



Применение абстрактного синтаксического анализа для трансляции динамически формируемых строк

Автор: Григорьев Семён Вячеславович

Санкт-Петербургский государственный университет
Математико-Механический факультет
Кафедра системного программирования

25 апреля 2012г.

Встроенные языки широко распространены.

- Встроенный SQL
- Генерация HTML
- DSL
- ...

Описание проблемы

- Динамически генерируемые строки – тоже код на некотором формальном языке программирования, который необходимо соответствующим образом обрабатывать.

Описание проблемы

- Динамически генерируемые строки – тоже код на некотором формальном языке программирования, который необходимо соответствующим образом обрабатывать.
- После преобразований необходимо гарантировать:
 - ▶ синтаксическую корректность
 - ▶ корректность по отношению к изменениям "внешнего" кода
 - ★ Переименование объектов
 - ★ Удаление объектов

Описание проблемы

- Динамически генерируемые строки – тоже код на некотором формальном языке программирования, который необходимо соответствующим образом обрабатывать.
- После преобразований необходимо гарантировать:
 - ▶ синтаксическую корректность
 - ▶ корректность по отношению к изменениям "внешнего" кода
 - ★ Переименование объектов
 - ★ Удаление объектов
- Необходимо помнить что:
 - ▶ возможности "внешних" языков могут различаться
 - ▶ семантика языков может отличаться

Пример. Динамический SQL.

T-SQL:

```
IF @X = @Y
    SET @TABLE = '#table1'
ELSE SET @TABLE = 'table2'
SET @S = 'SELECT x FROM ' + @TABLE
        + ' WHERE ISNULL(n,0) > 1'
EXECUTE (@S)
```

PL-SQL:

```
IF lv_X = lv_Y
THEN lv_TABLE := 'tt_table1';
ELSE lv_TABLE := 'new_table2';
END IF;
lv_S := 'SELECT new_x FROM ' || lv_TABLE
        || ' WHERE NVL(n,0) > 1';
OPEN new_cursor FOR lv_S;
```

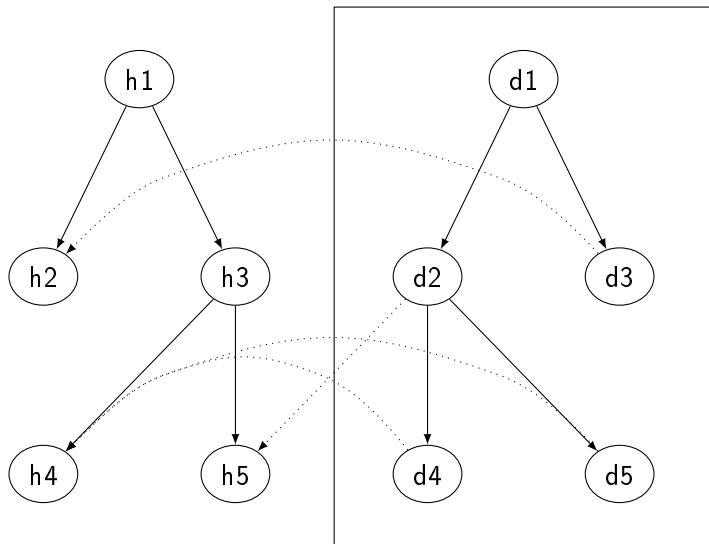
- Реинжиниринг информационных систем
 - ▶ Обработка встроенного SQL
 - ▶ Обработка динамического SQL в БД
- Разработка IDE с поддержкой рефакторинга встроенных языков.

Задача

- Разработать и реализовать библиотеку для трансляции динамически формируемых строк.
- Реализовать проверку результатов трансляции, минимизировав затраты на разработку.

- Лексический.
- Синтаксический.
- Основан на обработке компактного представления множества возможных значений формируемой строки:
 - ▶ регулярное выражение
 - ▶ граф
 - ▶ data-flow уравнение
- Используется для проверки синтаксической корректности динамически формируемых строк.
- Не подходит для трансформаций динамически формируемых строк, так как не сохраняет привязку к исходному коду.

Трансляция. Сохранение привязки.



Валидация результатов трансляции

- Необходимо проверять корректность трансляции статически.
- Необходимо минимизировать затраты на реализацию.

Валидация результатов трансляции

- Основная идея: если трансляция прошла успешно, то абстрактный парсер целевого языка должен получить столько деревьев, сколько получил абстрактный парсер исходного языка до трансляции.
- Практически всё можно переиспользовать.
 - ▶ Протягивание констант.
 - ▶ Абстрактный анализ.

- Реализована библиотека абстрактного анализа, расширенного для нужд трансляции.
- Проведена апробация на примере трансляции динамического SQL из T-SQL в PL-SQL.
- Реализована проверка корректности результатов трансляции.
- Продемонстрирована возможность переиспользования всех основных копонент для реализации проверки корректности результатов трансляции.

- F#
- FSharp PowerPack
- FsLex
- FsYacc
- YaccConstructor