Automated Classification of Sleep Stages using EEG/EOG and R&K Rules

Dennis Castleberry, Arnab Ganguly, Sahar Navaz, Mohammad Tohid, Manohar Karki

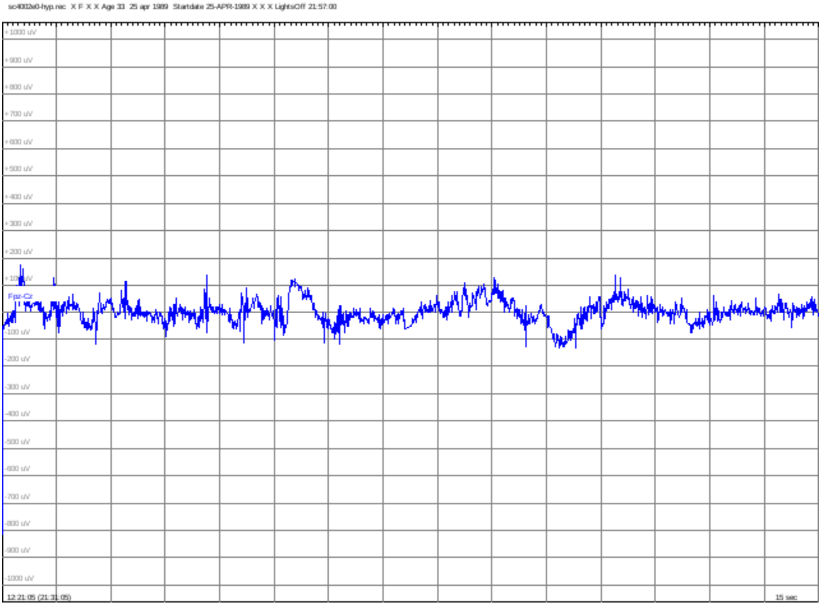
November 26, 2012

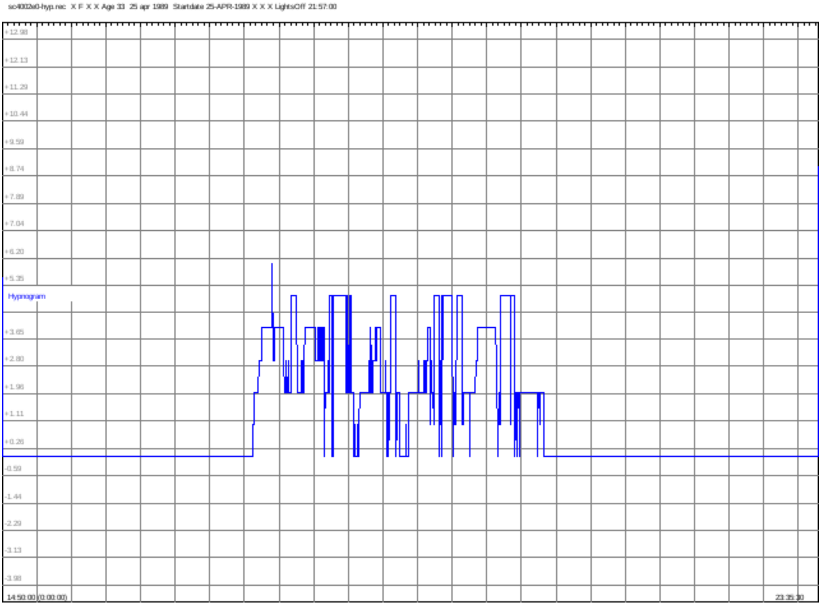
Problem Definition

- The problem is to use electroencaphalogram data to create an automated sleep scorer. Until recently, sleep psychologists have scored sleep stages by hand–a laborious and time-consuming task.
- Sleep stage scoring is a classification problem. We chose to create two classifiers: a rule-based classifier and a Naive Bayes classifier.

