Multi-layers Symbolic Fusion with conceptual graphs to support the Sleep Apnea Syndrome Diagnosis

Adrien UGON*, Jean-Gabriel GANASCIA*, Carole PHILIPPE†, Hélène AMIEL†, Pierre LÉVY†

*Laboratoire d’Informatique de Paris 6
4 place Jussieu
75252 Paris Cedex 05
France
{adrien.ugon, jean-gabriel-ganascia}@lip6.fr
†Hôpital Tenon (AP-HP)
4, Rue la Chine
75970 Paris Cedex 20
France
{pierre.levy, carole.philippe}@tnn.aphp.fr
amielhelene@gmail.com

Abstract—This paper describes a complete tool for supporting the diagnosis of Sleep Apnea Syndrome that uses multi-layers symbolic fusion. Parameters are extracted from recorded curves in the first step and then fused using domain knowledge, so that finally we obtain complex physiological events that indicate for each respiratory events that occured, its type, its start- and end-time, its onset context and its possible physiological consequences. During all the process, concepts and relationships are represented by using conceptual graphs.

From these complex events, contextual indexes are computed, giving to the physician a very precise description of pathology, something possible only by using fusion. Apnea-Hypopnea Indexes are compared with the physician’s results. The severity and specificity by severity class was compared to another similar system. All these results proved the efficiency of our approach.

Keywords—Symbolic Fusion, Sleep Apnea Syndrome, Support to diagnosis, Conceptual Graphs, Artificial Intelligence, OptiSAS.

I. INTRODUCTION

A. Symbolic Fusion

Symbolic Fusion has been proven to be efficient to fuse information from different sources, possibly heterogeneous. In [6], Belur DASARATHY advised to use different steps of symbolic fusion to increase the knowledge extracted from all sources. Each step allows to abstract the information got from the data recorded by the sensors framework to a semantic level where the description is rich enough to take a decision. Sensors generate large volumes of data that are often corrupted by noise and environment bias which can be solved by high-level fusion.

Through the different steps, knowledge from application domain is introduced. Symbolic Fusion offers the possibility to work directly on the concepts used in the application domain. The definitions of these concepts in the application domain allows to write the rules to extract them from the initial low-level data. The following steps will be used to implement the rules of fusion that connect the concepts to each other.

A symbolic approach needs to predefine a formalism under which the knowledge will be represented. In this work, we chose to use conceptual graphs to represent the information.

B. Conceptual Graphs

Conceptual graphs are a well known and widely used formalism for knowledge representation and reasoning that were first proposed by John F. SOWA in [23]. They are efficient to represent high level symbolic information. Moreover, the advantages of using conceptual graphs to proceed with high level fusion were demonstrated in [14]. The potential of conceptual graphs for reasoning is emphasized when it is applied to a domain with a clear ontology — with well defined concepts and hierarchy relations between them — and with clear fusion strategies [13].

In our case, conceptual graphs will only be used as a formalism for knowledge representation and not directly for reasoning.

C. Sleep Apnea Syndrome

The Sleep Apnea Syndrome is a sleep trouble characterized by repeated respiratory events that occur frequently during sleep. The gold standard medical examination to diagnose it is called polysomnography. It consists in recording many physiological parameters during a night. In particular, are recorded electroencephalograms (EEG) for brain activity, electrooculograms (EOG) for eyes movements, electromyograms (EMG) for muscle tone, respiratory flow, thoracoabdominal movements for respiratory effort, Blood Saturation in Oxygen ($SpO_2$), Pulse, electrocardiogram (EKG) for heart activity and Body position (cf Figure 1).
Sleep physicians, neurologists and pneumologists, have then to proceed to the scoring of data. During this tedious and time-consuming task, they look at curves split in short-time windows, in order to highlight physiological events that may have an impact on pathology. To support the diagnosis, we propose a system composed of several modules using symbolic fusion that increases, through different abstraction layers, the richness of the description model of the disease and its characteristics.

