An expert system supporting diagnosis in clinical psychology

R. Spiegel[1] & Y. P. Nenh[2]
[1]*Department of Experimental Psychology and Wolfson College, University of Cambridge, UK*
[2]*Department of Computing, Goldsmiths College, University of London, UK*

Abstract

We introduce an expert system prototype that is meant to support psychologists in finding out what disorders their clients might have. The expert system provides a user interface that permits the psychologist to enter a large variety of symptoms. The symptoms are linked to a database where records of these symptoms/symptom combinations as well as their underlying disorders are stored. The system provides fuzzy rather than deterministic feedback, i.e. instead of suggesting only one diagnosis it indicates all possible diagnoses and estimates the risk for each possible diagnosis individually. The system is not meant to replace psychologists, but rather to support them in generating hypotheses at an early stage of diagnosis.
Keywords: expert system, database, clinical psychology, user interface, symptom classification.

1 Introduction

Early career psychologists often find it challenging to diagnose disorders. What makes this task particularly difficult is the fact that many disorders show a large overlap in terms of their symptoms, e.g. there can be significant weight loss in an eating disorder such as anorexia nervosa, but the same can apply to an affective disorder such as major depression. Current diagnostic manuals such as DSM IV (*Diagnostic and Statistical Manual of Mental Disorders*) [1] or ICD 10 (*International Statistical Classification of Diseases and Related Health Problems, 10th Revision*) [2] are books listing a large amount of symptoms

Human Perspectives in the Internet Society: Culture, Psychology and Gender, K. Morgan, J. Sanchez, C. A. Brebbia & A Voiskounsky (Editors) © 2004 WIT Press, www.witpress.com, ISBN 1-85312-726-4

assigned to disorders. Due to the aforementioned symptom overlap, however, this format makes it difficult to establish a diagnosis, as the symptoms of different disorders constantly need to be compared with each other until a diagnosis can be formed. Another problem with using book-format is the fact that disorders are described on separate pages, and going back and forth to compare symptoms and symptom combinations can take a considerable amount of time.

An alternative solution would be an automated system that would permit the psychologist to enter symptoms and to get feedback from a database listing possible disorders. Such a system would mimic an expert's behaviour, in this case a clinical psychologist searching for symptoms/symptom combinations and underlying disorders in a diagnostic manual. Rather than manually applying this procedure by referring to the aforementioned manuals in book-format, this system would search through the entire database of symptoms and disorders. This would have a couple of advantages:

First, the automatic search would be much faster than a manual search. Second, this system would be far more reliable than a human expert, provided the symptoms are entered correctly into the database. The human expert might make an error in the large number of listed symptoms/symptom combinations. Third, novel research findings about symptoms and underlying disorders can be updated more easily in a database than in a book. Finally, the graphical user interface in the automated version allows the psychologist to enter symptoms by simply clicking on symptoms that are present and by getting immediate feedback. This saves time over writing down symptoms during the surgery and later assigning these symptoms to an underlying disorder. Moreover, an electronic record of the symptoms can be stored in the system and is immediately accessible to other healthcare professionals through the database. This is advantageous over making manual records of symptoms, which would have to be entered into a database later. All these examples save valuable time that can be spent on the care for the client. Before going into technical details, we would like to give a general introduction on expert systems and support this introduction by practical examples from our system.

1.1 Introduction to expert systems

The purpose of an expert system is to simulate professional practices of a human expert. In our example, the human expert would be a *clinical psychologist* and the professional practices would be *considering symptoms and symptom combinations to diagnose an underlying mental disorder*. A more formal definition is provided by Jackson [3]: "An expert system is a computer program that represents and reasons with knowledge of some specialist subject with a view to solving problems or giving advice." Because our system is applied by a clinical psychologist, the justification for this system is to improve productivity by supporting the psychologist in the early stages of diagnosing a mental disorder. To be of practical use, expert systems should be quick and reliable in performing the particular task they specialise in. Expert systems are very specialised and restricted to a particular area, e.g. an expert system simulating a