Background: Rechtschafffen & Kales

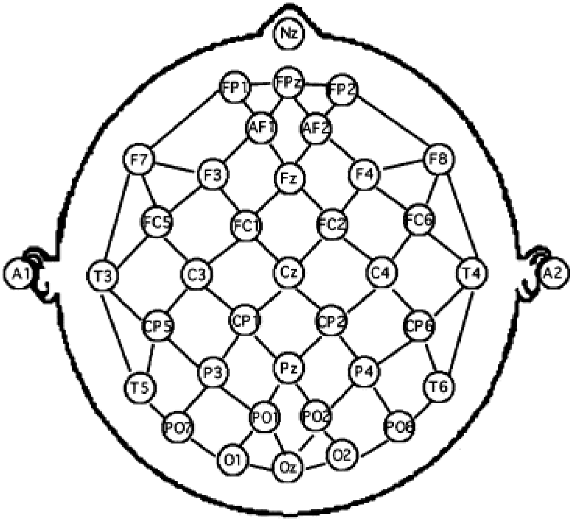
- A standardized set of rules for manually scoring EEG data was developed by Rechtschaffen & Kales. The EEG data is divided into 30-second intervals called **epochs** and independently scored using the rule set.
- The resulting graph of the stages is called a **hypnogram**.

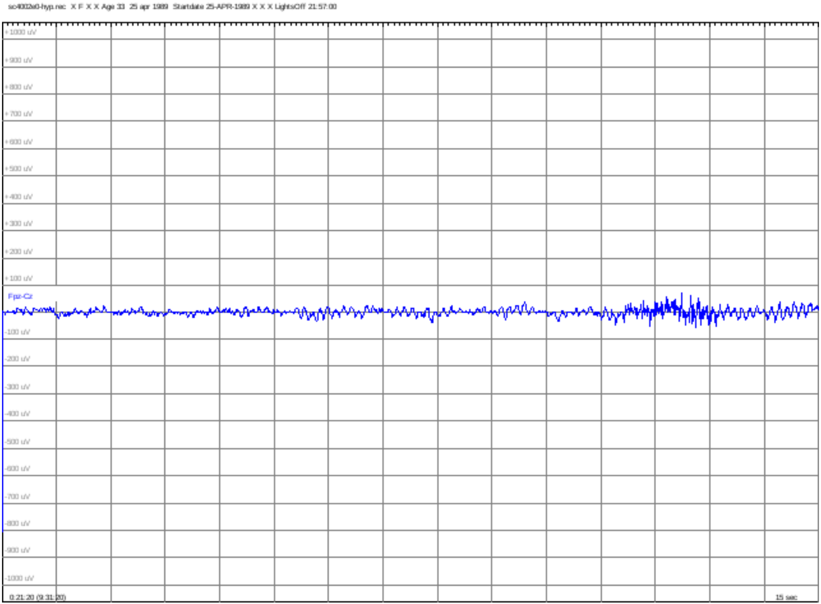

sc4002e0-hyp.rec X F X X Age 33 25 apr 1989 Startdate 25-APR-1989 X X X LightsOff 21:57:00
+1000 uV
+900 uV
+800 uV
+700 uV
+600 uV
+500 uV
+400 uV
+300 uV
+200 uV
Fpz-Cz
-100 uV
-200 uV
-300 uV
-400 uV
-500 uV
-600 uV
-700 uV
-800 uV
-900 uV
-1000 uV
12:21:05 (21:31:05)
15 sec

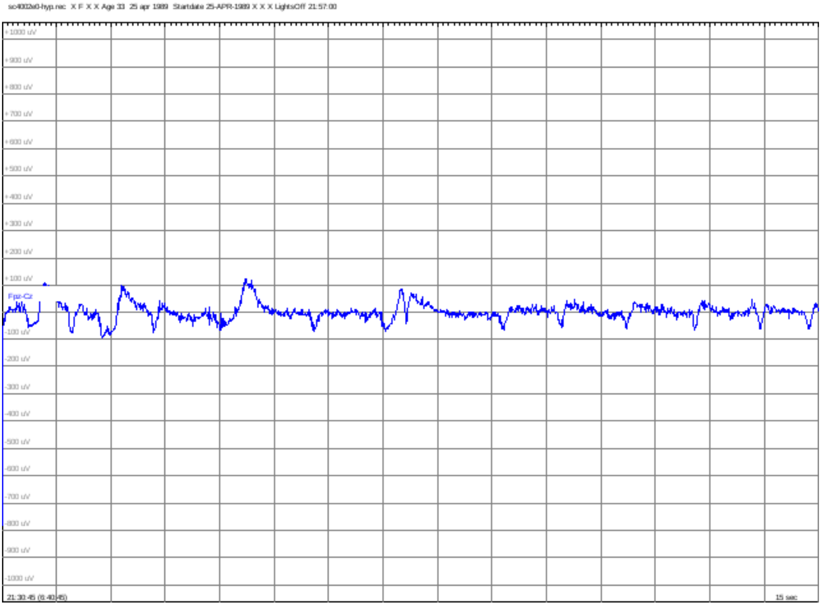

sc4002e0-hyp.rec X F X X Age 33 25 apr 1989 Startdate 25-APR-1989 X X X LightsOff 21:57:00
+12.98
+12.13
+11.29
+10.44
+9.59
+8.74
+7.89
+7.04
+6.20
+5.35
Hypnogram
+3.65
+2.80
+1.96
+1.11
+0.26
-0.59
-1.44
-2.29
-3.13
-3.98
14:50:00 (0:00:00)
23:35:30

Background: EEG

- Electroencephalogram (EEG) is a measure of electrical activity along the scalp.
- Certain waveforms manifest in the EEG output during different stages of consciousness. For example, a wave called a K-complex appears predominantly in Stage 2 sleep. Also, amplitude and frequency vary among the sleep stages.

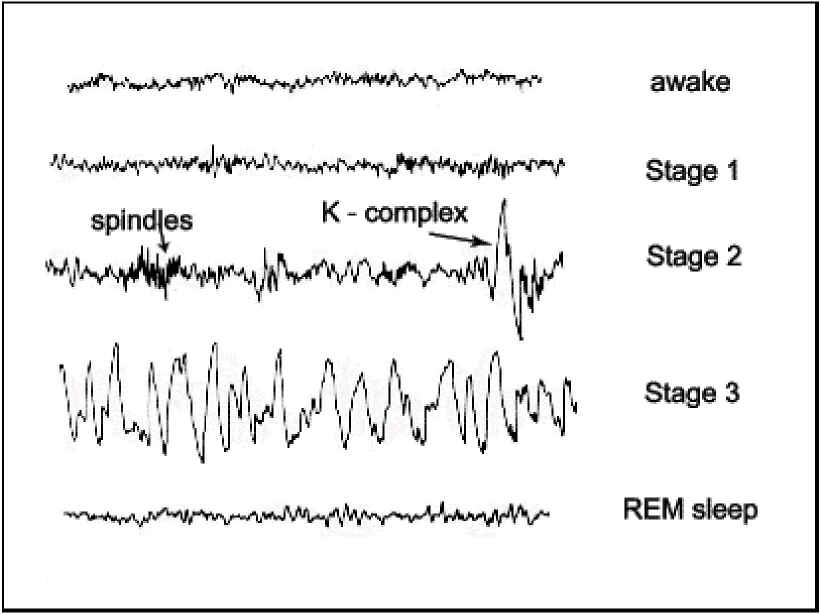

Nz
FP1
FPz
FP2
AF1
AF2
F7
F3
Fz
F4
F8
FC5
FC1
FC2
FC6
A1
T3
C3
Cz
C4
T4
A2
CP5
CP1
CP2
CP6
P3
Pz
P4
T5
T6
PO7
PO1
PO2
PO8
O1
Oz
O2


sc4002e0-hyp.rec X F X X Age 33 25 apr 1989 Startdate 25-APR-1989 X X X LightsOff 21:57:00
+1000 uV
+900 uV
+800 uV
+700 uV
+600 uV
+500 uV
+400 uV
+300 uV
+200 uV
+100 uV
Fpz-Cz
-100 uV
-200 uV
-300 uV
-400 uV
-500 uV
-600 uV
-700 uV
-800 uV
-900 uV
-1000 uV
0:21:20 (9:31:20)
15 sec


sc4002e0-hyp.rec X F X X Age 33 25 apr 1989 Startdate 25-APR-1989 X X X LightsOff 21:57:00
+1000 uV
+900 uV
+800 uV
+700 uV
+600 uV
+500 uV
+400 uV
+300 uV
+200 uV
+100 uV
Fpz-Cz
-100 uV
-200 uV
-300 uV
-400 uV
-500 uV
-600 uV
-700 uV
-800 uV
-900 uV
-1000 uV
21:30:45 (6:40:45)
15 sec