In [10], the American Academy of Sleep Medicine defined in 2007 the guidelines — Terminology and Rules — to use for the scoring of sleep. By nature, the complete interpretation needs several sequential steps to first, identify sleep stages, then to recognize respiratory and other physiological events and, finally, to link some of them to each other according to predefined rules.

The support to diagnosis of Sleep Apnea Syndrome is thus a good application domain to apply symbolic fusion since it benefits from consensual guidelines – the manual of terminology and rules – that should be used by every physician in sleep medicine.

To diagnose a Sleep Apnea Syndrome, a sleep physician refers, among other things, to indexes associated to physiological events that occur during sleep. These indexes are defined as the average hourly frequency of occurrence of an event:

$$index\ (event) = \frac{NB_{Event} (Occurrences)}{Total\ Sleep\ Time}$$

In this paper we calculate contextual indexes for a given event, that are the average hourly frequency of occurrences of this event when it arises in a given Sleep Context. We call the onset sleep context for a giving event the couple of states Sleep Stage/Body Position in which the sleeper was at the onset of the event. Contextual indexes were defined as infoxels in a previous work [27] that presented OptiSAS, a visualization method that inspired this method. Currently physicians do not benefit from indexes as precise as our contextual indexes. Most of the time, they have indexes for each type of event – apnea, hypopnea, desaturations and micro-arousals – and sometimes they also have the apnea-hypopnea indexes by sleep stage in a section and by body position in another section.

II. METHOD

To support the calculation of contextual indexes from recording curves, we used a system composed of different layers, each of them using symbolic fusion.

From signal to indexes, data are processed through different modules so as to obtain a rich description of the disease. To achieve this goal, domain knowledge is introduced from step to step, allowing to benefit from the contribution of each recorded curve using the knowledge that experts of the domain use in their own interpretation. We propose here to focus successively on each step of the whole processing.

To be more explicit, we will illustrate our method with an example from the real polysomnographic curves of Figure 2.

**A. From Data to Features**

The goal of the first step is to analyse numeric curves to extract physiological events. There are several events to extract and, for each, it is necessary to use a different method. To apply symbolic fusion in the next steps, the output should be in the form of conceptual graphs. The extraction of features can be done using conventional methods.

In that study, we considered as features sleep stages and events that should be scored, according to the AASM Manual for the Scoring of Sleep and Associated Events [10], visually by a sleep physician. Sleep stages are states of body and mind activities during sleep. Apneas and hypopneas are, respectively, complete cessations and diminutions of respiratory flow followed by a desaturation or an arousal. A Desuration is a decrease of Oxygen level; it as a physiological consequence of an apnea or a hypopnea. A Micro-arousal is a short awakening lasting between 3 and 10 seconds. We used here different algorithms for each type of event; arousals were scored visually by a physician.

---

1*Total Sleep Time* is defined as the whole duration of time during which the patient was asleep.
1) Sleep stages scoring: The scoring of sleep stages is usually done on 5 or 6 neurological curves — 2 or 3 EEG, 2 EOG and 1 EMG — by epochs of 30 seconds. The AASM defined 5 sleep-stages:

- **W**: Wakefulness
- **N1**: light sleep 1
- **N2**: light sleep 2
- **N3**: deep sleep
- **R**: Rapid Eyes Movements (REM) Sleep

To recognize them automatically, we used a binary decision tree. As input we defined 24 parameters, 8 for each type of curve. As output of them are obtained from a frequent analysis, filtering signals by alpha, beta, theta and delta waves with filter banks [4]. Learning and training were processed with Weka. We got the best results with REPTree associated with Bagging metaclassifier [26].

At the end of the first step, we have many conceptual graphs that describe the sleep stages by epochs of 30 seconds.

