Proceedings European Case Based Reasoning Conference, ECCBR 04, Madrid.

Case Representation and Retrieval in the Diagnosis and Treatment of Obstructive Sleep Apnea: A Semio-fuzzy Approach

Mila Kwiatkowska[1] and Stella Atkins[2]

[1]Department of Computing Science, University College of the Cariboo
Kamloops, BC V2C 5N3, Canada
mkwiatkowska@cariboo.bc.ca
[2]Department of Computing Science, Simon Fraser University
Vancouver, BC V5A 1S6, Canada
stella@cs.sfu.ca

Abstract. Obstructive sleep apnea (OSA) is a common sleep disorder, especially in middle-aged and obese patients. However, the diagnosis and treatment of OSA is complex, particularly in patients with mild or moderate symptoms and co-occurring medical problems. Furthermore, sleep disorders are interconnected with diverse factors, such life style, shift work, and psychological problems. Case-based reasoning (CBR) is a natural approach in such a highly complex domain. CBR allows for the retrieval and analysis of successful outcomes, as well as, failures in the treatment. To address the complexity of the domain knowledge, the authors applied a semio-fuzzy framework, an approach combining fuzzy logic and semiotics, in the development of a prototype CBR system. The system has been implemented in an educational sleep disorders clinic to assist respiratory students in the retrieval of prototypical cases, and has been particularly helpful in analysis of co-occurring medical problems and the treatment process.

1 Introduction

The case-based reasoning (CBR) paradigm has been successfully used in various medical fields, from lung disease, through cardiology, eating disorders, to diabetes and Alzheimer's disease [4,10,12]; nevertheless, a search of the sleep medicine literature has revealed only one example of CBR, used for sleep staging [11].Yet, in practice, the diagnosis and treatment of sleep disorders rely heavily on the analysis of typical as well as exceptional cases. For many patients, the diagnosis of the common disorder, such as insomnia, is relatively unproblematic, however, the non-pharmaceutical treatment requires long-term psychological and life-style changes. Furthermore, patients with mild or moderate symptoms, multiple sleep disorders, and co-occurring medical problems require individualized diagnosis and specialized treatment based on analysis of physiological, psychological, behavioral, and social factors. In addition, sleep disorders often develop gradually – many patients initially have mild symptoms or symptoms worsening periodically. In general, patients in sleep disorder clinics vary from uncomplicated cases to diagnostic “gray” zones and specialized cases with a long history of multiple diagnoses. Therefore, rule-based reasoning may be successfully utilized for the diagnosis of a single sleep disorder; however, the diagnosis and long-term treatment of more complicated cases would benefit from case-based reasoning or an approach that combines both paradigms.

This paper concentrates on the use of CBR in the diagnosis and treatment of a specific sleep disorder called obstructive sleep apnea (OSA). We are particularly concerned about patients with mild and moderate symptoms and with the co-occurrence of other medical problems. In particular, we focus on female patients who require individualized approach to diagnosis and treatment. The high variability among female patients, caused among many other factors by the menopausal changes, requires a case-based approach.

To demonstrate the benefits of CBR in the diagnosis and treatment of OSA, the authors developed a small prototype system, called Somnus, providing case storage and retrieval for educational purposes in the Sleep Disorders Clinic at the University College of the Cariboo (UCC). The Somnus system has been evaluated by students from the Respiratory Therapy Program, who worked in the clinic during their practicum.

The subsections 1.1, 1.2, and 1.3 provide a brief introduction to OSA and its diagnosis. The rest of the paper is organized as follows: Section 2 discusses the semio-fuzzy framework used for the domain knowledge representation. Section 3 describes the materials used for the construction of the system and the methods used for the case structure and the database architecture. Section 4 gives examples of case retrieval. Section 5 discusses the implementation issues and the system evaluation. The last section, Section 6, presents the conclusions and the directions for future work.

1.1 Obstructive Sleep Apnea (OSA)

Obstructive Sleep Apnea (OSA) is a common and serious respiratory disorder afflicting approximately 4% of middle-aged men and 2% of middle-aged women. OSA is caused by the collapsing of the soft tissues in the throat as the result of the natural relaxation of the muscles during sleep. The soft tissue blocks the air passage and the sleeping person literally stops breathing for 10 seconds or even 60 seconds. "Apnea" means "without breath," and occurs only during sleep, a condition that may go unnoticed for years. OSA is associated with increased risk for hypertension, congestive heart failure, coronary artery disease, and psychological disorders. Furthermore, sleep deprivation, caused by OSA, leads to several social consequences, such as motor vehicle accidents, poor job performance, job related injures, and, in general, a decreased quality of life [8].

1.2 The Diagnosis and Treatment of OSA

The gold standard for OSA diagnosis is an overnight in-laboratory polysomnography (PSG) involving continuous recordings of EEG, ECG, EOG, EMG, airflow, breathing effort, and oxygen saturation. However, PSG is a labor-intensive and costly test performed in specialized sleep disorders clinics only [8]. In Canada, most provinces have few clinics located in major cities, which limits significantly the access for patients from rural areas. Specifically, the province of British Columbia has five clinics and the waiting period for PSG varies from several months to over a year.

On the other hand, there is a growing interest in diagnostic alternatives, such as ambulatory home studies, clinical decision rules, and telemedicine. The UCC Sleep Disorders Clinic uses a gradual approach to the diagnosis. First, the patient maintains a sleep diary and answers sleep questionnaires. During the first visit, the students perform clinical assessment and take patient's medical history. Next, the patient undergoes home overnight oximetry and, subsequently, respiratory monitoring. At this point, the results are reviewed and the severe uncomplicated OSA cases are identified for the trial treatment. The patients with milder symptoms and inconclusive results are referred for full PSG studies and consultations with other clinics.

1.3 Gender Differences in the Diagnosis and Treatment of OSA

Although OSA is less frequent among women, insomnia, anxiety and depressive disorders are more frequent in women than in men, in all age groups. The frequency of insomnia is even greater in peri- and post-menopausal women. Studies of the middle aged population (45-55 years) [1] indicate that 29.8% of females suffer from insomnia as compared to 15.8% of males. In the female patients, OSA frequently coexists with insomnia, which in turn is often a symptom of anxiety and depression [14]. Additionally, recent studies of over 6,000 participants indicate that women report their sleep related symptoms differently than men [2]. For example, women are more likely to report 'tiredness' than the 'sleepiness.'

Therefore, the diagnosis and treatment of female OSA patients require a comprehensive approach and a careful analysis of each case. The following three factors must be considered: (1) frequent co-occurrence of other problems, for example, insomnia and depression, (2) hormonal status: pregnancy, lactation, monthly hormonal changes and menopausal changes, (3) psychological and social issues influencing the patient's compliance with the treatment.

2 Domain Knowledge Representation

The OSA diagnostic process uses subjective and objective data of varied granularity, uncertainty and precision. To address the diversity and complexity of data, the authors used a framework combining the fuzzy logic approach for modelling of the case features and the semiotic approach for the modelling of their measures. This combined framework, called a *semio-fuzzy* approach, was initially introduced by the authors in a paper presented at the NAFIPS 2004 conference [9]. This section describes briefly the fuzzy logic and the semiotic representations using an example of one of the most important symptoms of OSA – daytime sleepiness.

Fuzzy Logic. Since its introduction by Lotfi Zadeh in 1965 [15], fuzzy logic has been used in several medical applications [3,6] and in case-based reasoning [7]. The authors used two basic fuzzy concepts: linguistic variables for the representation of features and fuzzy membership functions for the specification of normal and abnormal sets. These two concepts are illustrated by the symptom of excessive daytime sleepiness [5]. Sleepiness can be measured only indirectly using two types of measures: subjective and objective. The most often used measure is a subjective self-administered questionnaire, the Epworth Sleepiness Scale (ESS), which is composed of eight questions to measure the general level of sleepiness in terms of probability of falling asleep during daytime activities, such as sitting and reading, sitting and talking, watching TV, driving a car. Each item has a response score between 0 and 3: never-0, slight chance-1, moderate chance-2, and high chance-3. The maximum score is 24. Typically, a score of 11 and above is recognized as 'excessive' daytime sleepiness, and a score above 20 as 'severe' sleepiness.
The symptom of sleepiness is defined by the linguistic variable L:

$$L = \langle X, T(X), U, M \rangle \tag{1}$$

Where:

X is the name of the variable, X = Daytime Sleepiness,

$T(X)$ is the set of possible terms (values) for sleepiness,

T(Daytime Sleepiness) = {*normal*, *excessive*, *severe*},

U is the universe of discourse, in this example, the ESS scale $U = [0,24]$,

M is a set of membership functions defining the *normal*, *excessive*, and *severe* sleepiness: $M = \{\mu_{normal}, \mu_{excessive}, \mu_{severe}\}$.