Human Perspectives in the Internet Society: Culture, Psychology and Gender, K. Morgan, J. Sanchez, C. A. Brebbia & A Voiskounsky (Editors) © 2004 WIT Press, www.witpress.com, ISBN 1-85312-726-4

broker could not be applied to diagnose mental disorders. The specialised area is called *Knowledge Base* (Jackson [3]) and has been developed after careful study of the human expert, i.e. it usually contains knowledge provided by the human expert. In our case, this knowledge was provided by the careful study of diagnostic manuals. Moreover, the first author was both a computing professional and a psychologist with detailed education in clinical psychology. Other psychologists and healthcare professionals were also asked for advice. The knowledge gained from the human expert can be implemented into the system by a set of if-then rules (Alter [4]): "If certain conditions are true, then certain conclusions should be drawn." According to Alter [4], an ideal expert system should contain the following components:

- A *user interface* that lets the user enter facts and receive feedback. In our case this is a graphical user interface where the psychologist (user) can enter symptoms (facts) and get an estimated diagnosis of likely disorders (feedback).
- A *database* where facts (in our case: symptoms and diagnosis) are stored.
- A *knowledge base* that makes use of *if-then* rules. These rules were provided by an expert (psychologist).
- An *inference engine* that links the rules from the knowledge base with the facts from the database (in our case symptoms) to infer a result (in our case: a diagnosis) or to give alternative feedback (e.g. you have not specified any symptoms yet). The inference engine also makes use of *if-then* rules.
- An *explanation* offered to users who want to know how the inference engine draws conclusions (in our case: the reason why a diagnosis was made).

1.2 Introduction to diagnosis in clinical psychology

Clinical psychologists often refer to the two most widely accepted diagnostic manuals [1,2] when forming a diagnosis. There are a range of different disorders such as Generalised Anxiety Disorder, Obsessive-compulsive Disorder, Post-traumatic Stress Disorder, Borderline-type Personality Disorder, Anorexia Nervosa, Bulimia Nervosa, Acute Schizophrenia-like Psychotic Disorder, etc.

Possible causes of mental disorders can be organic, i.e. when a somatic illness such as a brain tumour is responsible for a mental disorder. Alternative causes could be reactive (e.g. someone suffering from depression due to the death of a loved family member) or endogenous (when the underlying cause is unknown). It is important to realise that the diagnosis can be the same irrespective of the underlying cause, e.g. someone can suffer from a major depression due to a change of neurotransmitter activity in the brain or due to an event that has no organic cause, such as the death of a loved one. Several authors have tried to define mental disorders. One example is that abnormal can be

Human Perspectives in the Internet Society: Culture, Psychology and Gender, K. Morgan, J. Sanchez, C. A. Brebbia & A Voiskounsky (Editors) © 2004 WIT Press, www.witpress.com, ISBN 1-85312-726-4

defined as "deviating from the norm or average" according to Gross et al. [5]. On the other hand, clinical psychologists usually avoid to classify clients into categories such as normal or abnormal. Rather, they assume that there is a continuous difference between people with and without a disorder, and differences may exist in terms of the degree specific symptoms are present. For a discussion see Davison and Neale [6] or Baumann and Perrez [7]. In order for something to be considered as deviating from the average it usually has to persist for a certain amount of time. If something is very transient, it might not be considered a mental disorder. One example would be a football supporter who feels depressed after his/her favourite team lost the cup finals. Although this person might show symptoms of depression right after the match, these symptoms might not persist in the same intensity for several weeks. Consequently, it would be wrong to assign a mental disorder to this person. An expert system simulating diagnosis in clinical psychology should therefore also consider the duration of these symptoms.

Another feature of diagnosis in clinical psychology is that the statistical manuals usually pre-classify disorders into specific categories, e.g. ICD 10 has groups such as *Schizophrenia, schizotypal and delusional disorders* or *Mood disorders*. DSM IV has categories such as *Eating disorders* or *Anxiety disorders*. This makes it easier for a clinical psychologist who might suspect that their client has an eating disorder. S/he can use one category when comparing different eating disorders. Having given an introduction to diagnosis in clinical psychology, we will now refer to the implementation of the system.

2 Implementation of the expert system

What follows is a detailed explanation of how the earlier mentioned five components (User Interface, Database, Knowledge Base, Inference Engine, Explanation) were implemented. At this stage, the expert system has the form of a prototype that certainly needs further testing and possible elaboration.

2.1 The graphical user interface

The interface was created in html code and PHP was used to connect the interface with the database. One reason for selecting html is that it can be viewed in any web-browser. Therefore, it should be accessible in health care institutions all around the world. This should also apply to institutions that have difficulties to afford the latest computing equipment. In order to access the expert system, the user (in our case a clinical psychologist) has to log onto the password-protected system. Once the user is logged on, s/he will receive an overview with a description about the purpose of the system and how it can be used. Based on the earlier mentioned diagnostic manuals [1,2], the symptoms are divided into categories, such as mood related symptoms, eating related symptoms or anxiety related symptoms. Each category is highlighted and once the psychologist clicks on a category, a list of symptoms appears on the screen. An example of three

Human Perspectives in the Internet Society: Culture, Psychology and Gender, K. Morgan, J. Sanchez, C. A. Brebbia & A Voiskounsky (Editors) © 2004 WIT Press, www.witpress.com, ISBN 1-85312-726-4

symptoms that appear in the category *mood-related symptoms* can be found in Figure 1.

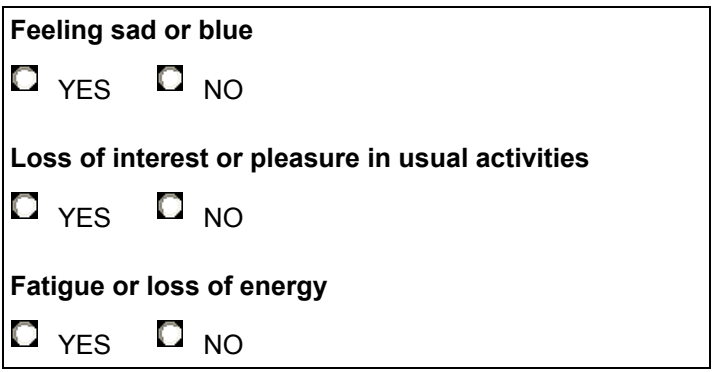

Figure 1: Example of symptom display on the interface.

In addition to the list of symptoms, the interface offers other features such as a Body Mass Calculator. The Body Mass Index BMI (kg/m^2) gives a numerical guideline whether a person's weight is within the normal range or not. The BMI adds to eating-related symptoms when diagnosing an eating disorder. It is calculated automatically when the psychologist enters the client's height and weight. JavaScript was used to implement the function that calculated the BMI.

Bipolar Disorder

High risk
1
2
3
4
5
Low risk

Figure 2: Graphical representation of risk for a specific disorder, e.g. bipolar.

Having entered all symptoms, the psychologist clicks on the submit button and the form with all indicated symptoms is submitted to the database. What happens to the submitted form in the database will be described later, but referring to the interface, the psychologists receives immediate feedback about a possible diagnosis. The psychologist is reminded that the system only performs an estimate and no final diagnosis. Moreover, the diagnosis is formulated in

Human Perspectives in the Internet Society: Culture, Psychology and Gender, K. Morgan, J. Sanchez, C. A. Brebbia & A Voiskounsky (Editors) © 2004 WIT Press, www.witpress.com, ISBN 1-85312-726-4

terms of risks but not as a final answer, e.g. *based on the symptoms, there is a high risk that the client suffers from a bipolar disorder (=manic depression). 8 symptoms matched the description of a bipolar disorder*. Moreover, a graphical representation on a scale from 1 (high risk) to 5 (low risk) can be obtained for the risk assessment of different disorders (Figure 2 displays an example). This enables a quick visual comparison between a range of possible diagnoses. A numerical representation is also available.