From the example of Figure 2 was built the conceptual graph of Figure 3. From top to bottom, the first curve gives the position; here the patient slept in supine position. Then, there are 2 EEG derivations curves : C3-A2 and C4-A1. Then, there are 2 EOG derivations curves : EOG-L and EOG-R. The next curve is the EMG curve. Then we can see the Respiratory Flow. Then there are the curves of thoracoabdominal movements. Finally, there is the \( O_2 \)-Saturation curve. On these curves, the sleep physician scored 3 informations : first we know the sleep-stage; the patient slept here in R (REM) Sleep-Stage. Then the patient scored an obstructive apnea on the Respiratory Flow curve. Finally, he scored a desaturation on the \( O_2 \)-Saturation curve and a micro-arousal on the neurological curves. We can thus conclude by looking at these curves that the patient, at 5:37:21, had an obstructive apnea, while he was sleeping in supine position and in R Sleep-stage; this apnea lead to a desaturation and to a micro-arousal. This is a full description of what happened physiologically at that time for this patient.

![Fig. 3. Example of Sleep Stage Conceptual Graph](image)

2) Detection of Respiratory Events: There are 2 types of respiratory events to detect [10]:

- **apneas** are defined as “a drop in the peak respiratory flow by \( \geq 90\% \) of baseline that lasts at least 10 seconds”.
- **hypopneas**\(^2\) are defined as “a drop in the peak respiratory flow by \( \geq 50\% \) that lasts at least 10 seconds, followed by a \( \geq 3\% \) desaturation or associated with arousal”.

We can see that, by definition, hypopneas result from the fusion of events from 3 types of events “diminutions of respiratory flow”, “desaturations” and “arousals”. In that case, the fusion strategy is a simplification by semantic equivalence.

For both respiratory event types, it is necessary to extract events from the *Respiratory Flow* curve. We defined 2 types of such events: “Cessations of Respiratory Flow” and “Diminutions of Respiratory Flow” that correspond to the criteria used in the definition of Apneas and Hypopneas. To detect these events, we first extract all respiratory cycles from respiratory flow curve. Then, we define a new curve that indicates the respiratory state at every moment of the night. During the respiratory cycles, the state is set to *Normal Air Flow*; otherwise it is set to *Stopped Air Flow*. To each respiratory cycle is associated a respiratory amplitude. The reference amplitude is then defined as the average amplitude of the 3 previous respiratory cycles. If criteria of Hypopneas or Apneas are met, we change the respiratory state during the respiratory cycle. We finally check if the duration criterion is met, that is if the event lasts 10 seconds; if yes, an event — Cessation of Respiratory Flow or Diminution of Respiratory Flow — is created with its start-time and its end-time.

From the example of Figure 2 was built the conceptual graph of an “Cessation of Air Flow” ; this is given on Figure 4.

![Fig. 4. Example of Cessation of Air Flow Conceptual Graph](image)

In [10], apneas are associated to a type depending on the presence or the absence of respiratory effort. Thus, for every Cessation of Air Flow, the 2 curves of thoracic and abdominal events are analysed to decide if there is presence, or absence, of respiratory effort in the two halves of the duration of the event:

- if there is presence of respiratory effort in the 2 halves of the event, the Cessation of Air Flow is interpreted, by equivalence, as an Obstructive Apnea;
- if there is absence of respiratory effort in the 2 halves of the event, the Cessation of Air Flow is interpreted, by equivalence, as a Central Apnea;
- if there is absence of respiratory effort in the first half and presence in the second half of the event, the Cessation of Air Flow is interpreted, by equivalence, as a Mixed Apnea;
- otherwise, the Cessation of Air Flow is interpreted as an Undefined Apnea.

This type of Apnea is not recognized by the AASM, but was used in our system when it was unable to take a decision. The physician may have to do it by himself, this would not be a time-consuming task.

From the example of Figure 2 was built the conceptual graph of an Obstructive Apnea; this is given on Figure 5. As we can see, there is a Fusion process. In that case too, the Fusion strategy is an interpretation by semantic equivalence.

\(^2\)There exists another definition of hypopneas in [10]
3) Detection of Desaturations: Desaturations are easy to detect as they are defined as a decrease of the $O_2$-Saturation of at least 3%.

From the example of Figure 2 was built the conceptual graph of a desaturation; this is given on Figure 6.

