



Синтаксический анализ для поиска в метагеномных сборках

Автор: Семён Григорьев

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

11 мая 2016г.

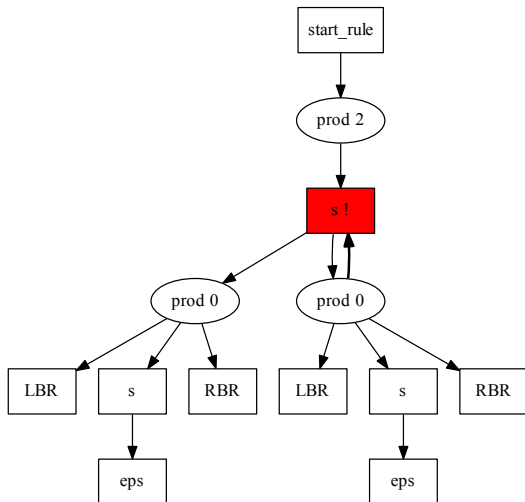
- Исследования в области лексического и синтаксического анализа
- Открытый исходный код
 - ▶ <https://github.com/YaccConstructor>
- Основной язык разработки — F#



Динамически формируемый код

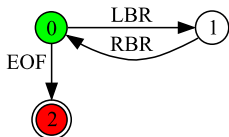
```
string res = "";  
for(i = 0; i < 1; i++)  
    res = "()" + res;  
use(res);
```

Лес разбора (SPPF)



Аппроксимация

("()")*



Грамматика

```
start ::= s  
s ::= LBR s RBR s  
s ::= ε
```

Задача: применение анализа строковых выражений для JavaScript eval

- Цель: трансляция стандартного eval в "безопасный"
 - ▶ Martin Lester. Information Flow Analysis for a Dynamically Typed Functional Language with Staged Metaprogramming
 - ▶ Martin Lester. Analysing Eval using Staged Metaprogramming
- Исследовательская задача: диплом, публикации
- Разбивается на 2 подзадачи
 - ▶ Получение аппроксимации
 - ▶ Трансляция SPPF в "безопасный eval"

Задача: использование SPPF в абстрактном синтаксическом анализе

- Абстрактный синтаксический анализ — один из подходов к анализу динамически формируемого кода
 - ▶ K. G. Doh, H. Kim, D. A. Schmidt. Static Validation of Dynamically Generated HTML Documents Based on Abstract Parsing and Semantic Processing
 - ▶ E. Verbitskaia, S. Grigorev, D. Avdyukhin. Relaxed Parsing of Regular Approximations of String-Embedded Languages
- Реализовать вычисление семантики по статьям
- Сравнить с нашим подходом
- А можно ли использовать SPPF
- Диплом, публикации

Средства для сертификационного
программирования

или

$$F\# + F^* = ?$$

- F^* (<https://www.fstar-lang.org/tutorial/>)
- Coq
- Agda
- ...

Задача: объединение $F\#$ и F^*

- $F\# + F^* = F\#^*$
- Парсер для $F\#^*$
- Транслятор из AST $F\#$ в AST F^*
- Разбивается на подзадачи (до 3 человек)
- Диплом (вся задача), публикации

Задача: поддержка $F\#^*$ в Microsoft Visual Studio

- Поддержать в модели проекта, редакторе, отладчике
 - ▶ Создание файлов, шаблоны
 - ▶ Подсветка синтаксиса
 - ▶ Сообщения об ошибках, подсветка ошибок
 - ▶
- Разбивается на подзадачи (до 4 человек)
- Диплом (вся задача), публикации

Задача: межъязыковое взаимодействие $F\#$ и F^*

- Использовать функции, написанные на F^* в $F\#(.NET)$
- Использовать функции, написанные на $F\#(.NET)$ в F^*
- Сохранить типизацию/вывод типов
- Разбивается на подзадачи (до 2 человек)
- Диплом (вся задача), публикации

- Почта: `rsdpisuy@gmail.com`
- Исходный код YaccConstructor:
`https://github.com/YaccConstructor`
- Google+ сообщество: `https://goo.gl/DuPWkM`