

Licencjacki projekt programistyczny 2011

Program do wykonywania obliczeń statystycznych związanych z grą go

Dokumentacja programisty

Wojciech Jedynak

Wrocław, 28 czerwca 2011

Spis treści

1	Wprowadzenie	3
1.1	Cel dokumentacji	3
2	Organizacja projektu	3
2.1	Ogólny opis	3
2.2	Konfiguracja programu	3
2.2.1	Opis formatu pliku CONFIG	3
2.3	Baza danych	3
2.3.1	Tabela go_stat_data	4
2.4	Struktura modułów	4
2.4.1	Data.SGF.Parsing i Data.SGF.Types	4
2.4.2	Transformations	4
2.4.3	SgfBatching	4
2.4.4	Lang	4
2.4.5	Configuration	4
2.4.6	Pages	4
2.4.7	DB	4
2.4.8	Server	4
2.4.9	Main	4
2.5	Pozostałe pliki	5
3	Kompilacja i testowanie	5
4	Wykorzystane biblioteki i narzędzia pomocnicze	5
4.1	Biblioteki, pakiety, moduły	5
4.1.1	Śledzenie zależności (Cabal)	5
4.1.2	Serwer HTTP (Happstack)	5
4.1.3	Testowanie (HUnit, QuickCheck, test-framework)	5
4.1.4	Zarządzanie bazą danych (HDBC, HDBC-postgresql, HDBC-sqlite3)	5
4.1.5	Analiza leksykalna (Parsec)	5
4.1.6	Generowanie html (xhtml 3000)	5
4.1.7	Pozostałe (filemanip, strict, mtl)	5
4.2	Narzędzia	5
4.2.1	Git i portal http://github.com	5
4.2.2	Latex	5

1 Wprowadzenie

1.1 Cel dokumentacji

Celem niniejszego dokumentu jest takie przedstawienie struktury projektu, aby umożliwić jego modyfikacje oraz utrzymywanie programistom, którzy znają Haskell, ale nie należeli do początkowego zespołu. Wykaz użytych bibliotek powinien być pomocą, gdy instalacja oprogramowania nie powiedzie się i konieczna będzie kompilacja programu ze źródeł. Dodatkowo życzeniem autora jest nakreślenie wykonanej pracy tak, aby zainteresowane osoby były w stanie (w razie potrzeby) na wykorzystanie opisanych tu rozwiązań w swoich projektach.

2 Organizacja projektu

2.1 Ogólny opis

Program został napisany niemal w całości w Haskellu. Uruchamiany jest za pomocą wiersza poleceń, a komunikacja z użytkownikiem odbywa się za pomocą interfejsu WWW. Lista katalogów, które stanowią kolekcję plików SGF zapisywana jest do pliku konfiguracyjnego, wstępnie przetworzone (*znormalizowane*) gry są przechowywane w bazie danych. Dialog z użytkownikiem może odbywać się w języku polskim bądź angielskim.

2.2 Konfiguracja programu

W programie potrzebujemy przechowywać dwie informacje: jakiej bazy danych używa (chce używać) użytkownik i gdzie znajduje się jego kolekcja zapisów partii go, które chciałby analizować naszym programem. Do przechowywania ww. danych używamy bardzo prostego, autorskiego formatu. Funkcje związane z wczytywaniem, analizą leksykalną oraz zapisywaniem znajdują się w module Configuration (w pliku `src/Configuration.hs`).

2.2.1 Opis formatu pliku CONFIG

Format pliku jest opisywany przez poniższą gramatykę w postaci EBNF:

```
config ::= declaration*
declaration ::= db | dirs | '-' *anything*
db = dbserver ':' dbversion
dbversion = sqlite3 ';' path ';' | postgresql ';'
dirs = gamedirs ':' path ';'
```

Przykładowy plik konfiguracyjny:

```
-- new config
--dbserver:postgresql;
dbserver:sqlite3;/home/wojtekdatabases/games.db;
gamedirs:/home/wojtek/data/;
```

2.3 Baza danych

Program pozwala na użycie baz SQLite3 oraz PostgreSQL. Możliwa jest zmiana decyzji co do tego, która z nich jest używana; wymagane jest wówczas przebudowanie zawartości, gdyż protokół komunikacyjny bazy PostgreSQL nie jest kompatybilny z protokołem bazy SQLite3. Ani użytkownik ani programista nie muszą zajmować się ręczną administracją bazy danych: służy do tego odpowiedni moduł (DB w pliku `src/DB.hs`).

