

SignalML 2.0

Paweł Kordowski
Zakład Fizyki Biomedycznej FUW
Pawel.Kordowski@fuw.edu.pl

Postawienie problemu:

- EEG i komputery – korzyści i trudności
- Porównywanie metod analizy danych
(różnice między pacjentami, wymagana duża weryfikacja)
- Dostęp do wyników badań

Pomysły:

- Jeden, wspólny format danych
(bez szans)
- Konwersja między formatami
(wiele różnych opisów formatów + wiele implementacji + straty)
- Metaformat danych, czyli SignalML
(standardowe opisy formatów + jedna implementacja)

Opis sygnałów:

- Częstość próbkowania (odstęp czasu)
- Liczba próbek
- Liczba kanałów i ich nazwy
- Pozycje sensorów
- Sygnał (ASCII, float, int, little(big)endian, różne długości i ułożenia próbek, prefiksy szesnastkowe, ósemkowe)
- Tagi/adnotacje
- Informacje o referencji - montażu
- Jednostki, gain i offset
- Id badanego, data zebrania danych

SignalML – technologie:

- XML

(nowe formaty bez rekompilacji, znaki lokalne – jednostki)

- XML Schema

(opis składni SignalML)

- XPath

(parsowanie innych plików XMLowych)

- Python

(składnia wyrażeń, typy zmiennych)

- Perl

(wyrażenia regularne)

- NumPy

(opis ułożenia bitów w plikach binarnych)

SignalML – zalety:

- Nowe formaty bez rekompilacji
- Brak konwersji – odczyt bezpośrednio z formatu
- Brak duplikacji danych
- Zwiększenie popularności formatów opisanych w SignalML
- Licencja Creative Commons

SignalML 2.0 vs 1.0:

- Obsługa wielu plików o różnych opisach
- Mapowanie próbek przy pomocy wyrażeń, które mogą zawierać parametry
- Specyfikowanie formatu próbek
- Parametry opisane przez wyrażenia
- Opis plików:
 - binarnych (pozycja i długość)
 - tekstowych (linia, numer słowa, wyrażenia regularne)
 - XMLowych (XPath)

SignalML – praktyka:

- Własny parser na podstawie opisu SignalML
- Svarog jako gotowa implementacja, własne algorytmy jako pluginy
- Rozwijane implementacje w Pythonie i Java

SignalML – informacje:

- SignalML Wiki

<http://signalml.org/wiki/>

- Svarog repo

<http://escher.fuw.edu.pl/git/svarog>

- Durka, P. J. and Ircha, D. (2004).

SignalML: metaformat for description of biomedical time series.

Computer Methods and Programs in Biomedicine,
76:253–259.

SignalML

przykład:

```
<?xml version="1.0"?>
<format>
  <header>
    <format id='PE-EASYS'/>
  </header>
  <file extension='*.d' type='binary' >
    <param id='datatype_width'>
      <expr>4</expr>
    </param>
    <param id='mapping' type='int'>
      <expr>(number_of_channels * sample_number + channel_number) * datatype_width + 16 * data_offset</expr>
    </param>
    <param id='magic'>
      <format>|S3</format>
      <offset>0</offset>
    </param>
    <assert id='magic_ok'>
      <expr>magic == "EAS"</expr>
    </assert>
    <param id='number_of_channels'>
      <format>>i1</format>
      <offset>16</offset>
    </param>
    <param id='sampling_frequency' type='float' units='Hz'>
      <expr>_sampling_frequency/100</expr>
    </param>
    <param id='_sampling_frequency'>
      <format>>i4</format>
      <offset>18</offset>
    </param>
    <param id='calibration_gain' type='float' units='µV'>
      <expr>_calibration_gain/100</expr>
    </param>
    <param id='_calibration_gain'>
      <format>>i1</format>
      <offset>25</offset>
    </param>
    <param id='data_offset'>
      <format>>i2</format>
      <offset>28</offset>
    </param>
  </file>
</format>
```


<Dziękuję za
uwagę />