzy set A is described by:

$$\mu_A : X \longrightarrow [0, 1] \quad (2.1)$$

The membership function describes the certainty that element x belongs to set A . Fuzzy sets can both be applied to discrete and continuous domains. A fuzzy set A is described in terms of its members and the corresponding degrees of membership. For a discrete fuzzy set A this will be:

$$\sum_{i=1}^n \mu_A(x_i)/x_i \quad (2.2)$$

For a continuous fuzzy set this will be:

$$\int_{i=1}^n \mu_A(x_i)/x_i \quad (2.3)$$

These notations come from fuzzy set theory and are distinct from algebraic notation.

A fuzzy set is modeled by a membership function. Some popular membership functions are the triangular, trapezoid and Gaussian membership functions. A membership function can be of any shape but must be normalized between 0 and 1. Operation on fuzzy sets are done with the use of special fuzzy operators.

2.2.2 Fuzzy Operators

Operators from ordinary logic have been adopted within fuzzy logic and are described in (Engelbrecht, 2002; Wikipedia, 2008b). Often these are named after the operators from set theory. The author however prefers to use the vocabulary from classical logic.

Two two-valued fuzzy sets are equal if and only if (iff) the sets have the same universe of discourse, and $\mu_A(x) = \mu_B(x)$ for all $x \in X$. A two-valued set A does contain a set B ($A \subset B$) iff $\mu_A(x) \leq \mu_B(x)$ for all $x \in X$. The complement (negation) of a fuzzy set consists of the same elements. The degree of membership for the elements in the complementary set is defined by:

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x) \quad (2.4)$$

The intersection (conjunction) of two fuzzy sets can be modeled in different ways. The so called Zadeh operators make use of the Gödel t-norm:

$$\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x)), \forall x \in X \quad (2.5)$$

An alternative is to use the product t-norm for a so called strong conjunction:

$$\mu_{A \cap B}(x) = \mu_A(x) \cdot \mu_B(x), \forall x \in X \quad (2.6)$$

The union (disjunction) of two fuzzy sets can be implemented with the use of the maximum t-conorm:

$$\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x)), \forall x \in X \quad (2.7)$$

A common alternative is to use the probabilistic sum:

$$\mu_{A \cup B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x) \cdot \mu_B(x), \forall x \in X \quad (2.8)$$

2.2.3 Fuzzy Inferences

With the use of the operators described and a fuzzy inferencing system it is possible to make fuzzy inferences based on linguistically described knowledge. This knowledge takes the form of fuzzy rules. Fuzzy rules are implications and take the form “IF antecedent THEN consequent”. Within a fuzzy system many rules can be defined and they all have their own influence on the consequents.

The components of a Mamdani fuzzy inferencing system are shown in figure 2.1. The crisp input will be translated to a fuzzy set in the fuzzification process. This fuzzification process extracts the fuzzy set based on the membership functions and non-fuzzy inputs. The inferencing process draws inferences by evaluating fuzzy rules. In the defuzzification stage the results from the fuzzy rules are aggregated into a crisp output value.

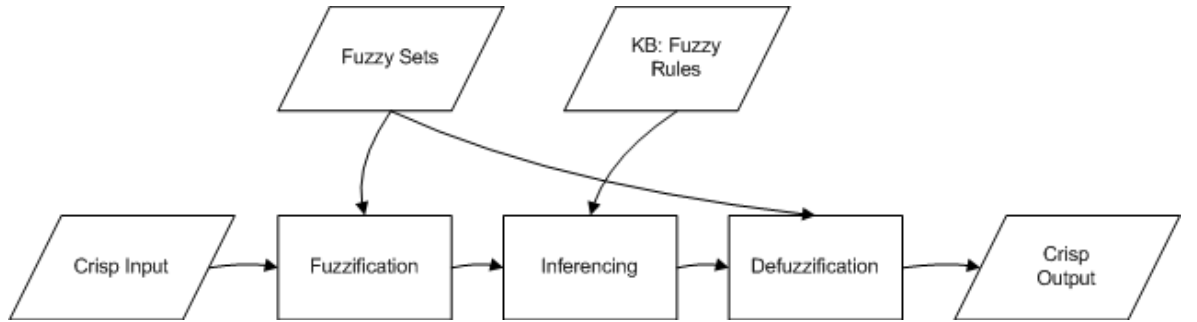


Figure 2.1: Schema of a Fuzzy Inferencing System

An alternative to the Mamdani system is the system proposed by Takagi-Sugeno-Kang (TSK). Although the TSK system is normally less computationally expensive it is less manageable because of the way it presents the knowledge. Jassbi et al. (2006): “In terms of use, the Mamdani FIS is more widely used, mostly because it provides reasonable results with a relatively simple structure, and also due to the intuitive and interpretable nature of the rule base.”

2.3 Related Work

Some dating sites already make use of AI and some related work on fuzzy logic in relation to personality has been done in the past. A few dating sites will be discussed that are either successful (well-known) or those that make use of similar technologies.

2.3.1 Online Dating

Online dating is undertaken by many people and much money is spent in this business. According to (Myers, 2002) Americans alone did spend 214 million dollar on internet dating sites in the first half of 2003. In online conversations people tend to disclose more, are more honest, and less posturing. Myers also cites some research that points out that romantic relationships formed on the internet last on average for at least two years. Whitty (2008) stresses some positive points of online dating: “[...] cyberspace presents a unique opportunity for people to relate online - an experience that is sometimes difficult or impossible to achieve in any other medium”.

There are a few websites that make use of some AI in their dating concepts other than in the form of classic expert systems. OkCupid (2008) uses a kind of statistical matching but might also use fuzzy logic to determine certain classifications on user profiles. Their way of gathering knowledge by doing quizzes and answering questions from an extensive library of questions (probably over a thousand) is perceived as enjoyable by some users, of which the author and someone he knows from that website. OkCupid is not fully focussed on people that look for relationships but seems to be targeted to a wide audience that want to spend time online and find people with similarities.

The Dutch web site PaiQ (2008) is a dating application that uses artificial intelligence as their main selling point. PaiQ makes use of artificial neural networks. How these are applied is not certain but one theory is that feedback after chatting for a while is used as a form of unsupervised learning and related to personality characteristics and habits that can be expressed while signing up. Besides that it has a photo rating module in which photos of two users are presented and one can be favoured over the other. This is not only a good way to get information on physical attractiveness but it is enjoyable as well. PaiQ limits the user in its possibilities and only assigns a match once in a while for non paying members. Photos of matches only become visible to a match after chatting for a while. This last feature is not always appreciated by its users and it can be debated whether it is advantageous or not. However PaiQs approach to online dating can clearly be considered as modern and innovative and the chatting feature that can also be used for non paying members is well appreciated. Opinions are based on the authors experience and that of a few other users that were contacted.