The set of membership functions M is extended by the semiotic approach.

Semiotics. Semiotics, the study of signs [13], defines 'sign' as any entity carrying some information and being used in a communication process. Since both the signs and the communication are present in all sciences, the semiotic approach has been used in almost all disciplines. The authors use the classical Peirce's semiotic triangle: *object*, *representation*, *interpretant* to represent feature (object), measure (representation), and diagnostic use of the measure in context of multiple factors (interpretant). In our example, 'sleepiness' is considered as an *object*, the ESS questionnaire is used as its *representation*, and the interpretation of the results for two population groups (general, female) corresponds to the *interpretant*. Each population group has specific membership functions for the normal, excessive, and severe sleepiness. The membership functions for the female population are represented by Figure 1. The X-axis corresponds to ESS score and Y-axis corresponds to the membership degree. Since studies [2,14] indicate that females typically underreport their sleepiness on the ESS scale, the threshold value for excessive sleepiness was lowered to 9 (from 11 for the general population) and for the severe sleepiness lowered to 16 (from 20 for the general population).

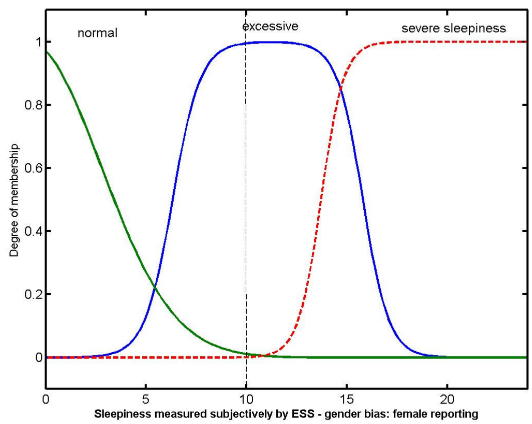

Fig. 1. Membership functions for the female population

3 Materials and Methods

3.1 Materials: the Clinical Records

The clinical data were obtained from the UCC Sleep Disorders Clinic, a teaching clinic providing training for students in the respiratory and polysomnography programs. The authors used 37 complete records of female patients diagnosed with OSA. The patients were referred by family doctors, underwent the PSG or overnight home study (e.g. oximetry, respiratory monitoring), and were diagnosed by a sleep specialist. The clinic maintains paper-based patient records. A typical complete record has several pages, and includes: clinical pre-assessment, sleep and bed-partner questionnaires, sleep log, results from PSG and home studies, diagnosis, and history of treatment. The data required for the prototype were manually extracted from the charts and entered into the database.

3.2 Methods: Case Structure

The cases are generalized into prototype cases and exceptional cases. A case C is defined by a set of feature-value pairs F, set of diagnoses D, set of treatments T, and an annotation A:

$$C = \langle F, D, T, A \rangle \quad (2)$$

Features. The case features are classified into eight groups: demographic, anatomical, physiological, life style, coexisting medical conditions, day time symptoms, nocturnal symptoms, and respiratory disturbance measures (see Table 1).

Table 1. Features associated with Obstructive Sleep Apnea cases

Demographic Features	***Daytime symptoms***
Age, Gender	*Excessive Daytime Sleepiness*
Anatomical Features	*Non-restorative Sleep*
Body Mass Index (BMI) kg/m^2	*Morning Fatigue, Headaches*
Neck Circumference	*Morning Dry Mouth or Sore Throat*
Crowding of Oropharynx	***Nocturnal symptoms***
Physiological Features	*Snoring*
Blood Pressure (BP), Heart Rate (HR)	*Snorting, Gasping, Chocking*
Life Style Factors	*Observed Apneas*
Sleep Hygiene	*Frequent Awakenings, Nocturia*
Smoking, Alcohol and Drug Abuse	***Respiratory disturbance Measures***
Coexisting Medical Conditions	*Respiratory Disturbance Index (RDI)*
Hypertension, Depression, Diabetes,	*Apnea-hypopnea Index (AHI)*
Asthma	*Desaturation Index (DI)*

Diagnoses. Sleep diagnoses are grouped into three categories: (1) sleep disorders (e.g. OSA, periodic limb movement disorder, insomnia, narcolepsy), (2) life style (e.g. inadequate sleep hygiene, smoking, alcohol and drug abuse) and (3) related medical problems (e.g. hypertension, depression, diabetes, obesity).

