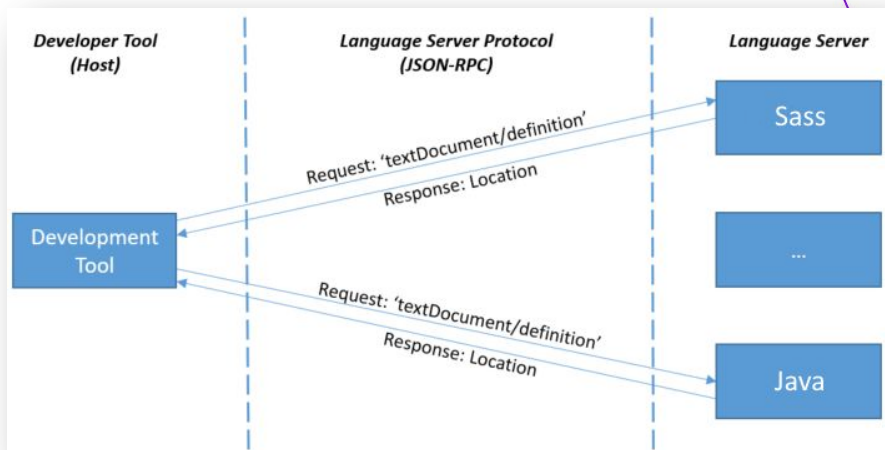


ІІТМО

Разработка метода межъязыкового семантического анализа текстов программ

Подготовил: Орловский М.Ю
Системное и прикладное программное обеспечение
Научный руководитель: Логинов И.П.

Анализ сценариев использования



Для анализа сценариев использования решено рассмотреть LSP.

LSP (language server protocol) – протокол, обобщающий языковые концепции для получения общих методов взаимодействия с программой в контексте инструментальных средств разработки

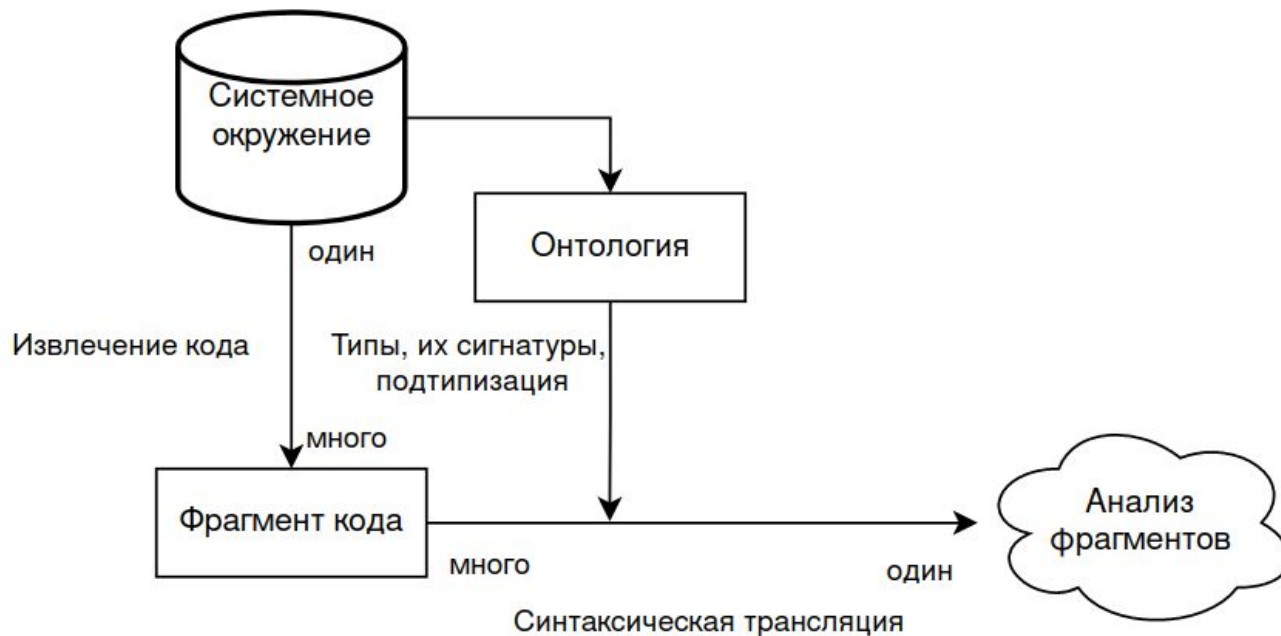
Зависимость функциональности от необходимой информации



