



Software Engineering

Teamprojekt

Projektbeschreibung, Wintersemester 2018/2019

Abgabetermin: 31. Januar 2019

Projekt 1 Open source Graphstruktur von SBML Dokument-Inhalten

1.A Hintergrund

Die „KEGG PATHWAY“-Datenbank [1] erfreut sich seit langem großer Beliebtheit und wird häufig als erste Informationsquelle für die Erstellung von Modellen biologischer Prozesse genutzt. KEGG stellt seine Stoffwechsel- und Signalwege in einem eigenen XML-basierten Format [2] namens KGML zur Verfügung, dessen Hauptzweck darin besteht, die visuelle Darstellung dieser Karten im Lehrbuchstil maschinenlesbar zu speichern. Der Schwerpunkt von KEGG liegt damit eher auf der Identifizierung allgemeiner Mechanismen, als spezifische Prozesse für betrachtete Organismen oder Gewebe abzubilden.

Die Software *KEGGtranslator* [4] wandelt KGML-Dateien in mehrere besser dokumentierte Standard-Formate um, darunter BioPAX [5], SBGN-ML [6] und SBML [7]. In diesem Prozess werden viele zusätzliche Informationen von KEGG abgerufen und integriert, die in den ursprünglichen KGML-Dateien nicht enthalten sind. Dadurch sind die resultierenden Modelle viel umfassender als die original KGML-Dateien und können in anderen Workflows und Programmen eingesetzt und wiederverwendet werden. Zur internen Repräsentation und Erstellung der graphischen Darstellungen der Daten verwendet KEGGtranslator in der aktuellen Version 2.5 die nicht quelloffene Bibliothek yFiles [8]. Um die Software in künftigen Releases vollständig quelloffen anbieten zu können, muss diese Bibliothek daher durch eine quelloffene Variante ersetzt werden.

Als Teil von Google Open Source [9] werden die Google Core Bibliotheken für Java [10] entwickelt. Diese beinhalten mit dem *Graph-Package* [11] ein mächtiges Datenmodell zur Modellierung von Graphen.

1.B Ziele

Ziel dieses Projektes ist die Entwicklung einer eigenständigen Software, die die folgenden Bedingungen erfüllt:

- Verwendung von Java (mind. Version 8)
- Einlesen von SBML Dokumenten des „Level 3 Version 1 Core“ [12]
- Einlesen von SBML Dokumenten mit der Erweiterung „Qualitative Modelle“ [13]
- Verwendung der Bibliothek JSBML [14] zur Verarbeitung der SBML-Daten
- Verwendung der Bibliothek „Google Guava Graph“ zur Speicherung der eingelesenen Daten
- Graphische Benutzeroberfläche zur Steuerung der Software und Darstellung der Graphen

1.C Referenzen

- [1] KEGG PATHWAY, <http://www.genome.jp/kegg/pathway.html>, Zugriff am 19.Oktober 2018.
- [2] XML, https://de.wikipedia.org/wiki/Extensible_Markup_Language, Zugriff am 19.Oktober 2018.
- [3] KGML, <http://www.genome.jp/kegg/xml/>, Zugriff am 19.Oktober 2018.
- [4] KEGGtranslator, visualizing and converting the KEGG PATHWAY database to various formats <http://www.cogsys.cs.uni-tuebingen.de/software/KEGGtranslator/doc/index.html>, Zugriff am 19.Oktober 2018.
- [5] BioPAX, <http://biopax.org>, Zugriff am 19.Oktober 2018.
- [6] SBGN-ML, <http://www.sbgn.org>, Zugriff am 19.Oktober 2018.
- [7] SBML, <http://sbml.org>, Zugriff am 19.Oktober 2018.
- [8] Graphbibliothek yFiles, <https://www.yworks.com>, Zugriff am 19.Oktober 2018.
- [9] Google open source, <https://opensource.google.com/>, Zugriff am 19.Oktober 2018.
- [10] Google core libraries for Java, <https://github.com/google/guava>, Zugriff am 19.Oktober 2018.
- [11] Google Guava Graphs, <https://github.com/google/guava/wiki/GraphsExplained>, Zugriff am 19.Oktober 2018.
- [12] SBML Specifications, <http://sbml.org/Documents/Specifications>, Zugriff am 19.Oktober 2018.
- [13] Qualitative Models for SBML, http://sbml.org/Documents/Specifications/SBML_Level_3/Packages/qual, Zugriff am 19.Oktober 2018.
- [14] JSBML, <http://github.com/sbmlteam/jsbml>, Zugriff am 19.Oktober 2018.

Zögern Sie im Falle von Schwierigkeiten nicht

- in den Büros C320, C321 oder C305 vorbeizukommen,
- (0 70 71) 29-7 04 59, 7 04 60 oder 7 04 15 anzurufen,
- oder eine E-Mail an {draeger,hamm,tiede}@informatik.uni-tuebingen.de zu schreiben.