2.3.2 Fuzzy Inferencing

Ören and Ghasem-Aghaee (2003) did some research concerning the implementation of the OCEAN model with the use of fuzzy logic. In their article they describe how five traits, described by the OCEAN model, can be derived from 30 facets to determine personality types. These thirty attributes were distilled from contemporary psychology and each of them is called a personality facet. In their article they describe the personality descriptors based on levels of their facets that are either low, medium or high. Based on the outcomes for each type a personality type can be derived. Furthermore the article discusses different ways to graphically or linguistically represent personalities.

The same authors present in (Ghasem-Aghaee and Ören, 2003) a way to model these personality descriptors with the use of fuzzy logic. They define fuzzy sets for personality traits and process them with the use of fuzzy logic. A personality is described as a combination of the big five personality

traits: openness, conscientiousness, extroversion, agreeableness, and negative emotions. These five descriptors each have six facets that contribute to their expression in a personality.

“Linguistic variables were introduced by Zadeh (1973) and this term is used to describe some concepts that usually have vague or fuzzy values. The linguistic terms low, medium, high, are assumed to be from the term set for a linguistic variable like negative emotionality of personality traits and are interpreted as fuzzy subsets of some universe U of negative emotionality” (Ghasem-Aghaee and Ören, 2003).

Ghasem-Aghaee and Ören (2003) suggest the mapping of fuzzy sets by means of a vector following from equation 2.2. A triangular membership function for medium worry withing the domain of discourse 'degree of worry' $[0,100]$ is described as:

$$(0.0/10, 0.0/20, 0.0/30, 0.2/40, 1.0/50, 0.2/60, 0.0/70, 0.0/80, 0.0/90, 0.0/100)$$

For low and high worry trapezoid functions are used that reach till the end of the domain. Furthermore they describe the use of hedges but it is unclear whether they are used or not in their model. The fuzzy inferencing in the system follows a similar structure as shown in figure 2.1.

Ghasem-Aghaee and Ören (2003) derive personality descriptors with a 1 to 1 relation between the facet of the ocean model and its valuation low, medium or high and how it is described. This is modeled by fuzzy IF-THEN rules. Another set of rules is responsible to derive personality facets. For openness a person can be categorized into preserver, moderate or explorer till a certain degree. This is modeled by the conjunction of all aspects that constitute openness in one single rule for every category. Additionally they suggest that aspects can weighted which is not used in their implementation. Furthermore with another set of rules compound personality characteristics are determined. This is based on the personality factors determined in last step. In another stage personality modification are determined from emotional states by basic implications.

Chapter 3

Analysis and Design

The objective is to develop a dating application that uses AI for matching individuals. Artificial Intelligence is able to make computers do tasks that were otherwise be done by humans or tasks that require human intelligence. Matchmaking is such a task that originally needed the knowledge of experts, like psychologists, to give good advice on future relationships. The field of expert systems, especially where combined with techniques from artificial intelligence like fuzzy inferencing or neural networks provide great opportunities to make online dating applications more intelligent and therefore more capable of matchmaking.

In this project a dating application is designed that uses a modern interface and applies artificial intelligence in the form of fuzzy logic as part of the matchmaking process. This chapter describes the design of some components of the dating application and analyses how these can be implemented in a functional, effective and enjoyable way.

The prototype that is described is build up of different modules. First there is some knowledge from the literature. This knowledge is acquired and the accompanying user data is also acquired. A simple expert system can make an estimation based on these data for every combination of a female and a male single person for them to end up in a relationship.

Secondly fuzzy reasoning can be used to argue with interests and personality. A combination of fuzzy logic, mathematical inferences and classical reasoning can result in a match percentage that tends to distinguish good potential partners from worse ones.

Finally, much more can be done with data where there is no present relationship known. People can give lists of what they like. Since other people might not have the same things or have different wordings for those and since these lists can be enormous it is not easy to model the relationship between characteristics of different users. Clustering these linguistic terms with the use of artificial neural networks was considered at the start of the project but is not implemented because of its complexity and the limited amount of time available within this project.

The following sections discuss the considerations and design choices. Firstly the matchmaking process is discussed with its use of fuzzy rules. Secondly the proposed user interaction is analysed and a design is shown. Furthermore privacy and security is shortly discussed and choices are made concerning the storage of knowledge and user data.

3.1 Fuzzy Matchmaking

As described in 2.1 the OCEAN model can be used to describe personality. As a basis for the matchmaking the OCEAN model will be used to derive the five personality factors based on the

model that was proposed in (Ören and Ghasem-Aghaee, 2003; Ghasem-Aghaee and Ören, 2003).

Besides the OCEAN model the Myers-Briggs Type Indicator (MBTI) can be implemented in a later stage. A hybrid implementation to describe personality might reduce errors in classification, both because it relies on more input variables and because ideas from one or the other method for type indication might work out better. In this project the OCEAN model will be used. From a practical viewpoint this will be more easy to implement partly because a good basis is described by Ören and Ghasem-Aghaee (2003); Ghasem-Aghaee and Ören (2003).

A similarity measure of interests and attitudes will be used in addition. Matching based on similarity of preferences like hobbies or sports is seen on many dating sites. It makes sense that a spouse with similar interest will constitute a good relationship. This similarity in interest might both be efficient because you can combine strengths and enjoyable because you probably like to share your experiences in that area. Although most sites focus on sports, hobbies and going out, other interest can also be matched like interest in politics, certain magazines, television programs, films, and music. Although this similarity matching is mostly beyond the scope of the prototype it is of interest to implement in a later stage. It can also be thought of to make it more easy for the user by sharing data between different web services, like importing music preferences from Last.fm or importing film preferences from IMDB.

In addition matchmaking should take place based on classical parameters that are used in most dating sites such as age, gender and height. Depending of the character of these parameters an appropriate way has to be chosen to integrate it into the matching algorithm.

3.2 User Interaction

Where using websites or more in general web applications was a limiting factor in the past it provides less and less restrictions today. As a result it can be seen that more and more applications make the move from the desktop to interfaces on the world wide web. Dating sites do not always make use of these possibilities provided by modern web browsers.

