

Frontend eines Agilent-Parsers erstellen

Belegarbeit

eingereicht am Fachbereich

Informatik

der Hochschule Zittau/Görlitz (HAW)

als Prüfungsleistung im Fach

Data Mining 2

vorgelegt von:

Christof Ochmann (35989)

Ingo Körner (40586)

Görlitz, 07. Februar 2013

Betreuer: Prof. ten Hagen

Abstract

In diesem Projekt...

Inhaltsverzeichnis

Literaturverzeichnis	V
1 Allgemeines	1
1.1 Einleitung	1
1.2 Aufgabenstellung	1
1.3 Relevanz des Forschungsgegenstandes	2
1.4 Der aktuelle Wissensstand	2
2 Umsetzung	3
2.1 Analyse	3
2.2 Entwurf	3
2.3 Text File Encoding	5
2.4 Unstimmigkeiten im Agilent-Format	5
A Arbeitsaufteilung	6
B Eigenständigkeitserklärung	7

Abbildungsverzeichnis

2.1	Analyseklassendiagramm Agilent-Parser	3
2.2	Entwurfsklassendiagramm Agilent-Parser	4

Abkürzungsverzeichnis

JVM Java Virtual Machine

Literaturverzeichnis

- [1] Martin, Robert C. (2008): Clean Code: A Handbook of Agile Software Craftsmanship. Prentice Hall International
- [2] Freeman, Eric (2007): Entwurfsmuster von Kopf bis Fuß. O'REILLY
- [3] <http://www.cs.waikato.ac.nz/ml/weka/arff.html> (08.06.2012)

1 Allgemeines

1.1 Einleitung

Ziel dieses Projektes ist es, das Frontend für einen Agilent-Compiler zu schreiben. Dabei soll das Frontend Agilent-Logdateien einlesen und sie in ein Intermediate-Format umwandeln.

1.2 Aufgabenstellung

In diesem Projekt sollen gegebene Agilent-Logdateien geparkt werden. Aus den geparkten Zeilen soll ein Baum in einem Intermediate-Format erstellt werden. Der erzeugte Baum wird im Hauptspeicher des Rechners gehalten. Die Weiterverarbeitung des Baumes im Intermediate-Format erfolgt in einem anderen Projekt und ist nicht Gegenstand dieser Arbeit. In diesem Projekt muss das Intermediate-Format nicht entworfen werden. Statt dessen wird es von Felix Deutschmann und Daniel Horbach übernommen. Ein Baum im Intermediate-Format wird immer nur aus einer Agilent-Logdatei erzeugt, d.h. es soll nicht ein Baum aus zwei oder mehreren Agilent-Logdateien erzeugt werden. *Agilent – Logdatei – > Agilent – Parser – > Baum im Intermediate – Format.*

1.3 Relevanz des Forschungsgegenstandes

Der Forschungsgegenstand dieser Arbeit ist, ein Compiler-Frontend für Agilent-Logdateien zu erstellen. Der Forschungsgegenstand ist relevant, da bisher kein Compiler-Frontend für das Umwandeln von Agilent-Logdateien in das Intermediate-Format vorliegt. Darüberhinaus müssen für das Erstellen des Frontends technische Probleme gelöst werden. Ziel der Forschung ist es, einen Frontend zu entwerfen, dass das Agilent-Logformat in ein Intermediate-Format überführt.

1.4 Der aktuelle Wissensstand

Noch nicht vorhandene Kenntnisse über das Parsen von Agilent-Logdateien werden aus der Format6.pdf gewonnen. In der Format6.pdf wird das Agilent-Logformat beschrieben.

2 Umsetzung

2.1 Analyse

Ziel dieses Projektes ist es, ...

In Abbildung 2.1 auf Seite 3 ist das Analyseklassendiagramm vom Agilent-Parser zu sehen.

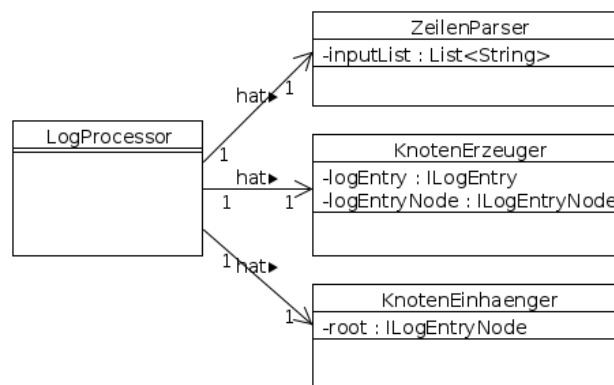


Abbildung 2.1: Analyseklassendiagramm Agilent-Parser

2.2 Entwurf

Ziel dieses Projektes ist es, ...