Furthermore, the interface allows the psychologist to enter any additional information in text format. This information will be stored in the database and can be used to exchange with other health professionals who are involved in the care for the client.

2.2 The database

Postgres was used as the database management system. Postgres is a relational database management system with the advantage that it can be extended to handle rules and objects similar to those in object oriented programming languages. Detailed information on Postgres can be found in [8]. After the psychologist clicks on the submit button, a connection to the database is established and the form with the client's symptoms is sent to the database. This database already contains information about all symptoms and all disorders, i.e. for each possible disorder, a list of symptoms is stored. Using the symptoms the psychologist indicated on the form, a query is created to find out how many of these symptoms overlap with the symptoms listed for each disorder. The number of overlapping symptoms indicates how well the symptoms on the submitted form match with each possible diagnosis. What has to be further considered is that occasionally some symptoms are better diagnostic features for a particular disorder than other symptoms, e.g. the symptom *low mood for at least 2 weeks* is a more important diagnostic feature for depression than the symptoms *fatigue or loss of energy*. Consequently, symptoms can have a different weighting for the same disorder (in this case depression). Alternatively, the same symptoms might differ in the way they match with a particular disorder, e.g. the symptom *excessive exercise* applies to both anorexia and bulimia nervosa, but it is a better indicator for anorexia nervosa, so it should be rated higher when the database searches through the symptom list describing anorexia nervosa. Instead of summing up the mere number of symptoms, it is therefore necessary to calculate the sum of weighted symptoms with respect to a disorder. The weighting is closely related to the knowledge base, which is based on experts estimating how the symptoms can be related to the disorder.

2.3 The knowledge base

The knowledge base is typically inspired by verbal reports from experts stating how they apply their knowledge to carry out a particular task. Subsequently, this knowledge is translated into a set of rules that can be programmed. It has to be kept in mind, however, that a real diagnosis in clinical psychology might be more complex than the set of rules that are eventually implemented into the

Human Perspectives in the Internet Society: Culture, Psychology and Gender, K. Morgan, J. Sanchez, C. A. Brebbia & A Voiskounsky (Editors) © 2004 WIT Press, www.witpress.com, ISBN 1-85312-726-4

expert system. According to Jackson [3] "few experts will provide a well-articulated sequence of steps that is guaranteed to terminate with success in all situations", so this expert system should only be seen as a tool to support the psychologist rather than as a tool that replaces expert knowledge. Real expert knowledge might consist of a large number of implicit skills and experiences that experts are not able to fully verbalise.

One example of translating expert knowledge into a knowledge base are the symptoms of depression (Gross et al. [5]), which is defined by persistent low mood for a duration of at least two weeks, plus five of the following symptoms:

- Significant weight loss or gain.
- Inability to sleep or difficulties staying asleep.
- Loss of energy or fatigue.
- Agitation, restlessness or irritability.
- Loss of pleasure or interest in usual activities.
- Pessimism or hopelessness.
- Recurrent thoughts of suicide, death or suicide attempts.

These symptoms of depression can be translated into a set of rules quite easily, but it gets more complex when symptoms have to be weighted. This is because there is hardly any quantitative measurement that would allow to weigh symptoms in a way that would be translatable into expert systems. Few psychologists know exactly how to weigh the symptom *exercise excessively* in anorexia nervosa as compared with bulimia nervosa. Even though clients with anorexia nervosa might be more likely to have this symptom, not all psychologists might agree in terms of the weighting that this symptom should be given. Although we relied on advice from psychologists, any future elaboration of this system will need to rely on further expert advice. After the symptoms have been translated into rules, these can be implemented. This happens in the inference engine.