The aim is to develop a dating application that provides an intuitive and enjoyable way for the interaction with the user. A novel design for a user interface is shown in figure 3.1. The implementation of this interface is beyond reach for the project but can be implemented partially or in full in a later stage. The graphical user interface design can be implemented with Adobe AIR or Adobe Flash. It shows the dating assistant on top that guides the user and asks questions for the system to obtain useful knowledge from the user to make choices on the matches the show in the interface. Based on the number of questions answered and the suitability of the match based on the algorithm the user obtains more possibilities for contact through chat, email or other means. Matches that are more similar move nearer to the user in that is shown in the middle. The matches represented by small photos with icons for interaction can move at low speed within a sphere to make the interface more alive/enjoyable. It can also be considered that certain parameters are used for the angle in which the match is project from the middle. For the implementation this is beyond reach but the general ideas can be used for the GUI (graphical user interface) for the dating application. These main ideas are that a distinction is made visible between matches that are more preferable and those that are less preferable. This can be done by means of a list.

Using the application needs to be both enjoyable and easy without losing its aim to result in serious relationships on the long term. Users must be able to provide the system with a photo and to see photos of users that match well since physical attraction is one of the main factors that constitutes a good match as it turned out in the literature review. Where possible the interface both aims to be

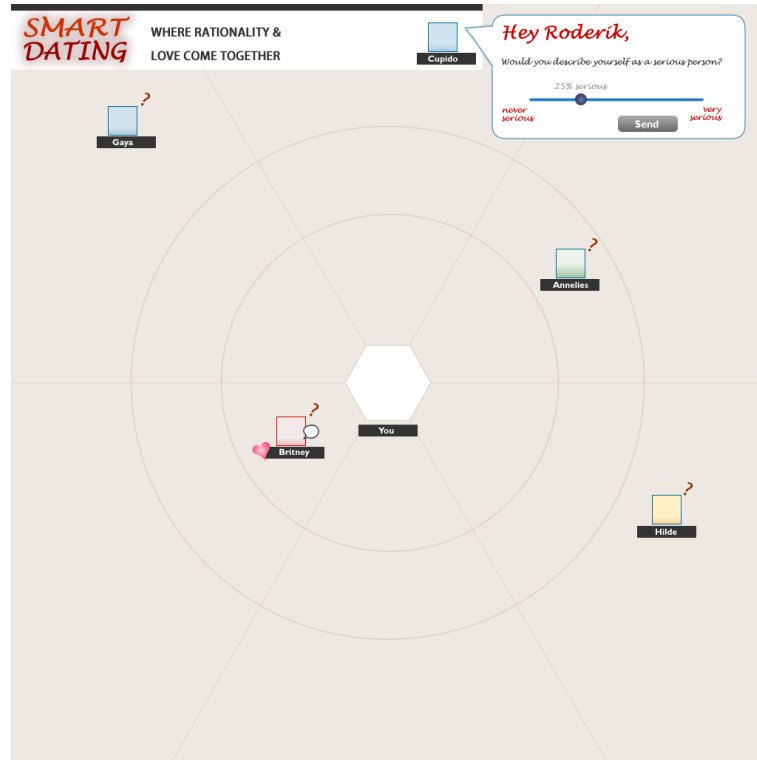


Figure 3.1: Design for a Graphical User Interface

modern, accessible and intuitive in its use. Also the user interface will be developed with extension with other modules in the future in mind.

3.3 Privacy and Security

Privacy and security are major concerns when applications are available through the world wide web and contain personal information. Especially in applications considering online personals the system has to deal with a lot of personal data. Although by signing up to a dating site the user is aware that he or she has to express himself, he or she must also have some trust in how these data is used.

In the first place the users are made aware that the data they submit can be used for purposes related to the aims of the website. By software means it will be made impossible to become a member/user without explicitly confirming that the data that is provided can become public domain.

However in practice not all this data will become into the public domain. The data will only be shown to other users at the website itself. Currently it is chosen not to show the surname and date of birth of the user to other users or any other visitor in order to make the user much harder to identify. Till what extent knowledge extracted from questions concerning preferences, personality and interests will be shown in user profiles is to be determined in a later stage. The user profiles, other than in the matches view, is not shown in the prototype.

Photos and other user data will only be accessible to users that are logged in and that have used the dating site recently. Even though it is possible to request photos directly by bypassing the user interface this will not result in information that is shown when someone is not logged in.

The login framework needs to be reliable and secure. It is decided not to use a secure protocol to transfer data. The data exchanged with the website is unencrypted and can in theory be intercepted but this is very unlikely to happen.

A full terms of service statement and privacy policy will be available when the website finishes the prototype stage and officially goes live. Users that sign up during the testing phase confirm that all their data can be used and can also request to redraw their data from the website if they do not like to have it stored online anymore.

3.4 Storing User Data and Inference Knowledge

The dating application has to deal with different types of data. Knowledge on how to infer matches from user data has to be stored, knowledge on how to question the user, on a users state or settings and answer to questions queried by the application have to be stored as well.

User data is stored in the database and is protected with an administrator password. Without this password it is not possible to obtain data from the database other than through normal use of the application. For security reasons it is chosen to have database access functions embedded in classes that are above the web root. Authentication is discussed in the implementation chapter.

Most data is stored in a database. A database is an efficient way to store data, it is both fast and easily manageable. The design of the database is shown in the diagram in appendix A. Where necessary, the structure will be discussed in the implementation chapter. An overview of the tables and their purpose is shown in the table 3.1.

NAME	PURPOSE
page	Information for page requests. It stores which template to use for a request
auth	Authenticated users
user	Settings for a user that is signed up
photo	Details about uploaded photos
question	List of all questions in the system
question-type	Configuration for a certain type of question
answer-type	Result of answered questions of the same type

Table 3.1: Tables in the database

An exception to the data in the database are the fuzzy rules. These are stored in fuzzy inferencing classes named after their purpose. Having all information on a fuzzy inferencing system withing a class file named after its purpose makes it easier to manage and easier to see in a glance how the inference system behaves.

Chapter 4

Implementation

This chapter discusses the implementation of the fuzzy inferencing and the dating application as a whole. Not all parts will be discussed. The photo upload, authentication and the dating assistant will be discussed as part of the graphical user interface. But what distinguishes this dating site from others is not only the user interface but also the fuzzy matchmaking. As part of the fuzzy matchmaking the fuzzy inference system for the five factor model is discussed.

4.1 The Fuzzy Inferencing System

The Fuzzy Inferencing System is responsible for drawing inferences based on the knowledge within domains. Fuzzy Inference models can be implemented in different technologies. MATLAB is well known for its functionality for Fuzzy Logic but it is not easy to integrate into the website. An alternative was to use Perl. The CPAN site¹ has several libraries available for fuzzy sets and fuzzy inferencing within their AI branch. Perl is relatively fast and runs on the web server. Since the main logic is implemented in PHP and in order to obtain full understanding of what happens during the processing I decided to write my own PHP5 classes for the parsing and evaluation of fuzzy systems. This is by far the most flexible and insightful approach.

