



YaccConstructor

Задачи на осенний семестр 2017

Автор: Семён Григорьев

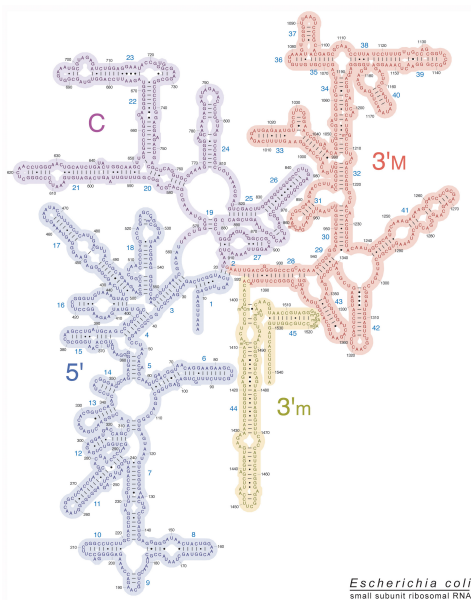
Лаборатория языковых инструментов JetBrains
Санкт-Петербургский государственный университет
Математико-механический факультет

8 сентября 2017г.

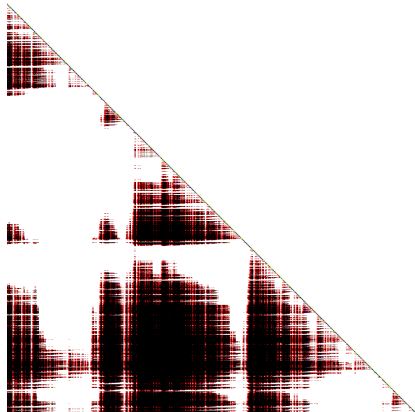
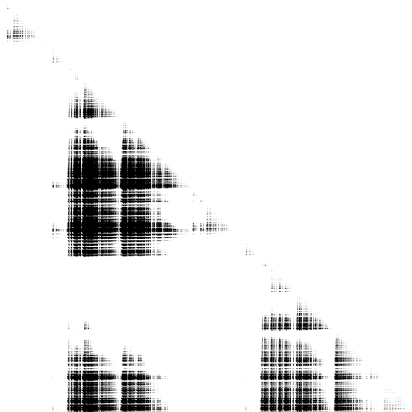
- Исследования в области формальных языков
- Открытый исходный код
 - ▶ <https://github.com/YaccConstructor>
- Основной язык разработки — F#

Применение методов “машинного обучения” к классификации цепочек (16s) на основе их вторичной структуры

Вторичная структура



Примеры изображений



- Обработка (классификация, кластеризация и т.д.) изображений, представляющих информацию о вторичной структуре
- Обработка (классификация, кластеризация и т.д.) векторных представлений данных о вторичной структуре
- Подробное описание задач: <https://goo.gl/A2хуае>
- “Направления развития”: курсовая, диплом, публикации

Требования к знаниям и навыкам

- Методы и алгоритмы обработки данных, статистические методы, машинное обучение, обработка изображений
- Умение читать и понимать научные статьи
- Умение читать и понимать чужой код

Синтаксический анализ графов на основе матричных операций

Задачи

- Цель: изучение возможностей построения эффективных алгоритмов синтаксического анализа графов
- Задачи
 - ▶ Исследование алгоритма синтаксического анализа линейного входа, основанного на матричных операциях, позволяющего повысить эффективность использования параллельных вычислений
 - ▶ Получение оценки сложности алгоритма синтаксического анализа графов в случае использования разреженных матриц для представления данных
 - ▶ Исследование возможности использовать алгоритма приближённого алгоритма перемножения матриц для приближённого синтаксического анализа графов
 - ▶ Оптимизация синтаксического анализа для DAG-ов и “почти DAG-ов”
- Подробное описание задач: <https://goo.gl/MB18nX>
- “Направления развития”: курсовая, диплом, публикации

Требования к знаниям и навыкам

- Теоретическая подготовка в области алгоритмов и анализе сложности, основы (линейной) алгебры
- Умение читать и понимать научные статьи
- Умение читать и понимать чужой код
- Умение писать научные статьи/отчёты
- Знакомство с функциональным программированием (F#, OCaml, Haskell и т.д.)
- Навыки работы с Git/GitHub

Парсер-комбинаторы для запросов к графовой базе данных

- Цель: использование парсер-комбинаторов для запросов к графовым базам данных
- Задачи
 - ▶ Интеграция библиотеки парсер-комбинаторов на Scala с графовой базой данных Neo4J
 - ▶ Подготовка данных и проведение экспериментального исследования полученного решения
- Подробное описание задач: <https://goo.gl/EaYKvY>
- “Направления развития”: курсовая, публикации

Требования к знаниям и навыкам

- Знакомство с языком программирования Scala
- Умение читать и понимать чужой код
- Базовые знания в области формальных языков, синтаксического анализа, теории графов
- Навыки работы с Git/GitHub

- Почта: `rsdpisuy@gmail.com`
- Исходный код YaccConstructor:
`https://github.com/YaccConstructor`