2.4 The inference engine

The inference engine interacts with the previously mentioned components in the following way: when the symptoms are clicked and the form is submitted to the database, the database will proceed with this information because each symptom is stored as a fact in the database. Moreover, each fact is attached to rules in the inference engine. These rules were previously generated based on the information from the knowledge base. What follows are two examples of a rule in text format:

IF the weighted amount of anorexia nervosa-related symptoms is larger than or equal to 14, there is a high risk of anorexia nervosa. IF the weighted amount is smaller than 14 but larger than or equal to 10, there is a medium risk of anorexia nervosa. IF the weighted amount is smaller than 10 but larger than/equal to 5 there is a small risk. ELSE there is no risk of anorexia nervosa.

Human Perspectives in the Internet Society: Culture, Psychology and Gender, K. Morgan, J. Sanchez, C. A. Brebbia & A Voiskounsky (Editors) © 2004 WIT Press, www.witpress.com, ISBN 1-85312-726-4

An alternative rule would be:

IF there are 4 or more symptoms related to bulimia AND including the symptom "binge and purge" AND the body mass index is less than 18 THEN there is a high risk of bulimia nervosa.

PHP was applied to implement the rules and SQL SELECT statements were applied to calculate the sum of the weighted symptoms.

2.5 The explanation

Once the final results have been calculated by the inference engine, stored in the database, and sent back to the user interface, it would be good to have more than the mere output of the disorder. A confidence rating indicating how much the user can rely on this diagnosis would certainly be useful. This is partly provided in the earlier mentioned risk estimate (Figure 2), but could also give more detailed feedback, e.g. the exact number of symptoms which suggested a particular disorder. For example: there were 5 symptoms related to depression and 2 symptoms related to a Bipolar disorder (=manic depression).

3 System evaluation

Having implemented an expert system it is certainly necessary to test whether it works. There are different techniques to perform an evaluation. One is to test reliability, i.e. whether the system implementation is flawless. For example: are all the calculations performed in the correct way? Though only a prototype at this stage, the system was thoroughly tested and no error could be found. More specialised types of reliability are the concepts of *retest reliability* and *inter-rater reliability*. The former tests whether the diagnosis is the same if the same psychologist enters the same symptoms at 2 independent points of measurement. The latter tests whether two psychologists end up with the same diagnosis when entering symptoms for the same client. Whilst *retest reliability* was tested and is fulfilled in 100 percent of the cases, our prototypical system is still evaluated in terms of *inter-rater reliability*. It has to be said, though, that retest reliability is much easier to test in our case. If the system is correctly implemented then it will always give the same diagnosis provided the same values are entered at different times. This is because the underlying system will perform the same operations and therefore obviously output the same results. Inter-rater reliability, on the other hand, requires a detailed case study with different clinical psychologists who interact with the same client. This in turn requires a large research budget to carry out the evaluation. Another important technique is to determine *validity*, i.e. does the system really provide a valid diagnosis given the symptoms. This is an important test, because in case the symptom weighting or the rules in the inference engine are incorrect, the system will not provide a valid diagnosis. We performed two tests to cross-validate our prototypical system. In the first test, an

Human Perspectives in the Internet Society: Culture, Psychology and Gender, K. Morgan, J. Sanchez, C. A. Brebbia & A Voiskounsky (Editors) © 2004 WIT Press, www.witpress.com, ISBN 1-85312-726-4

expert with training in clinical psychology was asked to indicate what symptoms would be expected for a person with a particular disorder (e.g. a bipolar disorder). This expert then had to answer the questions related to the symptoms (of which a subset is shown in Figure 1). The expert was told to answer the questions in a way that the diagnostic result would indicate a high risk for a bipolar disorder and low risks for any other disorder. This case of cross-validation worked surprisingly well and the expert succeeded in this task. The second test was to generate a random answer pattern for the questions. If this expert system works, then a random answer pattern should not indicate any high risk for any disorder. This test also succeeded, but it remains to be said that these tests are case studies and need to be more extensive once this system leaves the prototypical stage. Cross-validation will become a more challenging task once more symptoms/disorders will be entered into the database.