Używana jest jedna tabela o nazwie `go_stat_data`.

2.3.1 Tabela go_stat_data

Opis pól tabeli go_stat_data

Pole	Typ	NULL dozwolone?	Opis
id	PRIMARY KEY	nie	unikatowy identyfikator gry
winner	CHAR	nie	zwycięzca gry ('b' lub 'w')
moves	VARCHAR(700)	nie	znormalizowany przebieg rozgrywki
path	VARCHAR(255)	nie	względna ścieżka do gry
b_name	VARCHAR(30)	nie	pseudonim (nazwisko) czarnego
w_name	VARCHAR(30)	nie	pseudonim (nazwisko) białego
b_rank	VARCHAR(10)	tak	ranking czarnego
w_rank	VARCHAR(10)	tak	ranking białego

Jedyne pola, która nie są wymagane to pola b_rank i w_rank. Wynika to z tego, że na niektórych serwerach do gry w go nie jest wymagane podanie swojego orientacyjnego poziomu ani rankingu.

2.4 Struktura modułów

Lista modułów wchodzących w skład projektu:

2.4.1 Data.SGF.Parsing i Data.SGF.Types

Analiza plików SGF

2.4.2 Transformations

Przekształcenia matematyczne i normalizacja ruchów

2.4.3 SgfBatching

Konwersja plików SGF do formatu, który będzie łatwo zapisać w bazie danych

2.4.4 Lang

Komunikaty w języku polskim i angielskim

2.4.5 Configuration

Obsługa pliku CONFIG.

2.4.6 Pages

Tworzenie szablonów stron WWW.

2.4.7 DB

Zarządzanie bazą danych.

2.4.8 Server

Serwer HTTP.

2.4.9 Main

Punkt startowy aplikacji.

2.5 Pozostałe pliki

Pliki css, javascript, obrazki

3 Kompilacja i testowanie

4 Wykorzystane biblioteki i narzędzia pomocnicze

4.1 Biblioteki, pakiety, moduły

4.1.1 Śledzenie zależności (Cabal)

4.1.2 Serwer HTTP (Happstack)

4.1.3 Testowanie (HUnit, QuickCheck, test-framework)

4.1.4 Zarządzanie bazą danych (HDBC, HDBC-postgresql, HDBC-sqlite3)

4.1.5 Analiza leksykalna (Parsec)

4.1.6 Generowanie html (xhtml 3000)

4.1.7 Pozostałe (filemanip, strict, mtl)

4.2 Narzędzia

W niniejszej sekcji wymieniono i pokrótce opisano najważniejsze narzędzia, które pozwoliły ukończyć projekt.

4.2.1 Git i portal <http://github.com>

Git to system do kontroli wersji autorstwa Linusa Torvaldsa (TODO: sprawdzić pisownię). Github to portal, który pozwala na przechowywanie kodu źródłowego (ogólnie: repozytoriów kodu zarządzanych przez git) i udostępnienie go innym programistom. Poprzez użycie tych zasobów rozwiązałem kwestię składowania projektu i mogłem swobodnie eksperymentować: nietrafione zmiany można było wycofać jednym poleceniem.

4.2.2 Latex

System L^AT_EX pozwolił na utworzenie dokumentacji w formacie pdf, który jest standardem w informatyce.

5 Słownik

Czarny

Biały

plik SGF

interfejs WWW

serwer HTTP

Haskell

zapis gry

ranking (gracza)

znormalizowany przebieg gry

go

serwer do gry w go

gramatyka bezkontekstowa

postać EBNF