For the fuzzy inferences two main classes are developed. The functions are called from other classes, like the Match class that is responsible for the matchmaking or through an intermediate script responsible for an AJAX callback. The two classes that are developed for fuzzy systems are called *Fuzzy.class.php* and *FIS.class.php*. The naming might be a bit confusing since they can both be considered to deal with a fuzzy inference system (FIS). The FIS class implements all functionality for the implementation of Mamdani fuzzy inference systems containing of parts for the parsing and evaluation of fuzzy rules, the definition of membership functions and the process of defuzzification. The fuzzy class implements one or more fuzzy inference systems with the use of the Fuzzy class. The OCEAN model is implemented in six fuzzy inference systems resulting in a percentage that is used in conjunction with other systems to finally result in an overall match percentage.

The FIS class has most general functionality for fuzzy inference systems. An UML class diagram of this class is shown in figure 4.1. The OCEAN model is defined in the fuzzy class and makes use of this FIS class. It is discussed how the OCEAN model is defined and how the PHP classes for Fuzzy Logic work. The implementation of the OCEAN model is partly shown in appendix B. The theory behind this model is discussed in (Ghasem-Aghaee and Ören, 2003).

¹Comprehensive Perl Archive Network, <http://www.cpan.org/>

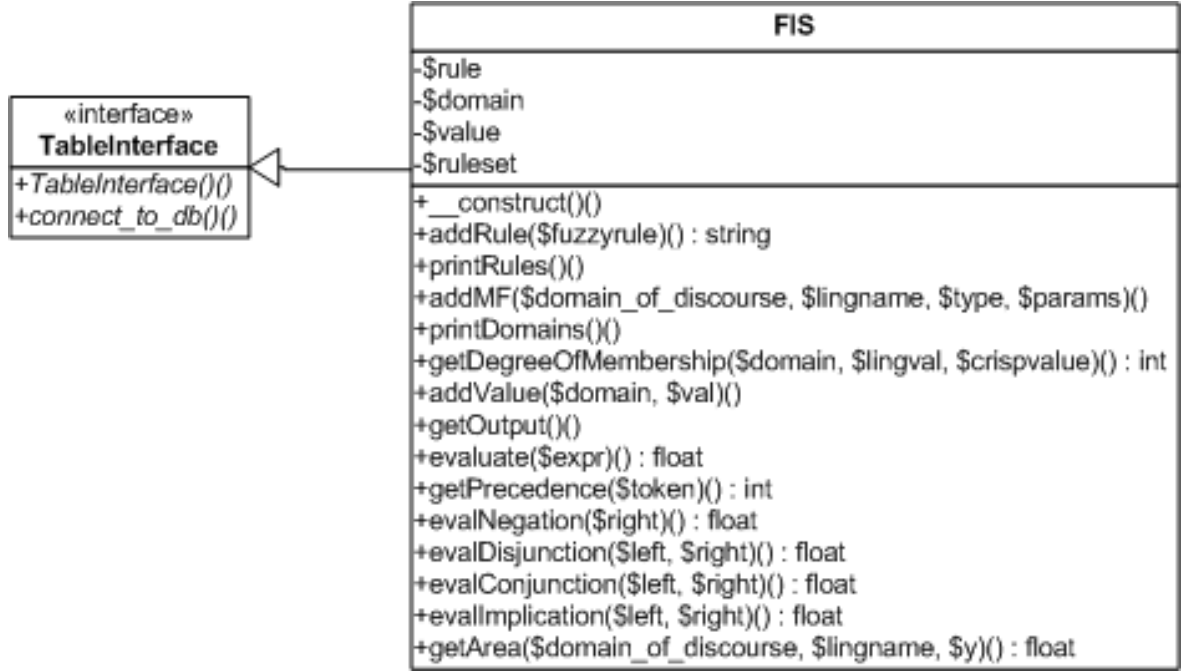


Figure 4.1: Class diagram for the FIS (fuzzy inference system) PHP class

When the calculation of negative emotionality (sometimes referred to as emotional stability) is considered it goes through some stages. Firstly the universes of discourse are defined. A shorthand function is use because many had the same character. It calls the member function addMF as follows:

```

$this->fis->addMF($name, 'low', 'trimf', array(-1.0, 0, 1.0));
$this->fis->addMF($name, 'high', 'trimf', array(0, 1.0, 2.0));

```

The addMF function creates a universe of discourse and adds fuzzy membership functions. In the example above a triangular membership function (trimf) is created that starts at -1.0 has a peak at 0 and is back at 0 at the end of the range. The triangular membership function for high is vertically mirrored for $x = 0.5$. This is modeled different than suggested by Ghasem-Aghaee and Ören (2003). The definition of the membership function is different because Ghasem-Aghaee and Ören (2003) define it at certain points in the range $[0,100]$ with a step size of 10 and in my class the membership function is defined using its characteristic parameters (start, peak, end), which was inspired by the way it is modeled in MATLAB.

Besides a triangular membership function a trapezoid membership function is also made available and the function leaves space to extent it with other membership functions like Gaussian ones, although those will be more computationally expensive. Another difference with the cited article is that Ghasem-Aghaee and Ören (2003) model three categories called low, medium and high. The medium membership is the smallest one and is centered within the domain as a triangular membership function. Their low and high memberships are at the sides of the domain and are modeled as a trapezoid with full membership on the edges for about 25/100 from their full domain of discourse. Modeling it this way you lose information from the crisp input since 2 (out of 100) is considered as much low as 24 (out of 100). Since there is no overlap as well you will not see a difference. Since the scale is considered to be highly linear there is also no aim of modeling a specific medium class for any other reason than that it might be more comfortable to describe the fuzzy rules in the way it was extracted by the authors in their article. To avoid this extra computational load it is therefore modeled as

described. It can also be considered that it is modeled by just one category but this is avoided on purpose because the modeling of the membership functions might be adjusted in the future to obtain better results. When it is not true that they can be mirrored in $y = 0.5$ then there is reason to use separate membership functions. Apart from that it might clarify the logic more in the fuzzy rules because it creates more freedom in the linguistic expressions. The output, negative emotionality, is modeled in a similar way as the inputs.

