

Okno główne programu podzielone jest na 3 obszary:

- z lewej strony Checkbox do wybierania składników wczytanych z pliku XML, które mają być użyte podczas podsumowania bazy,

- z prawej strony tabela z wynikami podsumowań, którą można dowolnie formatować. Przenosić kolejność tabel, zmieniać rozmiary kolumn oraz sortować podsumowania wg wartości danej kolumny,

- na dole obszaru okna mamy pasek narzędziowy składający się z

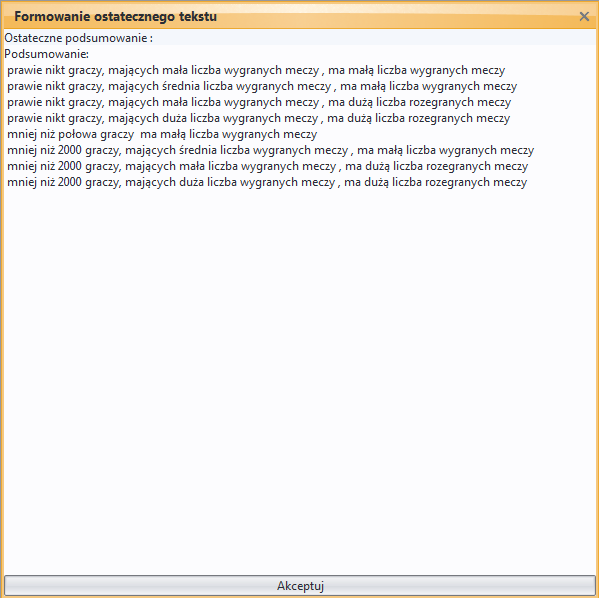
* Przycisku wczytaj dane – wczytujemy do pamięci programu dane z pliku .csv,
* Przycisku wczytaj drzewo – wczytujemy składniki czyli Kwantyfikatory Typu I, Kwantyfikatory Typu II, Kwalifikatory oraz Sumaryzatory z pliku XML,
* Spinner Pierwszy – określa możliwą ilość powiązanych kwalifikatorów w pojedynczym podsumowaniu,
* Spinner Drugi – określa możliwą ilość powiązanych sumaryzatorów w pojedynczym podsumowaniu,
* Przycisk ustaw wagi – wywołuje okno w którym ustawiamy wagi miar,
* Przycisk utwórz histogramy – tworzymy histogramy statystyk bazy oraz tworzymy wykresy funkcji przynależności zaznaczonych w Checkboxie składników,
* Przycisk Generuj – po wczytaniu bazy danych oraz składników XML umożliwia uruchomienie głównej funkcji programu czyli wygenerowanie podsumowań,
* Przycisk Popraw i zapisz- pozwala na edycję wybranych podsumowań oraz zapisanie ich do wybranego pliku .txt . Wybieranie podsumowań następuje po naciśnięciu na krotkę podsumowania w oknie głównym.

Okno ustawiania wag :



Po określeniu wag zatwierdzamy je poprzez przycisk Akceptuj i zamykamy okno.

Okno edycji podsumowań :



Po edycji podsumowań zatwierdzamy je poprzez przycisk Akceptuj, wybieramy plik tekstowy do zapisu danych i zamykamy okno.

Składniki ( kwantyfikatory, kwalifikatory, sumaryzatory ) z których korzystamy są definiowane w pliku elements.xml znajdującym się w głównym folderze projektu.

Format zapisu kwantyfikatora typu I :

<quantifierTypeOne type="relative">

<name>prawie nikt</name>

<universe type="continuous">

<start>0.0</start>

<end>1.0</end>

<step>0.01</step>

</universe>

<membership type="rectengular">

<a>0</a>

<b>0.15</b>

</membership>

</quantifierTypeOne>

Format zapisu kwantyfikatora typu II :

<quantifierTypeTwo type="absolute">

<name>więcej niż 100</name>

<universe type="continuous">

<start>0.0</start>

<end>10000</end>

<step>1</step>

</universe>

<universe2 typeofmembership="triangular" width="0.2">

<start>0.0</start>

<end>1.0</end>

<step>0.01</step>

</universe2>

<membership type="triangular">

<a>2.0</a>

<b>3.0</b>

<c>4.0</c>

</membership>

</quantifierTypeTwo>

Format zapisu kwalifikatora lub sumaryzatora:

<qualifier>

<name>mała liczba wygranych meczy</name>

<label>mało wygranych</label>

<attribute>wins</attribute>

<universe type="continuous">

<start>0.0</start>

<end>55.0</end>

<step>0.1</step>

</universe>

<membership type="rectangular">

<a>0.0</a>

<b>5.0</b>

</membership>

</qualifier>

Opis implementacji programu:

Projekt został utworzony w języju Java przy użyciu frameworka NetBeans IDE 7.01 korzystając z JDK 7. Oprócz standardowych bibliotek zawartych w tej technologii skorzystaliśmy z dodatkowych bibliotek :

- substance – do tworzenia modelu gui source: http://java.net/projects/substance/,

- trident – do tworzenia modelu gui source: <http://kenai.com/projects/trident/pages/Home> ,

- log4j-1.2.16 – do logowania danych na konsolę source: <http://logging.apache.org/log4j/1.2/download.html> ,

- lablib-checkboxtree-3.2 – do utworzeniach drzewa CheckBox source: <http://www.essi-lab.eu/projectsSites/lablib-checkboxtree/apidocs/overview-summary.html> ,

- jfreechart-1.0.14 – do tworzenia histogramów oraz wykresów punktowych source: <http://www.jfree.org/jfreechart/> ,

- jcommon-1.0.17 – biblioteka wymagana przy korzystaniu z niektórych funkcji jfreechart source: <http://jfree.org/jcommon/> ,

Podział na pakiety jest następujący :

- data – klasy związane z obsługą danych,

- gui – klasy związane z elementami gui,

- io – klasy związane z odczytem i zapisem danych w formatach .csv, .txt i .xml,

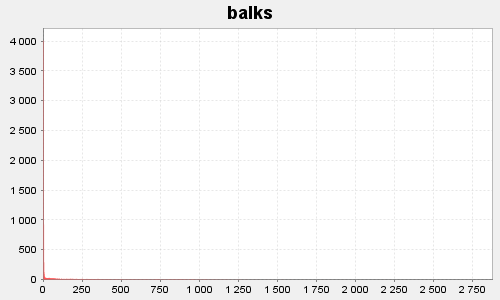
- main – klasy zarządzające działaniem programu,

- summarization – klasy związane z podsumowaniem, czyli implementacja miar ( typu I i typu II), elementów lingwistycznych ( zmiennej lingwistycznej, wartości lingwistycznej typu I i typu II oraz modyfikatora lingwistycznego), składników podsumowan ( wszystkie typy sumaryzatorów, kwalifikatorów i kwantyfikatorów ) oraz podsumowanie typu I oraz podsumowanie typu II,

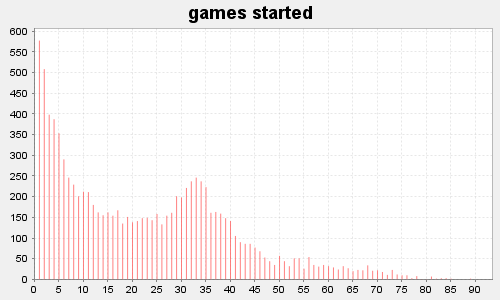
Opis zestawu danych używanego w projekcie :

Danymi w projekcie jest zestaw statystyk dotyczących baseballistów, pobrany ze strony <http://baseball1.com/statistics/> . Pierwotna baza posiada 40 tysięcy rekordów zawierajych 5 atrybutów zawierajace wartości własne oraz 25 atrybutów zawierające dane liczbowe możliwe do podsumowania. Zredukowany przez nas zestaw zawiera 10 tysięcy oraz 10 atrybutów uznanych przez nas na podstawie wytworzonych histogramów za najlepsze możliwe zbiory.

Histogram przykładowego odrzuconego przez nas atrybutu:



Histogram przykładowego wziętego przez nas pod uwagę atrybutu:



Lista wszystkich konkretnych rozważanych atrybutów:

1. Game started,
2. Complete games.
3. Earned runs,
4. Home runs,
5. Wins,
6. Losses,
7. Walks,
8. Strike outs,
9. Hits,
10. Oponnent batting average.