4 Conclusions

The results of our prototypical system suggest that it seems possible to implement a large scale project of this expert system. It remains to be said, though, that this requires advice from more experts and probably a large administrative team that is responsible for entering the symptoms into the database. Entering the thousands of disorders/symptoms as well as more extensive validity checks will necessary. As mentioned before, this system would be able to support clinical psychologists at an early stage of diagnosis. As encouraging as our initial results might be, the present system should by no means be over-interpreted in a way that it might be able to replace clinical psychologists. If one considers this system as a support tool, however, then it might be beneficial. Though electronic versions of diagnostic manuals exist (such as the electronic version of DSM IV, Pies [9]), these are computer aided versions that help the psychologist search electronically rather than manually, but do not contain expert system features. As a result, we believe that an automated expert system similar to our prototype could be a useful addition to current diagnostic manuals. Our prototype was predominantly based on a classical approach, where rules are used to represent expert knowledge. The idea of having an expert system supporting diagnosis could probably be further extended with ideas from other computational approaches. One group of candidates might be models implementing analogical reasoning, as psychologists might often see analogies between clients who suffer from the same disorder. The analogies between clients in terms of behaviour, expression and other symptoms might further enhance the applicability of our system. Error-based learning would be possible if the database stores the combinations of symptoms and the suggested diagnosis, the psychologist verifies or alters the diagnosis and feeds it back into the database. The more people were diagnosed with this system, the more would the database be able to learn about symptom combinations and the resulting diagnoses. Consequently, the system would be able to make analogies between the symptoms/disorders of previous and new clients. These analogy-making processes might also help when attempting to

Human Perspectives in the Internet Society: Culture, Psychology and Gender, K. Morgan, J. Sanchez, C. A. Brebbia & A Voiskounsky (Editors) © 2004 WIT Press, www.witpress.com, ISBN 1-85312-726-4

diagnose disorders with the automated system. Various analogy-making models that could be adapted for this purpose are described by Hofstadter [10], Mitchell [11] as well as Spiegel and McLaren [12].

References

[1] *Diagnostic and Statistical Manual of Mental Disorders - Fourth Edition (DSM-IV)*, American Psychiatric Association, Washington D.C., 1994.

[2] *International Statistical Classification of Diseases and Related Health Problems,* 10th Revision, Chapter 5: Mental Disorders, Geneva, World Health Organization, 1992.

[3] Jackson, P. *Introduction to expert systems*, Addison-Wesley: Harlow, UK, 1999.

[4] Alter, S. *Information Systems*, Prentice Hall: Upper Saddle River, NJ, 2002.

[5] Gross, R., Mcilveen, R., Coolican, H., Russel, J. & Clamp, A. *Psychology: A new introduction*. Hodder and Stoughton General: London, 2000.

[6] Davison, G.C. & Neale, J.M. *Abnormal Psychology*. Wiley: New York, 1998.

[7] Baumann, U. & Perrez, M. *Lehrbuch Klinische Psychologie: Grundlagen, Diagnostik, Aetiologie*. Huber: Bern, 1990.

[8] Postgres: http://db.cs.berkeley.edu//postgres-v4r2/postgres.faq.

[9] Electronic version of DSM IV according to R. Pies: http://www.mhsource.com/expert/exp1052101a.html.

[10] Hofstadter, D. *Fluid Concepts and Creative Analogies*. Basic Books: New York, 1995.

[11] Mitchell, M. Analogy-Making as Perception: A Computer Model. MIT-Press: Cambridge MA, 1993.

[12] Spiegel, R. & McLaren, I.P.L. Abstract and associatively-based representations in human sequence learning. *Philosophical Transactions of the Royal Society, (Biological Sciences)*, **358**, pp. 1277-1283, 2003.

Human Perspectives in the Internet Society: Culture, Psychology and Gender, K. Morgan, J. Sanchez, C. A. Brebbia & A Voiskounsky (Editors) © 2004 WIT Press, www.witpress.com, ISBN 1-85312-726-4