After defining the membership functions within the domains of discourse the knowledge is to be described. With the use of fuzzy rules it is tried to model the knowledge from the OCEAN model. Ghasem-Aghaee and Ören (2003) use a large rule that is the full conjunction of the six parameters that constitute negative emotionality. Depending on how the conjunction (or union) is modeled this might not be a good way to model it. For example if the union is evaluated with the use of the Gödel t-norm the minimum is taken from its constituents, however even if one parameter is very low it will most likely not mean that this should rule out all the other parameters. An alternative is to use the probabilistic product rule, that is used in this implementation. However in a large conjunction (many terms) a similar effect, though less drastically, will occur if one of the values is low (e.g. as a result of a mistake/outlier or just because the user is a bit different in one of the constituting factors). For this reason the disjunction (intersection) rule is added that compensates somewhat for this effect without reducing it to simple averaging. A slightly more computationally expensive approach might be adding (or replacing) simple implication rules for each parameter on its own. Future experiments need to be undertaken to find out which approach is more effective.

The rules in the knowledge base can be easily defined completely linguistically, no special characters are needed as instruction for the parser. The recursive parser is responsible for the evaluation of the rule and its constituents.

Finally the crisp valuations gathered from the users' answers are gathered from the database and the domains of discourse are assigned its current valuation. After which the defuzzification can take place by calling the `getOutput` function. The membership functions are scaled and the centre of gravity in combination with the mass measure (area) and results in a weighted crisp output. For other aspects of the OCEAN model a similar approach is undertaken. Using all outcomes a fuzzy match percentage is calculated based on similarity in personality factors. This is used as one of the constituents of the final match percentage that is used to present matches to the user.

4.2 The Graphical User Interface

As part of the data acquisition and interfacing with the system a novel GUI has been developed. A graphical user interface that is accessible, intuitive and enjoyable. In this section some aspects of this graphical user interface are discussed but to get the best impression a visit to the site² is highly recommended.

The main interface to the application (called the dashboard) is developed with the use of the EXT JS framework³. All modern web browsers have support for javascript (or ECMAScript) therefore javascript is good choice for the technology to use. Writing javascript is not particularly easy though and it is hard to debug since the result can not be seen in the source code. The support for ECMAScript which is highly similar to javascript is expressed by Lammarsch et al. (2008): "To provide interactivity, ECMAScript and the official W3C Document Object Model (DOM) need to be supported. This is the case for all major browsers. Several browsers have their own scripting dialects, but as they are all

²<http://www.aidating.co.uk/>, <http://www.aidating.nl/>, <http://www.datingatmosphere.com/>

³<http://www.extjs.com/>

based on the ECMAScript standard, one can avoid problems by sticking to it.”

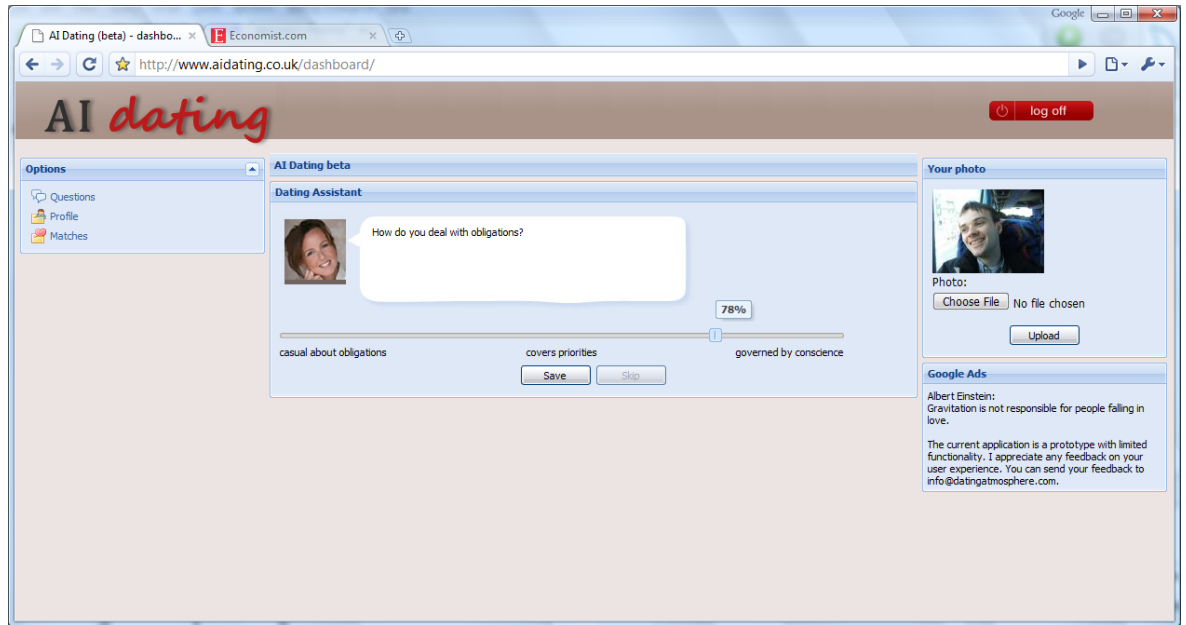


Figure 4.2: Graphical User Interface in Google Chrome browser

The EXT JS framework is one of several javascript frameworks that provides good support for AJAX (asynchronous javascript and xml) technology and it provides a wide variety of novel user interface components. Besides that it has an active forum and good developer support and the author has the impression that this framework will be one of the major frameworks for a longer time. The design and lay out of the dashboard is shown in figure 4.2.

4.2.1 The Dating Assistant

For the interaction between the site and the user it is chosen to use an analogy with the real life situation. An offline dating agency would have experts that ask you questions and give advice based on your answers. The questioning and answering principle as in a conversation is used in the concept of a dating assistant.

The dating assistant is shown in the centre of figure 4.2 where the dating assistant is asking a so called fuzzy question. The user is able to express their answer using a slider on a percentage scale which usually means the degree that a user agrees on a statement on its personality. As a guide there are linguistic terms defined for the lowest value, middle value and highest value on the scale. This type of user input is quite novel, it is adapted from the slider that was provided in a release of the framework during the development stage of the project itself. The control is not perfect and had to be adapted to make it suitable. A main critique on the slider control was that it is not modeled as a form field and the chosen value can therefore not be submitted as normal form control do. For this purpose a work around has been developed inspired by some suggested solutions by the EXT community.

The dating assistant is able to dynamically load questions from the database based on the priority and stage that is defined for a question. Only active questions will be considered. Questions of the same stage and priority can in theory be chosen in any order. Questions related to the OCEAN model are grouped within one stage to make it more likely that the inferences can be done, even when the user did not all the questions that are provided by the system. Questions can be of different types and the necessary configuration and control to answer them are loaded through an AJAX callback.

The main advantages of these callbacks is that the user interface itself (the page) does not have to be refreshed as a whole, which is both faster and it saves bandwidth. The controls itself are returned in a compact notation called JSON which stands for JavaScript Object Notation.