4) Micro-arousal: We did not have a method efficient enough to detect micro-arousals. Therefore we chose to ask the physician to visually score them.

From the example of Figure 2 was built the conceptual graph of the body position; this is given on Figure 7.

5) Body position: The body position does not require any interpretation since it is directly recorded. It is just necessary to represent every period of time with the same body position in a conceptual graph.

From the example of Figure 2 was built the conceptual graph of the body position; this is given on Figure 8.

B. From Features to Indexes

At the beginning of this second step, we have many conceptual graphs, extracted from polysomnographic curves, to our disposal. There exists some physiological link between these events that will be reconstructed using Fusion.

The main physiological events are the respiratory events. They occur in a given sleep context that may be an increasing factor for their onset by this patient. To know if it is the case, for each respiratory event, we will fuse the conceptual graph of the respiratory event and the one of the sleep context — body position and sleep stage — in which it arises. It is also important to know the physiological consequences of this respiratory event. The conceptual graphs of desaturations and arousals are thus fused to the conceptual graphs of respiratory events if the start-time of the arousal or the desaturation is between the start-time of the respiratory event and 20 seconds after its end-time. At the end of this step, each respiratory event will be associated to its sleep context and to its physiological consequences. We will call such type of event a complex event.

In that case, the fusion process is a completion operation.

In the example of Figure 2, in this step, conceptual graphs from Figures 3, 5, 6, 7 and 8 will be fused. The result is given on Figure 9.

Once every complex event obtained, in other words once every Respiratory event has been analysed and that its Onset context and its possible physiological consequences have been computed, we need to calculate indexes. To be as precise as possible, we generate contextual indexes. Given a Body Position and a Sleep-stage, we associate a Contextual Sleep Time. The contextual indexes are defined by the following formula:

$$ Contextual \ Index(\ Event) = \frac{NB_{Event, \ Sleep\ Context}(occurrences)}{Total\ Contextual\ Sleep\ Time} $$

where “Total Contextual Sleep Time” is the total time spent in a given sleep context (Body Position / Sleep Stage).

As “event” can be considered here either a single respiratory event or a complex event taking into account the respiratory event and all its physiological consequences.

III. RESULTS

We evaluated our method on a database of 70 patients’ polysomnographic data that was given to us by the Sleep Laboratory of Tenon Hospital (AP-HP) in Paris. These patients were recording using the Embla system and the scoring was done using the TM Somnologica software (Resmed). The patients are described in Table I. For all of them, we had the source curves and the Apnea Hypopnea Index (AHI), that is the average hourly frequency of occurrence of apneas or hypopneas during patient’s sleep.
TABLE I. DESCRIPTION OF TEST-DATABASE

<table>
<thead>
<tr>
<th></th>
<th>Mean (Standard-deviation)</th>
<th>Min-Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>55 (12.5)</td>
<td>26-81</td>
</tr>
<tr>
<td>F/M</td>
<td>23/47</td>
<td></td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.70 (0.11)</td>
<td>1.46-1.93</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>86.5 (19.6)</td>
<td>46-177</td>
</tr>
<tr>
<td>BMI (kg/m$^2$)</td>
<td>29.9 (6.0)</td>
<td>16.3-51.2</td>
</tr>
<tr>
<td>AHI</td>
<td>14.2 (13.7)</td>
<td>0.58</td>
</tr>
</tbody>
</table>

The figure 10 is a scatter plot on which we can compare the AHI got by the medical expert (on the x-axis) and the AHI got by our method (on the y-axis). To facilitate the analysis, the median axis and the severity classes appear on the figure. The color and the shape of the marker used for each patient indicate his severity class:

- $AHI < 5$ : normal patient;
- $5 \leq AHI < 15$ : mild SAS;
- $15 \leq AHI < 30$ : moderate SAS;
- $AHI > 30$ : severe SAS.