Treatments. The treatments are classified into five groups: (1) continuous positive airway pressure (CPAP) therapy, (2) corrective surgery, (3) use of oral appliances, (4) behavioral therapy, and (5) pharmacologic therapy. CPAP therapy is used for patients with moderate to severe OSA. The type, size, and pressure for the mask are established in a process called CPAP titrating. The behavioral therapy involves mainly the weight loss and changes in sleep hygiene.

Annotations. The completed cases are annotated by the medical professionals. The overall outcome and the patient's compliance level are recorded.

3.3 Methods: Somnus System Architecture for the Domain Base and Case Base

The Somnus system is composed of two databases: (1) domain base and (2) case base, illustrated respectively by Figure 2 and 3. The domain base represents the case features as linguistic variables with corresponding representations and interpretant's membership functions. It uses three relational tables called *Linguistic_Variable*, *Representation*, and *Interpretant_MF* (Figure 2). The case base is composed of three types of cases: (1) prototypical, (2) exceptional and (3) individual.

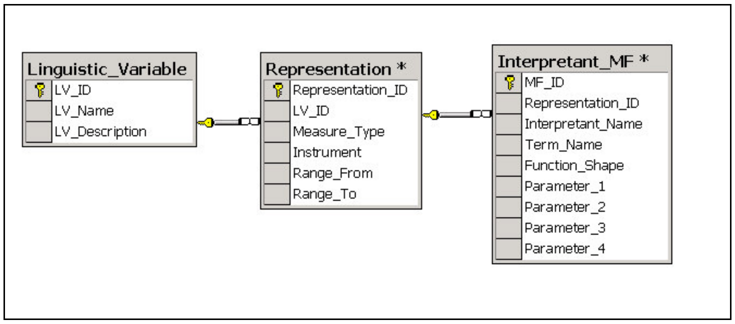

Fig. 2. Database schema for the domain base

The prototype cases (Prototype_Case table in Fig. 3) have 20 features corresponding to 20 dimensions (only 4 dimensions are shown). The exceptional cases have similar structure. The individual cases (Individual_Case table in Fig. 3) have 42 columns in total (only 14 columns are shown). The data for the individual cased were extracted from 37 records of female OSA patients. The prototype cases and exceptional cases were aggregated manually in consultation with the sleep specialist. Overall, Prototype_Case table has 20 records (prototype cases) and the Exception_Case table (not shown) has 2 exceptional cases.

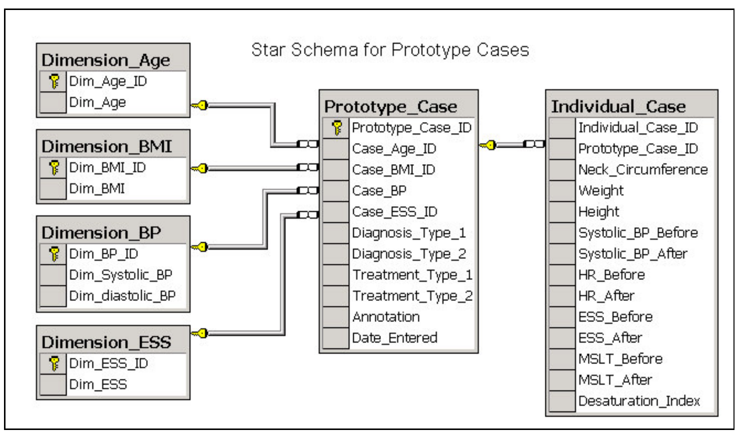

Fig. 3. Schema for the case base (selected dimensions and columns)

4 Case Retrieval

The retrieval process is based on three levels of abstractions: crisp values, fuzzy values, and general concepts. These three approaches are illustrated by Example 1. The retrieval interface uses SQL-like queries. The syntax of the queries is simplified to illustrate the overall idea. Example 2 illustrates a fuzzification function used for the interactive retrieval.