Although similar interfaces like the paper clip in Microsoft Office were not particularly popular it is believed that the interface with the use of an online character and a text balloon is intuitive and appreciated by the users. Besides that, it is supposed to give a more personal feeling and very suitable within this domain where it represent the dating agent as in offline dating agencies which is here modeled by means of the fuzzy expert system.

4.2.2 Authentication

Different scripts are available for user authentication. Some work with PHP, some with Perl, some work with server validation (like .htpasswd). Though the last one is quite secure and easy to implement it is not flexible, does not integrate easily in a site build up of modules and its look can not be customised. Many scripts are written by amateurs and presented on their websites in the form of a tutorial. Those scripts are often not proven to be secure. As an example a login script can suffer of a SQL injection attack if it is not properly protected.

It is chosen to make use of the PEAR::Auth library which is provided as a PEAR (The PHP Extension and Application Repository) module and is used at many websites. It is a well maintained library that lays a basis for authentication. Additional to the Auth module a script is written for the user to obtain a login handle. After signing up the user has to verify his email address as a means to make sure that he or she is not signing up someone else. In the verification e-mail a verification code with a randomly created floating point challenge code is sent for verification. After verification the entry is added to the 'auth' table and the user will be able to login to the dating application. The authentication module provides support for sessions. If a user is inactive for too long the session will expire so that a user that forgot to logout can't be misused after some time.

Since the AJAX methodology with callbacks to obtain data to update elements in the user interface is a potential risk since the callback can be directly called by a user bypassing the authentication or the dashboard interface all the callback scripts that return personal data are secured with a check for a valid authentication session. This way private data can not be requested by anyone without access to the dating site. A link to the login location or a custom log in form is provided when access is not granted.

4.2.3 Photo Upload

Physical attraction has a major influence on the likelihood of two people to match. Not only are attractive people more likely to be matched, people also seem to choose partners of similar attractiveness. Besides that, a photo gives a strong first impression. Maybe a too strong first impression, which is probably why the website PaiQ (2008) only shows photos after a conversation is going on for some time between two people. Photos can accompany a profile, the matches presented and can be used to rate attractiveness. In the current implementation a photo is used to accompany the matches presented and in order to give a more customised/personal feeling by showing it in your main (dashboard) screen.

The photo upload is given a space in the main application area. This is done so that people get confronted with it after every login and get stimulated to upload a photo of themselves and to keep this up to date.

The photo upload will only work for users that are logged in. Furthermore just after the upload of the file type and file size is checked. A thumbnail is generated in order to have all photos width of 150 pixels. In the database a reference to the photo is stored. It was also considered to store the photo in the database instead of on the file system but storage on the file system is faster and will not put a heavy load onto the database. All connections to the database are executed through the MDB2 database abstraction layer. This provides a lot of functionality and makes it more portable. In a later stage it also possible to smoothly migrate to another database.

Other users that have you in your match results will see your profile photo in their match overview.

Chapter 5

Testing

Given the time for the project extensive testing of the effectiveness of the dating application in suggesting matches for its users was beyond reach. However in the last weeks of the project about 500 people were invited through email to sign up for the dating site and some people were invited via websites. This resulted in just a view users with a meaningful profile and it is impossible to draw inferences based on this related to effect of the algorithm.

However testing is done during the implementation. When some functionality was developed it was tested by using it within some different situations to test if the algorithms produce the results that were expected. Unfortunately it happened often that small errors were made resulting in unexpected results or errors. Almost all of these errors are resolved by now and where they are not they have no importance for the functionality.

The fuzzy inferencing system has been tested by implementing the same models both in the developed model and by implementing it with a Perl library. The results were identical when max and min functions were used. This is done for two different models so it is likely that the fuzzy inferencing works well.

Besides that, test users pointed me at some errors in a few questions and a long initial load for users in India with internet connections that have lower transfer rates.

Given the high penetration and reliability of broadband connections in western countries and a main focus on this area the high initial load will not fundamentally change. There are some possibilities though to limit the load of the full javascript framework by making a custom build. This will be considered for the final release, not just to make it faster to load but also to save bandwidth. After the initial load the site performance should be acceptable since it only reloads necessary components with the use of AJAX technology.

User feedback is still encouraged and it is considered to implement a feedback form within the interface to get continuous feedback of users. This might lead to users expressing new ideas and current bottlenecks that can be addressed in a later stage.

Chapter 6

Conclusions, Evaluation and Further Work

6.1 Reflection and Conclusions

In a three months period the prototype for a dating site has been developed. Within this dating site a fuzzy inferencing system is applied as part of the matchmaking algorithm. Both the basic functionality of the dating application and the fuzzy inferencing system is successfully implemented. A literature review is undertaken that has resulted in knowledge I gained on fuzzy logic and notions on successful relationship prediction as described in literature from psychology. Furthermore a novel approach is taken for the development of the graphical user interface with a dating assistant that is dynamically updated with the use of AJAX technology. Experience is gained in the application of the EXT JS framework to build user interfaces. Even though some elements that are generally present at dating sites like profiles and contact facilities were not implemented a great amount of work is undertaken and a good basis is laid for matchmaking with artificial intelligence.

6.2 Further Work

There is a wealth of possibilities to be added to the dating site. Firstly, only a limited amount of important factors that are discussed in the background are yet implemented in the matchmaking. By extending the number of questions and adding extra fuzzy inference systems, classical inferences or mathematical inferences the matchmaking can be considerably improved. Especially geographic location, length and other characteristics can be modeled. It is also to be considered to determine personality not just based on the five factor model but to make a hybrid implementation that also uses the Myers-Briggs type indicator.

The dating site can also be extended with artificial neural network technology as discussed in the introduction. It will be interesting to see what results can be obtained with the use of artificial neural networks.

Modifications in the graphical user interface are needed and users must be able to give a small self-description. Also it thought of to add photo rating and profiles that show interests and the result of some fuzzy inferences (e.g. a personality description). A very important extension is that users must be enabled to contact each other.

Additional to this graphical user interface the Adobe AIR implementation that was suggested in

the analysis and design chapter can be implemented in the future.

A practical and necessary extension is a content management system to administer users data and to easily add, remove or edit questions in the system.

The developed prototype is great basis for novel dating sites.