Информация об областях видимости	Информация об областях видимости и типах
<ul style="list-style-type: none">• Подсветка символа• Сворачивание кода	<ul style="list-style-type: none">• Поиск объявления или определения• Определение иерархии вызовов• Поиск ссылок на идентификатор• Автодополнение символа

Специфично для языка: подсказки, рефакторинги, стиль кода

Неполная схема метода





Подзадачи:

- Извлечение информации о системе;
- Парсинг информации;
- Трансляция информации в обобщенную структуру.

Информация о системе

ІІТМО

```
include/  
├── metaLang99  
│   ├── assert.h  
│   ├── bool.h  
│   ├── choice.h  
│   ├── control.h  
│   ├── either.h  
│   └── eval  
│       ├── acc.h  
│       ├── eval.h  
│       ├── rec.h  
│       └── syntax_checker.h
```

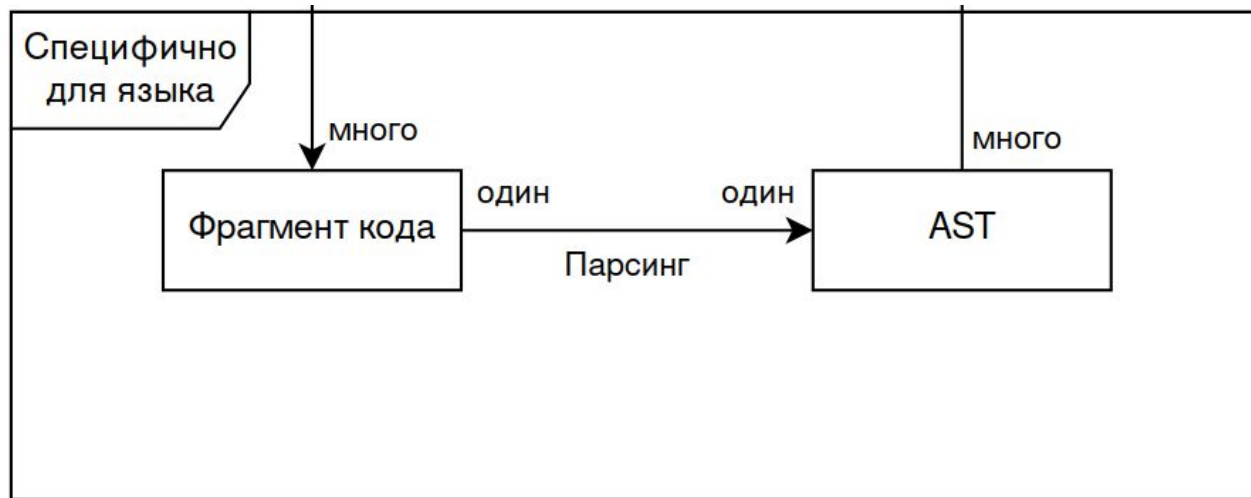
```
Apple ~ LOCAL_VAR="I am a Local Variable"  
Apple ~ export GLOBAL_VAR="I am a Global Variable"  
Apple ~ unset LOCAL_VAR  
Apple ~ unset GLOBAL_VAR  
Apple ~ echo $LOCAL_VAR  
Apple ~ echo $GLOBAL_VAR
```

Системной информацией может считаться:

- Содержимое файловой системы;
- Системные переменные;
- Ресурсы (сеть, процессор, конфигурация периферии).

То есть структурированные данные.

