

Übungen zu Algorithmische Bioinformatik: Netzwerke, Graphen und Systeme

Blatt 1

Abgabetermin: Freitag, 03.05.2019, 9 Uhr
Persönlich oder per Upload-Formular unter
www.bio.ifi.lmu.de/studium/ss2019/vlg_nginx/uebungsabgabe

Hinweise: Aufgaben, die für den Notenbonus zählen, sind als Bonus-Aufgabe gekennzeichnet. Pro Aufgabe gibt es 10 Punkte, außer es ist explizit anders angegeben. Bonus-Aufgaben müssen einzeln bearbeitet werden.

Aufgabe 1: (Entwurf der Datenstruktur, Bonus-Aufgabe):

Entwerfen und implementieren Sie eine Datenstruktur für das Arbeiten mit Graphen. Ihre Implementation muss detailliert und umfangreich genug sein, um damit folgende Aufgaben durchführen zu können:

- (a) Einlesen eines Graphen in den Speicher gegeben ein sinnvolles Eingabeformat.
- (b) Hinzufügen eines Knoten und zusätzlicher Annotation (z.B. Datenbank-ID).
- (c) Hinzufügen einer Kante gegeben der Endknoten und zusätzlicher Annotation (z.B. funktionelle Beschreibung).
- (d) Zählen der Kanten.
- (e) Bestimmen der Nachbarn eines Knoten.
- (f) Testen für zwei gegebene Knoten u, v (z.B. über ID) ob diese durch eine Kante verbunden sind.

Abzugeben sind:

- Ihr Programm (in einer zip/tar.gz Datei, siehe Abgabeformular). Das Programm muss im CIP-Pool laufen.

- Eine README Datei in der beschrieben ist, wie das Programm aufzurufen ist (in der zip/tar.gz Datei).
- Eine Beschreibung ihrer Datenstruktur (in einer pdf-Datei) mit Antworten auf die folgenden Fragen:
 - (a) Was sind die Laufzeitanforderungen Ihrer Implementierung der Datenstruktur für die oben angegeben Aufgaben in Θ -Notation?
 - (b) Wie groß sind die Platzanforderungen Ihrer Implementierung der Datenstruktur in Θ -Notation?

Aufgabe 2 (Protein-Interaktions Netzwerke, Bonus-Aufgabe):

Erweitern Sie Ihr Programm so, dass Sie ein Protein-Interaktions Netzwerk im PSI MITAB level 2.5-Format der BioGRID Datenbank (Database of Interacting Proteins, <https://thebiogrid.org>) einlesen können und folgende Aufgaben durchführen können:

- (a) Erzeugen Sie eine grafische Darstellung des Histogramms der Knotengrade eines Netzwerks (z.B. mit *gnuplot* oder *R*). Dies soll automatisiert erfolgen - also EIN Programmaufruf erzeugt Histogramm und Abbildung!
- (b) Visualisieren Sie einen (Ausschnitt eines) Graphen mit Graphviz, gegeben einer Knoten- und Kantenmenge (Dokumentation auf www.graphviz.org). Dies soll automatisiert erfolgen - also EIN Programmaufruf erzeugt die Graphviz-kompatible Eingabedatei und Abbildung!

Wenden Sie Ihre Implementation auf das BioGRID Network für Mensch an (erhältlich unter <https://downloads.thebiogrid.org/Download/BioGRID/Release-Archive/BIOGRID-3.5.171/BIOGRID-ORGANISM-3.5.171.mitab.zip>). Erzeugen Sie das Histogramm der Knotengrade dieses Netzwerks und visualisieren Sie den durch den Knoten SRSF3 und dessen Nachbarn induzierten Subgraphen. Hierbei sollen jeweils nur direkte Interaktionen berücksichtigt werden (psi-mi:“MI:0407”(direct interaction)).

Nötige Commandlineflags: `-file tab-Datei -o Ausgabe-Ordner ...`

Abzugeben sind:

- Ihr Programm (in der zip/tar.gz Datei). Das Programm muss im CIP-Pool laufen. Graphviz muss nicht mit abgegeben werden.
- Eine README Datei in der beschrieben ist, wie das Programm aufzurufen ist (in der zip/tar.gz Datei).
- Die mit ihrem Programm erzeugte Abbildungen (in der pdf-Datei).