![Comparison of expert and automatic analysis AHI for the whole dataset](image1)

We observe on this figure that most of the points are close to median axis, which indicates that the automatic analysis gives results in the same range as the expert. More generally the automatic analysis found the same severity class for most of patients; sometimes, the AHI was overestimated, and the marker is over the median axis. That is due to the fact that we chose to favor the recall\(^3\) to the precision\(^4\). Indeed, it is easier for the expert to remove irrelevant events, than to look for those that would have been ignored by the automatic analysis.

![First Case’s not detected apneas](image2)

There are nevertheless 2 patients for whom the AHI was underestimated. These are surrounded by a red circle. We will analyse more deeply these 2 cases.

First case is represented with a red square. According to the medical expert, the AHI was 11; according to the automatic analysis, the AHI was 4. The expert scored 17 apneas, among which 8 were correctly detected and scored as apneas, 5 were also detected but scored as hypopneas and 4 were not detected (Figure 11). The expert also scored 108 hypopneas among which 96 were correctly detected and identically scored and 12 were not detected (Figure 12). By looking at missed events, we can suppose that these events were considered too short.

![First Case’s not detected hypopneas](image3)

\[^3\text{Recall} = \frac{TP}{TP + FN}\]
\[^4\text{Precision} = \frac{TP}{TP + FP}\]
Second case is represented with a purple circle. His AHI was 36 according to the expert and 14 according to the automatic analysis. This large difference is due to the fact that the signal is very noisy as we can see on Figure 13. On this figure we can see that the respiratory flow channel is very noisy and that respiratory events are more clear on the abdominal movements curve (in blue). Moreover, we can see that thoracic movements were not correctly recorded and that the curve is abnormally flat.

Fig. 13. Example of not detected apnea for 2nd case

To evaluate further the results, we decided to evaluate the following measure, percentage difference between the AHI got by the expert and by our method:

$$\left| \frac{AHI_{auto} - AHI_{expert}}{AHI_{expert}} \right|$$

We can observe, that the difference does not exceed 40% for patients with $AHI \geq 5$. For normal patients ($AHI < 5$), the difference is between 60% and 160%, which is due to the low values of AHI in this class. This does not lead to a misdiagnosis, what we can see on Figure 10.

Fig. 14. Boxplot, by severity, of the percentage difference of AHI

Despite these evaluations, it has not been possible to evaluate complex events and contextual indexes, since we did not know them and thus could not compare to our ones. This should be done in a further evaluation.

IV. DISCUSSION

Many researchers worked on the problem of supporting the diagnosis of Sleep Apnea Syndrome. They proposed different solutions; most of them analyze a single channel. Some work on the electrocardiogram or RR-intervall time series [1], [2], [12], [16], [17], others study overnight oximetry [15], [19], [20], [28] or mandibular movements [22].

There are also some work using several channels. In [9], Gabriela GUIMARAES et al. use multi layers fusion. First, parameters are introduced into a Self Organizing Maps from which are extracted some primitive patterns. Successions of primitive patterns are then interpreted as events; sequences of events are interpreted as temporal patterns. These temporal patterns correspond to real physiological events. In 2005, Tarik AL-ANI et al. in [3] on one hand and Paula Tamara POZZO-MENDOZA et al. in [21] proposed 2 similar solutions that analyze 2 signals: the Respiratory Flow and the Pulse Transit Time. Parameters are extracted from each signal separately. The decision is then taken by an Artificial Neural Network.

All these methods just aim at supporting the respiratory analysis and don’t work on neurological curves; a precise diagnosis needs thus the recognition of Sleep Stages to be done visually by a physician. However, there also exist complete methods that proceed both to respiratory analysis and recognition of Sleep Stages.

First of them, Isabel MILHO and Ana FRED presented in [18] a web supported development tool specific for medical diagnosis, based on Bayesian networks and tested on Sleep Apnea Syndrome. The second system is called PSG-Expert and was presented in [8]. It is an expert system in the domain of sleep disorders exploring polysomnographic data that implements a knowledge editor including a fuzzy fact editor and a rules editor. The third one is called SAMOA and is presented in [5] by Mariano J. CABRERO-CANOSA et al. It is a help tool for automatic SAS diagnostic system that incorporates both conventional programming and artificial intelligence techniques. Respiratory and neurological curves are analyzed separately with different modules. At the end, following results are generated:

- existence, or nonexistence, of a Sleep Apnea Syndrome (if yes, its type is given);
- global AHI;
- index by Apnea Type;
- number of events by their type;
- number of desaturations and resaturations;
- distribution of values of Oxygen Saturation;
- length of each sleep stage.