Example 1

Crisp Values. The user searches for the treatment history (text field) of individual cases, in which the sleepiness has decreased by 5 points on ESS scale or the systolic blood pressure has dropped by 40 mmHg. The following SQL query is used:

```
SELECT Tretment_history
   FROM  INDIVIDUAL_CASE
 WHERE (ESS_Before – ESS_after > 5
          OR Systolic_BP_Before – Systolic_BP_After > 40);
```

Fuzzy Values. The user searches for the prototype cases for obese patients (measured by Body Mass Index) with excessive daytime sleepiness (measured by ESS and 'biased' for female subpopulation (interprentant)). The following SQL query is used:

```
SELECT ALL
   FROM PROTOTYPE_CASE
 WHERE fuzzy_search ('BMI', 'obese')
          AND fuzzy_search ('ESS', 'excessive', 'Interpretant', 'female');
```

General Concepts. The user wants to analyze the types of instruments used for measuring sleepiness. The following SQL query is used:

```
SELECT Instrument
  FROM Representation R, Linguistic_Variable LV
WHERE LV.LV_Name = 'sleepiness' AND R.LV_ID = LV.LV_ID;
```

Example 2

Fuzzification Function. The following example illustrates the fuzzification of features. A new undiagnosed case has the following crisp features *F* (a set of feature-value pairs):

F = {(gender, female), (age, 45), (BMI, 39), (ESS,15)}

The user expands the search for similar diagnosed cases by using a fuzzification function, *Fuzzy*, which returns the following values:

Fuzzy (F) = {(gender, female), (age, middle), (BMI, obese), (ESS, excessive)}.

5 Implementation Issues and System Evaluation

Implementation issues can be grouped into three types: (1) representation of the case features and their measures, (2) database modelling, and (3) interface design.

The representation process involved selection of diagnostic features and their operationalization. The set of 20 features classified into 8 groups was established for this implementation based on consultations with the sleep specialists and research in sleep medicine literature. However, this is not a complete list of OSA features used in predictive models and, furthermore, this list should be standardized, using, for example, SNOMED terminology. The other aspect of the features, the operationalization, was very difficult to achieve for qualitative and subjective symptoms, such as excessive daytime sleepiness or morning fatigue. Therefore, in order to develop the first prototype, the authors decided to use a limited number of the basic features and their primary measures only.

The database modelling involved the relational model for the domain knowledge, and the multidimensional model (star schema) for the case representation. These two databases were created using MS SQL Server 2000. The data warehouse approach to case base was chosen for its rapid retrieval based on specific dimensions; however, this approach remains to be tested for a larger number of dimensions and cases.

The user interface in the first prototype of the Somnus system is limited to SQL-like statements providing a quick access to data; however, the use of the system is restricted right now to a small group of users familiar with the database schema.

The existing retrieval process has been evaluated by five respiratory students and two instructors. The evaluation involved functional testing and user feedback gained from structured interviews. The functional testing used three types of queries: (1) a search for cases based on fuzzy similarities, (2) a search for cases with a specific diagnosis and treatment, and (3) an analysis of the treatment outcomes. Each student evaluated two new cases (female patients). The students rated each query on a scale from 3 to 0: very useful (3), useful (2), somewhat useful (1), and not useful (0). Overall, the students created 25 different queries, most of them with one or two features only. On average, students rated the results as useful. The highest rating was assigned to queries searching for the coexisting medical conditions. The user feedback process involved both the students and the instructors. All participants agreed that the exploration of similar cases is beneficial to teaching and learning process, and the results from "vague" queries helped them to analyze large "clusters" of patients. On the other hand, most participants expressed a need for a "more user-friendly interface."

6 Conclusions and Feature Work

In this paper, we discussed the importance of case-based reasoning in the diagnosis and treatment of obstructive sleep apnea in a specific subpopulation - female patients with multiple sleep disorders and coexisting medical problems. Furthermore, we described a semio-fuzzy framework used to address the complexity of domain knowledge and to support fuzzy-based retrieval.