Background: The R&K Rules

- (W): there is a low voltage within the $10 - 30\mu V$ range and mixed frequency.
- (1): low-voltage mixed-frequency EEG with highest amplitude in the $2 - 7\mu V$ range.
- (2): presence of sleep spindles and K-complexes at least three minutes apart
- (3): 20-50% of the record should contain waves with $2Hz$ or lower and with amplitudes above $75\mu V$.
- (4): similar to (3), except amplitudes greater than $75\mu V$ appear in 50% of the record.
- (R): similar to (W), except that the EOG activity is high.
- (MT): if the EEG signals are unclear due to amplifier blocking or muscle activity.


awake
Stage 1
spindles
K - complex
Stage 2
Stage 3
REM sleep

Background: Waveforms

- There are two waveforms of concern for the Stage 2 classification:
  - **Sleep spindles** have a duration of at least $.5s$ with frequencies in the $12 - 14Hz$ range.
  - **K-complexes** are sharp negative spikes followed by slow positive peaks, at least $.5s$ in duration, usually with higher amplitude than the surrounding waves.
- We noted that, although the Rechtschaffen & Kales rules are clearly-defined, the sleep spindles and K-complexes are less so. Their interpretation is often subjective and based upon contextual information, which is problematic for manual scoring.

Objective

- Bearing this in mind, our objective was to compare a direct rule-based automated scorer with a Naive Bayes classifier to see the extent to which the Naive Bayes classifier would compensate for measurement errors.
- For this, we would need to compute the attribute values ourselves, as well as devise an automated scorer faithful to the R&K rules.

Hypothesis

- Suppose a rule-based classifier uses a set of attributes A.
- We hypothesized that a Naive Bayes classifier using the a subset of A would outperform our direct rule-based classifier.
- Our hypothesis was based upon the assumption that our waveform detection algorithms would yield a low accuracy.
- The corresponding stages would not be detected using a direct rule classifier where the waveform detectors fail; however, since the Naive Bayes classifier uses all the attributes A to profile the stage, it may compensate for the waveform detection errors.

Data

- We retrieved data from the Sleep-EDF database, available on-line at: http://www.physionet.org/physiobank/database/sleep-edf/
- Each EEG recording has the following data:
  1. Time stamp
  2. Fpz-Cz EEG
  3. Pz-Oz EEG
  4. EOG (electrooculogram)

Method: Extracting the Attributes

- Using the rule set, we derived several attributes to evaluate per-epoch to be used with a Naive Bayes classifier:
  - absolute value of the average of the amplitude
  - sums of amplitudes of frequency ranges
  - number of sleep spindle(s) detected
  - number of K-complexes detected

  With the per-epoch values for these attributes, a manual scorer can follow the R&K rules to assign a stage to the epoch.

Method: Calculating the Attributes

- We used a straightforward approach for calculating the average amplitudes per epoch.
- For the frequency items, we used a Fast Fourier Transform (FFT) on the epoch, then integrated over certain frequency ranges referred to in the R&K rule set:
  1. $0 - 2Hz$
  2. $2 - 7Hz$
  3. $7 - 12Hz$
  4. $12 - 14Hz$
- This yields an uneven "histogram" giving the sums of the amplitudes for those frequency ranges.
- Since each sleep stage has a unique frequency and amplitude signature, we would expect these values to adequately characterize the epoch.

Method: Sleep Spindles

- The method for detecting sleep spindles is as follows:
  1. A bypass filter is run for the frequency range between $12 - 14Hz$. This isolates the amplitudes of the wave which fall within that frequency range.
  2. For each decisecond, the root mean squared is computed.
  3. If the root mean square is greater than an experimentally pre-determined threshold (in this case, 10) for five deciseconds ($.5s$), then the waveform is labelled as a sleep spindle.
- Our sleep spindle detector achieved a plateau accuracy of 64%.