Процесс парсинга



Парсинг вовлекает генератор парсеров и языковую грамматику.
Формат AST – структурированное дерево (S-выражения)

Синтаксическая трансляция

	Production	Code Snippet
1	$Number \rightarrow Sign\ List$	$$$ \leftarrow \$1 \times \$2$
2	$Sign \rightarrow +$	$$$ \leftarrow 1$
3	$Sign \rightarrow -$	$$$ \leftarrow -1$
4	$List \rightarrow Bit$	$$$ \leftarrow \1
5	$List_0 \rightarrow List_1\ Bit$	$$$ \leftarrow 2 \times \$1 + \$2$
6	$Bit \rightarrow 0$	$$$ \leftarrow 0$
7	$Bit \rightarrow 1$	$$$ \leftarrow 1$

В общем случае функция, отображающая синтаксические элементы (AST) в исполняемый код для над этими элементами

Для совместного анализа информации о различных языках необходимо выбрать единое представление.

Таким представлением является онтология.



Краткое содержание онтологии:

- Семантика межъязыковых связей;
- Грамматики и соответствующие парсеры для разных языков;
- Синтаксические трансляции для разных языков;
- Система типов.

В качестве системы типов стоит избрать хорошо изученные формализмы из теории типов. Таковыми являются различные расширения лямбда-исчисления.

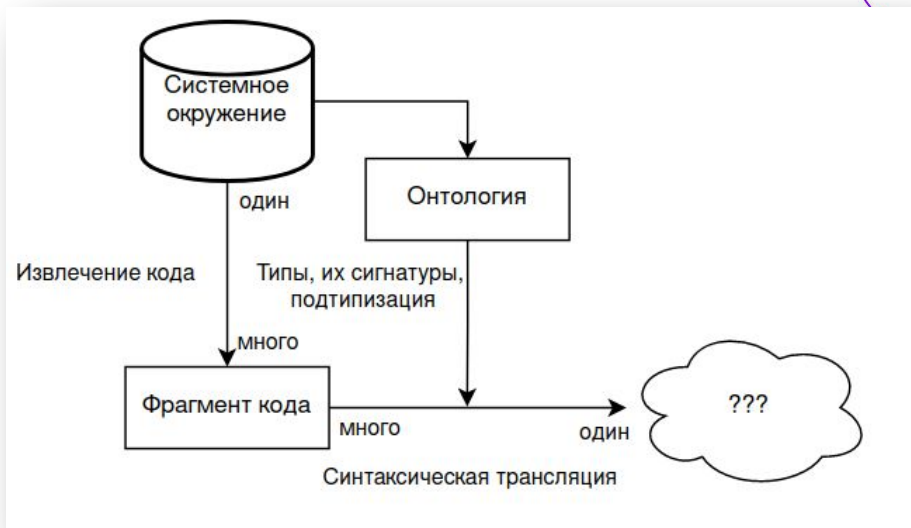
Виды систем:

- λ_{\rightarrow}
- $\lambda_{<}$
- System F
- System F_<

Практичная система: $\lambda_{<}$

Идеальная система: System F_<

О фрагментах



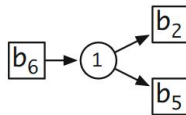
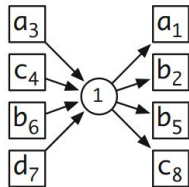
В ходе исследования было предпринято много попыток выработки формализма фрагментов.

Итоговым вариантом стал метод логического вывода, основанный на структурированных утверждениях.

Но всё изменилось...