This system was extended by Ángel FERNÁNDEZ-LEAL et al. in [7] with a treatment of temporal information.
Finally, Ana Claudia Tonelli de Oliveira et al. tested in [24] a portable respiratory monitoring called Somnocheck, but its method is not revealed.

As we can see, there are lot of works on the support to diagnosis of Sleep Apnea Syndrome, but, although fusion is inherent in the domain — by the definition of events or by the physiological link that join them to each other to understand the pathology — results are often given into separate sections. In addition, patients used for evaluation are different from each other. For these reasons, it is difficult to compare presented methods to our one.

However, we decided to compare our method to Somnocheck, tested in [24] on 149 patients. We disposed from the sensitivity and specificity of each method on different datasets and grouped by AHI ranges. Results are presented in Table II. In the last column, we removed the patient with noisy signal (Figure 13).

<table>
<thead>
<tr>
<th>AHI range</th>
<th>Tonelli de Oliveira</th>
<th>UGN (70 patients)</th>
<th>UGN (69 patients)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHI ≥ 5</td>
<td>96.2</td>
<td>64.7</td>
<td>93.8</td>
</tr>
<tr>
<td>AHI ≥ 15</td>
<td>81.3</td>
<td>82.6</td>
<td>93.9</td>
</tr>
<tr>
<td>AHI ≥ 30</td>
<td>80.0</td>
<td>92.1</td>
<td>97.8</td>
</tr>
</tbody>
</table>

We observe that we got better results for patients with moderate and severe Sleep Apnea Syndrome, who are the patients for whom a misdiagnosis would have the most serious consequences.

The high combinatory complexity and richness of generated contextual indexes need an associated visualization for the physician to benefit from them. This one was given in the fifth chapter of [25].

Among the advantages of our approach, the use of several layers and modules offers the possibility to improve the global efficiency by changing only one subpart of the system. Most of other approaches fuse all the informations in one operation, for example with an artificial neural network, and everything should be done again if a new source of information were added or in case of any changing of the rules. With our approach, additional sources of information may contribute to improve the existing rules, by completing or updating them.

The problem of noisy data can be fixed by a preliminary analysis that could try to clean the signal or decide, if it fails in it, to alert the user that at least one recorded signal is too noisy for the tool to be efficient in processing the data of the given patient. It should be also discussed with experts how to deal with missing data.

To help fully the physician and propose a diagnosis of a Sleep Apnea Syndrome or a treatment the system needs some more information, as, for Example, the score of the patient on the Epworth Sleepiness Scale [11] or other information. The choice of treatment may depend on other factor : age, gender, BMI7 and cardiovascular comorbidities. We could improve the system by adding these informations so that the system proposes a treatment. By using a symbolic representation of the data, the system offers the flexibility necessary to work with heterogeneous sources; this is another advantage. This will be done in future work.

### V. Conclusion

In this paper we propose a new method for supporting the diagnosis of Sleep Apnea Syndrome. The system is composed of different layers. First, events are extracted from recorded signals. These events are mostly inspired from physiological events defined by the American Academy of Sleep Medicine in [10]. In consequence, their recognition uses medical knowledge. Represented formally as conceptual graphs, events are then fused so as to obtain, for each respiratory event, its onset context and its physiological consequences. The main advantage of our method is the automatic processing of all curves, very rarely offered by other systems. This was made possible by using symbolic fusion. We proved that we got similar results as the human expert and that it got better results than other existing systems.

### REFERENCES


---

5Sensitivity = TP / TP + FN
6Specificity = TN / TN + FP

7Body Mass Index