Bibliography

- R. Bayne. *The Myers-Briggs Type Indicator: A Critical Review and Practical Guide*. Nelson Thornes, 1997.
- Lukas Brozovsky and Vaclav Petricek. Recommender system for online dating service. In *Proceedings of Znalosti 2007 Conference*, Ostrava, 2007. VSB.
- David M. Buss. Human mate selection. *American Scientist*, 1985.
- D.M. Buss. *The Evolution of Desire: Strategies of Human Mating*. BasicBooks, 1994.
- D.M. Buss, M. Abbott, A. Angleitner, A. Asherian, A. Biaggio, A. Blanco-Villasenor, M. Bruchon-Schweitzer, H.Y. Ch’U, J. Czapinski, B. Deraad, et al. International Preferences in Selecting Mates: A Study of 37 Cultures. *Journal of Cross-Cultural Psychology*, 21(1):5, 1990.
- Robert Callan. *The Essence of Neural Networks*. Prentice Hall, 1998. ISBN 013908732X.
- A. Dumitras and G. Moschytz. Understanding fuzzy logic: An interview with lotfi zadeh [dsp history]. *Signal Processing Magazine, IEEE*, 24(3):102–105, May 2007. ISSN 1053-5888. doi: 10.1109/MSP.2007.361606.
- Andries Engelbrecht. *Computational Intelligence: An Introduction*. Halsted Press, New York, NY, USA, 2002. ISBN 0470848707.
- M. Evans. *Love, an unromantic discussion*. Polity Press, 2003.
- Andrew Rocco Tresolini Fiore. Romantic regressions: an analysis of behavior in online dating systems. Master’s thesis, Massachusetts Institute of Technology, 2004.
- N. Ghasem-Aghaee and T.I. Ören. Towards fuzzy agents with dynamic personality for human behavior simulation. In *Proceedings of the 2003 Summer Computer Simulation Conference*, 2003.
- Jennifer L. Gibbs, Nicole B. Ellison, and Rebecca D. Heino. Self-presentation in online personals: The role of anticipated future interaction, self-disclosure, and perceived success in internet dating. *Communication Research*, 33(2):152–177, April 2006. doi: 10.1177/0093650205285368.
- Craig A. Hill. *Human sexuality: Personality and social psychological perspectives*. Sage Publications, Inc., 2008. ISBN 978-1-4129-0483-4.
- Günter Hitsch, Ali Hortaçsu, and Dan Ariely. What makes you click: An empirical analysis of online dating. Technical report, Society for Economic Dynamics, 2005.
- T.L. Huston. *Foundations of interpersonal attraction*. Academic Press, 1974.
- Peter Jackson. *Expert Systems*. Pearson Education Limited, 1998.

- J.J. Jassbi, P.J.A. Serra, R.A. Ribeiro, and A. Donati. A comparison of mandani and sugeno inference systems for a space fault detection application. *World Automation Congress, 2006. WAC '06*, pages 1–8, July 2006. doi: 10.1109/WAC.2006.376033.
- T. Lammarsch, W. Aigner, A. Bertone, S. Miksch, T. Turic, and J. Gartner. A comparison of programming platforms for interactive visualization in web browser based applications. *Information Visualisation, 2008. IV '08. 12th International Conference*, pages 194–199, July 2008. ISSN 1550-6037. doi: 10.1109/IV.2008.34.
- Choo Weisen Christopher Ledesma, Chua Wan Rong, Loh Mui Ling Felicia, Low Shee Ching Valerie, and Wu Wei Lin. The right one (tro) - computer matchmaking in the 21st century. 2005.
- Marcello Gallucci Marco Perugini. A hierarchical faceted model of the big five. *European Journal of Personality*, 11(4):279–301, 1997.
- M. Kyle Matsuba. Searching for self and relationships online. *CyberPsychology & Behavior*, 9(3): 275–284, 6 2006.
- Kibeom Lee Michael C. Ashton. A theoretical basis for the major dimensions of personality. *European Journal of Personality*, 15(5):327–353, 2001.
- David G Myers. *Social Psychology*. McGraw-Hill, 2002. ISBN 0-07-241687-5.
- OkCupid. Okcupid.com website, 2008. <http://www.okcupid.com/> [Online; accessed 10-June-2008].
- PaiQ. Paiq.nl website, 2008. <http://www.paiq.nl/> [Online; accessed 10-June-2008].
- Parship. Parship website, 2008. <http://www.parship.co.uk/> [Online; accessed 10-June-2008].
- T.I. Ören and N. Ghasem-Aghaee. Personality representation processable in fuzzy logic for human behavior simulation. In *Proceedings of the 2003 Summer Computer Simulation Conference*, 2003.
- Bradford W. Scharlott and William G. Christ. Overcoming relationship-initiation barriers: The impact of a computer-dating system on sex role, shyness, and appearance inhibitions. *Computers in Human Behavior*, 11(2):191–204, 1995.
- Veronica M. Scott, Karen E. Mottarella, and Maria J. Lavooy. Does virtual intimacy exist? a brief exploration into reported levels of intimacy in online relationships. *CyberPsychology & Behavior*, 9 (6):759–761, 2006. doi: 10.1089/cpb.2006.9.759.
- Shelley E. Taylor, Letitia Anne Peplau, and David O. Sears. *Social Psychology*. Prentice-Hall, Inc., 2000. ISBN 0-13-021336-5.
- Pascal Wallisch. Human mate choice as the psychologist views it: Findings, issues, problems and perspectives. n.p., 2002.
- Monica T. Whitty. Liberating or debilitating? an examination of romantic relationships, sexual relationships and friendships on the net. *Computers in Human Behavior*, 24(5), 2008.
- Wikipedia. Big five personality traits — wikipedia, the free encyclopedia, 2008a. <http://en.wikipedia.org/> [Online; accessed 24-August-2008].
- Wikipedia. Fuzzy set operations — wikipedia, the free encyclopedia, 2008b. URL http://en.wikipedia.org/w/index.php?title=Fuzzy_set_operations&oldid=193333237. [Online; accessed 9-September-2008].