Method: K-Complexes

- For the K-complex detection, we first extracted sets of five values: $t_{start}, t_{max}, t_{mid}, t_{min}, t_{end}$.
- We then pruned this set using restrictions given from the K-complex definition and the R&K rules:
  - $t_{min} - t_{mid} > t_{mid} - t_{max}$
  - $t_{end} - t_{start} > .5s$
- We also made generalizing assumptions for small fluctuations in the recording within an epoch:
  - $\frac{|A_{max} - M| + |M - A_{min}|}{2} > \frac{\sum_i^N |A_i|}{N}$
- We adjusted the thresholds experimentally in order to achieve a plateau accuracy of 54%.

Method: Rule-Based Classifier #1

- To provide a basis of comparison against our Naive Bayes classifier, we created a direct rule-based classifier based on a direct translation of the Rechtschaffen & Kales rules using our attribute values.
- Our first rule-based classifier was loyal to the R&K rules, using all the attributes called for in the stage classification.
- Our first classifier achieved an accuracy of 64%.

Method: Rule-Based Classifier #2

- Our second rule-based classifier attempted to circumvent the requirement for a “representative frequency” for the epoch.
- It mainly relied upon the detection of sleep spindles to distinguish between Stage 2 and Stage 3 sleep, and mean amplitude for the other stages.

Method: Rule-Based Classifier *(Continued)* #2

- Extracting useful data from the signal in each epoch:
  - Frequencies:
    - 0-2 Hz (stages 3 and 4)
    - 2-7 Hz (stage 1)
    - 12-15 Hz (stage 2)
  - Amplitudes:
    - $10 - 20\mu V$ (stage 0)
    - $75 - 100\mu V$ (stages 3-4)
    - $100 - 200\mu V$ (stage 1)
  - Dominance of each frequency (stages 3-4)
  - Spindles
  - K-Complex

Method: Rule-Based Classifier *(Continued)* #2

- Rule-Based Ordering

(freq => 2, amp $<$ 75, duration $>$ 50)	stage 3
(freq => 2, amp $<$ 75, duration $<$ 50)	stage 4
(2 $>$ freq $>$ 7, amp $>$200)	stage 1
(12 $>$ freq $>$ 14, spindle, kcomplex)	stage 2
(10 $>$ amp $>$ 30)	stage 1

- Class-Based Ordering

Method: Naive Bayes Classifier

- Initially, we used a Naive Bayes classifier with 20 attributes: mean, max, frequency histogram, number of sleep spindles, number of K-complexes. These attributes were computed for both EEG recordings, and mean was computed for EOG to provide training data to classify REM stages.
- Since the calculation of the attributes was time- consuming, we split an EEG record in half to use for our training and validation sets.
- However, this initial classifier produced an accuracy of 6%. After examining the correlation matrix for the 20 attributes, we determined that this set of attributes violated the Naive Bayes assumption (they were not probabilistically independent). Thus, we pruned the attributes based values from the correlation matrix.

Method: Naive Bayes Classifier

- Experimentally, we tuned the attribute set until we achieved a plateau accuracy of 73% using the frequency histogram.
- This accuracy outdoes the direct rule classifiers and, given the efficiency of computing the amplitude-frequency integrals, is competitive with existing classifiers (which have accuracy ranges from 70-80%)!

Results

- The following table summarizes the accuracies of our measures and classifiers:

Spindles	64%
K-complexes	54%
Rule Classifier 1	64%
Rule Classifier 2	66%
Naive Bayes Classifier	73%

Conclusion

- During our exploratory analysis of direct-rule and Naive Bayesian classifiers, we found that Naive Bayes classifiers can in fact compensate for measurement errors of attributes used in direct-rule classifiers.
- In addition, the accuracy of our Naive Bayes classifier has interesting implications for the R&K rules; namely, the amplitude-frequency integrals seem to contain sufficient information to classify the epochs without using many of the other attributes referenced in the rule set.

References

- Susmakova, Kristina. "Human sleep and sleep EEG." Measurement Science Review 4.2 (2004): 59-74.
- Devuyst S., Dutoit T., Stenuit P., and Kerkhofs M. "Automatic K-complexes detection in sleep EEG recordings using likelihood thresholds." IEEE Engineering in Medicine and Biology Society (2010).
- Steffen Gais, Matthias Molle, Kay Helms, and Jan Born. "Learning-dependent increases in sleep spindle activity." The Journal of Neuroscience 22.15 (2002).