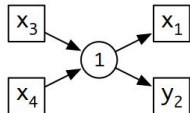
Графы областей (Scope-graphs)

```
def a1 = 0
def b2 = a3 + c4
def b5 = b6 + d7
def c8 = 0
```



```
def x1 = 1
def y2 = if x3 == 0 then 3 else x4
```

Scope graph



Declarations

$x_1^D : \tau_1$ $y_2^D : \tau_2$

Reference constraints

$x_3^R \mapsto \delta_1$ $\delta_1 : Int$
 $x_4^R \mapsto \delta_2$ $\delta_2 : \tau_4$

Type constraints

$\tau_1 \equiv Int$
 $Bool \equiv Bool$
 $Int \equiv Int$
 $\tau_3 \equiv Int$
 $\tau_2 \text{ is } \tau_3 \sqcup \tau_4$

Solution

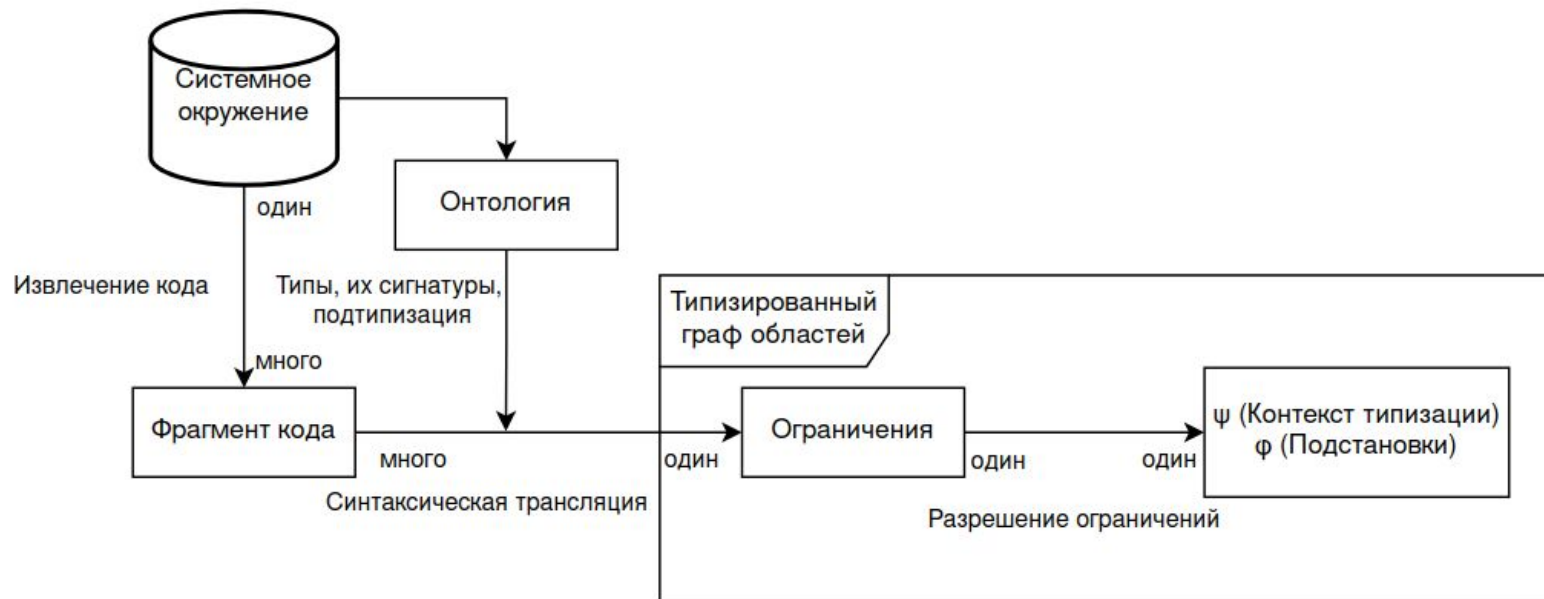
$\delta_1 = x_1^D$ $\delta_2 = x_1^D$
 $\tau_1 = Int$ $\tau_2 = Int$
 $\tau_3 = Int$ $\tau_4 = Int$

Статья “A Theory of Name Resolution [1]”, а также её продолжение

“Language-Independent Type-Dependent Name Resolution [2]”.

Языко-независимый формализм анализа областей видимости и ТИПОВ.

Итоговая схема анализатора



**Спасибо
за внимание!**

it'sMO *re than a*
UNIVERSITY

Github.com/uberde
ver
t.me: @uberdever