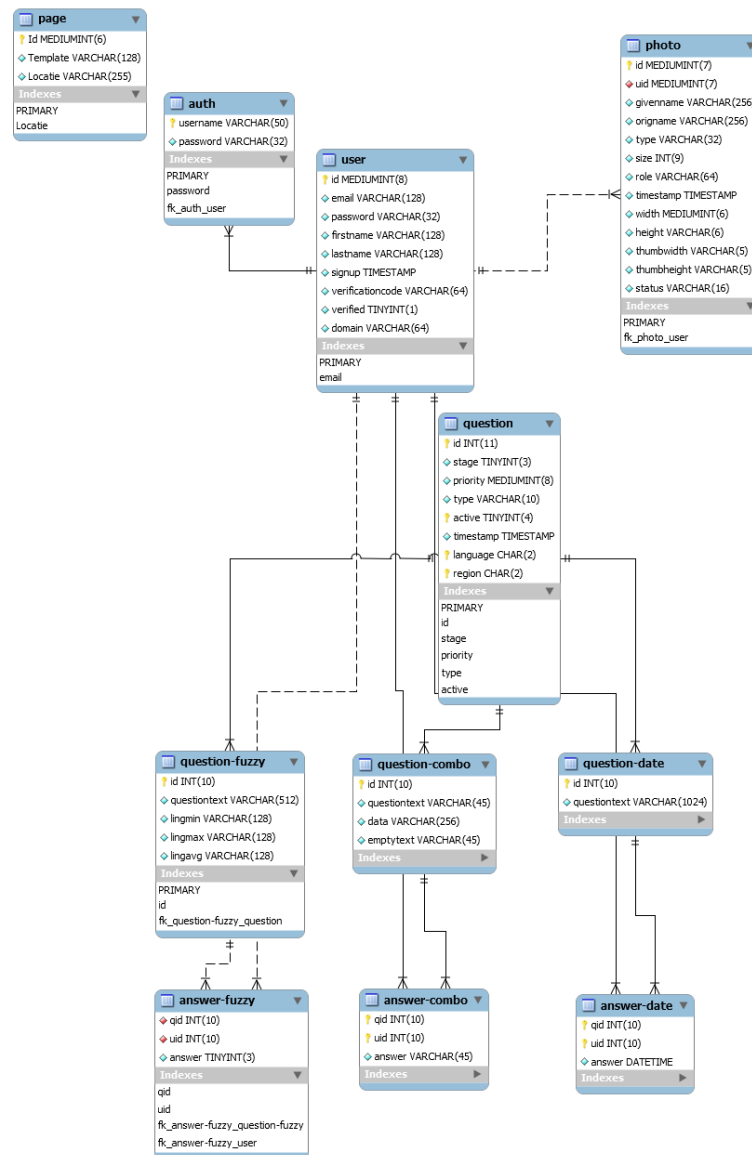
Lotfali Asker Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

Lotfali Asker Zadeh. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans. on Systems, Man, and Cybernetics*, SMC-3:28–44, 1973.

Jacek M Zurada. *Introduction to Artificial Neural Systems*. Pws Pub Co, 1992. ISBN 053495460X.

Appendix A

The Database Model



Appendix B

The Fuzzy OCEAN Model

The following code fragments are from the PHP class that implements the OCEAN model. Some functions are left out because they have high similarity with others. They will be available on the accompanying CD.

```
function getOpenness($do_fantasy, $do_aesthetics, $do_feeling,
    $do_actions, $do_ideas, $do_values) {
    [...]
}

function getConscientiousness($do_competence, $do_order, $do_dutifulness,
    $do_achievementstriving, $do_selfdiscipline, $do_deliberation) {
    [...]
}

function getExtraversion($do_warmth, $do_gregariousness, $do_assertiveness,
    $do_activity, $do_excitement_seeking, $do_positive_emotions) {
    [...]
}

function getAgreeableness($do_trust, $do_straightforwardness, $do_altruism,
    $do_compliance, $do_modesty, $do_tender_mindedness) {
    [...]
}

function getNegativeEmotionality($do_worry, $do_anger, $do_discouragement,
    $do_self_consciousness, $do_impulsiveness, $do_vulnerability) {
    $this->makeSimpleDomain('worry');
    $this->makeSimpleDomain('anger');
    $this->makeSimpleDomain('discouragement');
    $this->makeSimpleDomain('self_consciousness');
    $this->makeSimpleDomain('impulsiveness');
    $this->makeSimpleDomain('vulnerability');

    $this->fis->addMF('negative_emotionality', 'low', 'trimf', array(-100, 0, 100));
    $this->fis->addMF('negative_emotionality', 'high', 'trimf', array(0, 100, 200));

    $this->fis->addRule("If worry is high and anger is high and discouragement is high and self_consciousness is
    high and impulsiveness is high and vulnerability is high then negative_emotionality is high");
    $this->fis->addRule("If worry is low and anger is low and discouragement is low and self_consciousness is low
    and impulsiveness is low and vulnerability is low then negative_emotionality is low");

    $this->fis->addRule("If worry is high or anger is high or discouragement is high or self_consciousness is high
    or impulsiveness is high or vulnerability is high then negative_emotionality is high");
    $this->fis->addRule("If worry is low or anger is low or discouragement is low or self_consciousness is low or
    impulsiveness is low or vulnerability is low then negative_emotionality is low");

    $this->fis->addValue('worry', $do_worry);
    $this->fis->addValue('anger', $do_anger);
    $this->fis->addValue('discouragement', $do_discouragement);
    $this->fis->addValue('self_consciousness', $do_self_consciousness);
    $this->fis->addValue('impulsiveness', $do_impulsiveness);
    $this->fis->addValue('vulnerability', $do_vulnerability);

    return $this->fis->getOutput();
}

function getOceanMatch($ocean1, $ocean2) {
    $diff0 = abs($ocean1[0] - $ocean2[0]);
    $diffC = abs($ocean1[1] - $ocean2[1]);
    $diffE = abs($ocean1[2] - $ocean2[2]);
    $diffA = abs($ocean1[3] - $ocean2[3]);
    $diffN = abs($ocean1[4] - $ocean2[4]);
}
```

```

$this->makeSimplePercentDomain('dOpenness');
$this->makeSimplePercentDomain('dConscientiousness');
$this->makeSimplePercentDomain('dExtraversion');
$this->makeSimplePercentDomain('dAgreeableness');
$this->makeSimplePercentDomain('dNegativeEmotions');

$this->fis->addMF('ocean_match', 'low', 'trimf', array(-100, 0, 100));
$this->fis->addMF('ocean_match', 'high', 'trimf', array(30, 100, 170));

$this->fis->addRule("If dOpenness is low then ocean_match is high");
$this->fis->addRule("If dConscientiousness is low then ocean_match is high");
$this->fis->addRule("If dExtraversion is low then ocean_match is high");
$this->fis->addRule("If dAgreeableness is low then ocean_match is high");
$this->fis->addRule("If dNegativeEmotions is low then ocean_match is high");

$this->fis->addRule("If dOpenness is high then ocean_match is low");
$this->fis->addRule("If dConscientiousness is high then ocean_match is low");
$this->fis->addRule("If dExtraversion is high then ocean_match is low");
$this->fis->addRule("If dAgreeableness is high then ocean_match is low");
$this->fis->addRule("If dNegativeEmotions is high then ocean_match is low");

$this->fis->addValue('dOpenness', $diff0);
$this->fis->addValue('dConscientiousness', $diffC);
$this->fis->addValue('dExtraversion', $diffE);
$this->fis->addValue('dAgreeableness', $diffA);
$this->fis->addValue('dNegativeEmotions', $diffN);

return $this->fis->getOutput();
}

```

Appendix C

The Project Planning

The initial project planning.