We have constructed the first prototype of the Somnus system as proof of concept that CBR is a useful framework for the diagnosis and treatment of sleep disorders. Our experience with the system provided us with three insights. First, the CBR approach was helpful in the analysis of diverse OSA cases among the female population. Second, the identification of features for OSA is complex since it involves diverse definitions and clinical perspectives; however, the use of semio-fuzzy approach provided us with a uniform representation for quantitative and qualitative features and subjective and objective measures. And third, the CBR approach reflects the nature of an educational setting, in which the students study diagnosis and treatment on a case by case basis. Therefore, even a limited prototype of a CBR system demonstrates that such a tool is beneficial for training future health providers.

Currently, the Somnus prototype is limited to case retrieval; however, its functionality and interface will be significantly improved. Future research is planned for expanding the work in four directions: (1) development of a conversational style interface, (2) addition of new features and cases to the prototype database, (3) addition of two functionalities - case adaptation and automation of the case storage, and (4) evaluation of the system, using records from the University of British Columbia Sleep Disorders Clinic.

References

1. Backhaus J., Mueller-Popker K., Hajak G., Vonderholzer U., GBerger M., Riemann D., Hohagen F.: Prevalence, Diagnosis and Treatment of Insomnia in General Practice. Journal of Sleep Research 7 (1998) 13.
2. Baldwin C.M., Kapur V.K., Holberg C.J., Nieto C.R.F.: Associations between Gender and Measures of Daytime Somnolence. Sleep 27 (2004) 305-311.
3. Bardossy A., Blinowska A.: Fuzzy Reasoning in Pacemaker Control. In: Szczepaniak P, Lisboa P, Kacprzyk J, (eds.). Fuzzy Systems in Medicine. Heilderberg: Springer-Verlag (2000) 451-466
4. Bellazzi R., Montani S., Portinale L., Riva A.: Integrating Rule-Based and Case-Based Decision Making in Diabetic Patient Management.ICCBR LNAI 1650 (1999) 386-400
5. Curcio G., Casagrande M., Bertini M.: Sleepiness: Evaluating and Quantifying Methods. International Journal of Psychophysiology 41 (2001) 251-263.
6. Czogala E., Leski J.: Entropy and Energy Measures of Fuzziness in Ecg Signal Processing. In: Szczepaniak P, Lisboa P, Kacprzyk J, (eds.). Fuzzy Systems in Medicine. Heilderberg: Springer-Verlag (2000) 227-245
7. Dubois D., Prade H., Esteva F., Garcia P., Godo L., Lopez de Mantaras R.: Fuzzy Set Modelling in Case-Based Reasoning. International Journal of Intelligent Systems 13 (1998) 345-373.
8. Flemons W.W., Littner M.R., Rowley J.A., Gay P., Anderson W.M., Hudgel D.W., McEvoy R.D., Loube D.I.: Home Diagnosis of Sleep Apnea: A Systematic Review of the Literature. Chest 124 (2003) 1543-1579.
9. Kwiatkowska M., Atkins M.S.: A Semio-Fuzzy Approach to Information Fusion in the Diagnosis of Obstructive Sleep Apnea. Proceedings NAFIPS (2004) 680-685.
10. Marling C., Whitehouse P.: Case-Based Reasoning in the Care of Alzheimer's Disease Patients.ICCBR LNAI 2080 (2001) 702-715
11. Park H.J., Oh J.S., Jeong D.U., Park K.S.: Automated Sleep Stage Scoring Using Hybrid Rule- and Case-Based Reasoning. Computers and Biomedical Research 33 (5) (2000) 330-349.
12. Schmidt R., Montani S., Bellazzi R., Portinale L., Gierl L.: Cased-Based Reasoning for Medical Knowledge-Based Systems. Medical Informatics 64 (2001) 355-367.
13. Sebeok T.A. Signs: An Introduction to Semiotics: University of Toronto Press (1999).
14. Vonderholzer U., Al-Shajlawi A., Weske G., Feige B., Riemann D.: Are There Gender Differences in Objective and Subjective Sleep Measures? A Study of Insomniacs and Healthy Controls. Depression and Anxiety 17 (2003) 162-172.
15. Zadeh L.: Fuzzy Sets. Information and Control 8 (3) (1965) 338-353.