In Abbildung 2.2 auf Seite 4 ist das Entwurfsklassendiagramm vom Agilent-Parser

zu sehen.

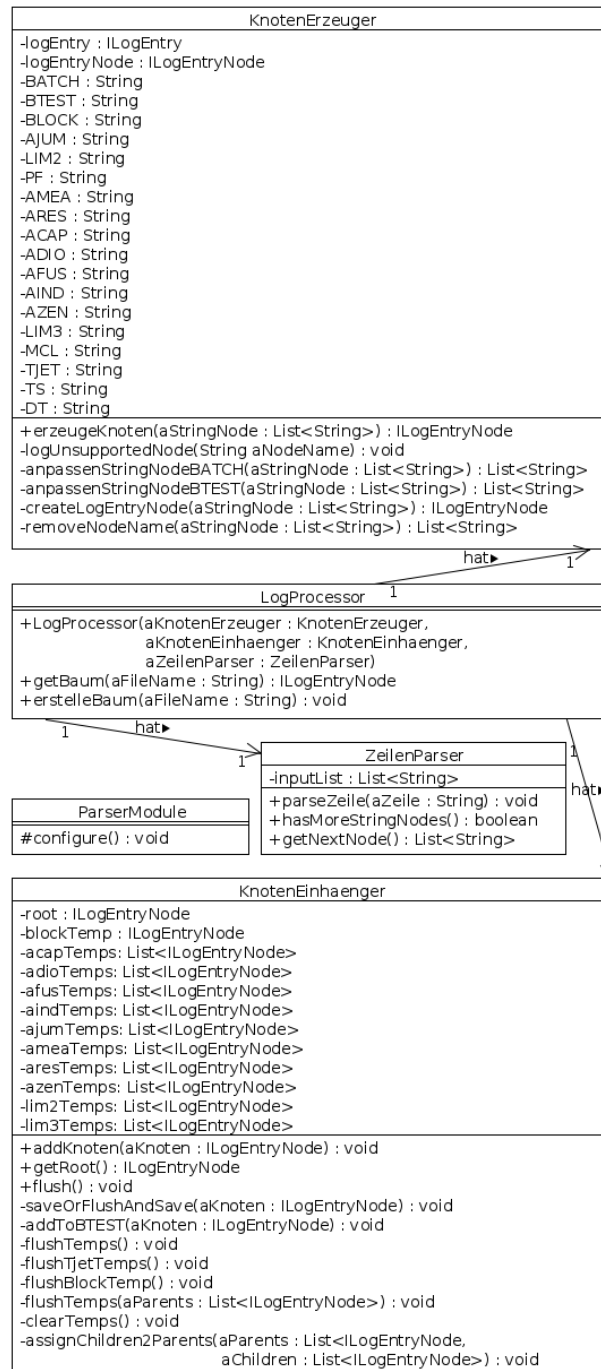


Abbildung 2.2: Entwurfsklassendiagramm Agilent-Parser

2.3 Text File Encoding

Unter Ubuntu 10.04 ist in Eclipse Juno das Text File Encoding standardmäßig auf UTF-8 gesetzt. Die kann beim Kopieren von Strings aus dem Agilent-Logformat zu Fehlern führen, da bestimmte UTF-8 Zeichen im Java-Editor unsichtbar sind. Um unsichtbare Zeichen im Java-Editor sichtbar zu machen, sollte unter *Eclipse* – *> Preference* – *> General* – *> Workspace* das “Text File Encoding” von UTF-8 auf ISO-8859-1 umgestellt werden.

2.4 Unstimmigkeiten im Agilent-Format

In der Format06.pdf auf Seite 35 stimmt bei BATCH die Anzahl der Attribute nicht mit der Legende überein. Auch in den tatsächlichen Agilent-Logdateien stimmt die Anzahl der Attribute nicht immer mit der Anzahl der Attribute überein, wie sie in der Format06.pdf beschrieben werden.

In der Format06.pdf auf Seite 40 stimmt bei BTEST die Anzahl der Attribute nicht immer mit der tatsächlich vorkommenden Anzahl in der Agilent-Logdatei überein.

A Arbeitsaufteilung

Arbeit	C. Ochmann	I. Körner
Abstract		0
Einleitung		1.1
Aufgabenstellung		1.2
Forschungsgegenstand		1.3
akt. Wissensstand		1.4
Zusammenfassung		??
Ausblick		??

Tabelle A.1: Aufteilung

B Eigenständigkeitserklärung

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe. Mir ist bekannt, dass jede Form des Plagiats mit der Note 5 (Betrugsversuch) bewertet wird.

Ochmann, Christof

Unterschrift:

Körner, Ingo

Unterschrift: