

# Прототип программного инструмента для анализа связности потока управления программ с открытым исходным текстом

Пустыгин А. Н.  
Зубов М.В.

Челябинск, 2017

# Общий алгоритм функционирования прототипа программного инструмента

---

- Преобразование исходного текста в универсальное промежуточное представления (для потока управления и диаграммы классов)
- Построение из универсального промежуточного представления эквивалентного представления в виде диаграммы потока управления на уровне методов и функций

# Другие подходы к проблеме

---

**XBLiteGen** – решает аналогичную проблему для BASIC.

## **Частичное решение проблемы:**

- Elsa, Bison, Flawfinder, GCC-XML
- scrML, JavaML, c2xml, C++2XML
- Dehydra, Treegydra
- Clang

# Эквивалентное представление потока управления программ Universal Control Flow Representation (UCFR)

предназначено для отображения конкретного участка выполнения программы — функции или метод класса, который назван управления функциональным блоком.

Является ориентированным графом общего вида, допускающий циклы.

Условные операторы представляются отдельными вершинами, в отличие от классических моделей, где операции перехода отображаются только в виде ребер графа.

Отображение такого графа будет больше соответствовать блок-схеме алгоритма, а так же позволит хранить и анализировать условия, по которым выбирается направление выполнения программы.

# Множество вершин графа эквивалентного представления

---

- project — исследуемый проект целиком;
- method — метод класса в проекте;
- function — функция в проекте (для процедурных языков);
- block — блок кода в потоке управления;
- flow — переход между блоками кода;
- tryExcept — блок обработки исключений;
- finally — блок, выполняющийся независимо от исключений;
- for — цикл со счетчиком;
- if — условный оператор;
- while — цикл с условием;
- call — вызов функции или метода.

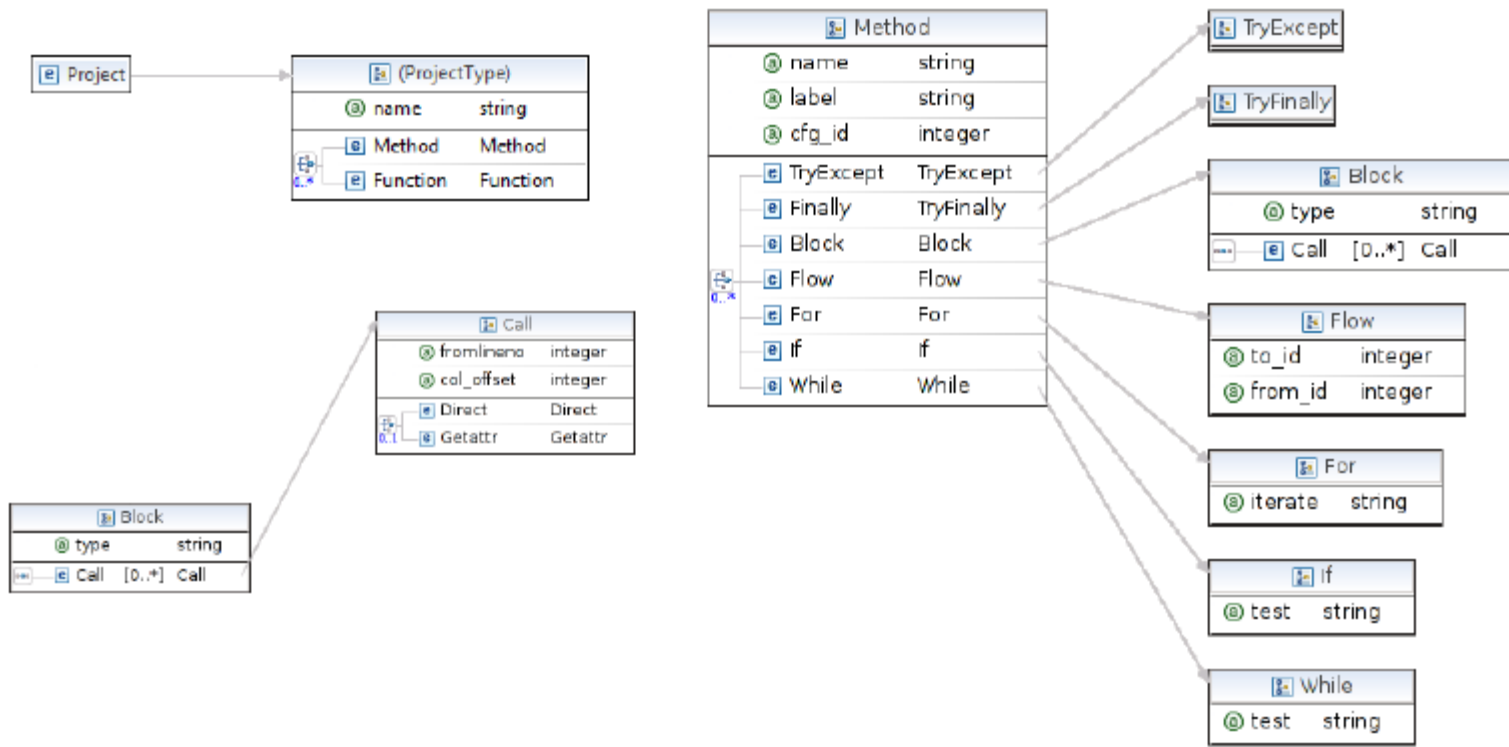
# Наличие узлов представления в избранных языках

Язык	project	method/ function	block/ flow	tryExcept	finally	for	if	while	call
Java	+	+ методы	+	+ try/catch	+	+ for/ foreach	+	+ while/ dowhile	+
Python	+	+ методы/ функции	+	+ try/except	+	+ foreach	+	+ while	+
C++	+	+ методы/ функции	+	+ try/catch	-	+ for	+	+ while/ dowhile	+
PHP	+	+ методы/ функции	+	+ try/catch	+ (с версии 5.5)	+ for/ foreach	+	+ while/ dowhile	+

# Машинный формат эквивалентного представления Universal Control Flow Representation (UCFR)

**Для описания такого представления был выбран текстовый язык разметки XML, использование которого позволяет сохранять эквивалентное представление в текстовый файл и проводить его анализ (возможно предварительный) вручную. Кроме того, любой XML-документ может быть описан XML-схемой с помощью специального языка [1]. Использование XML и их схем для описания форматов передачи данных является рекомендацией Консорциума Всемирной Паутины (W3C)**

# XML -схема формата файла эквивалентного представления *Universal Control Flow Representation (UCFR)*





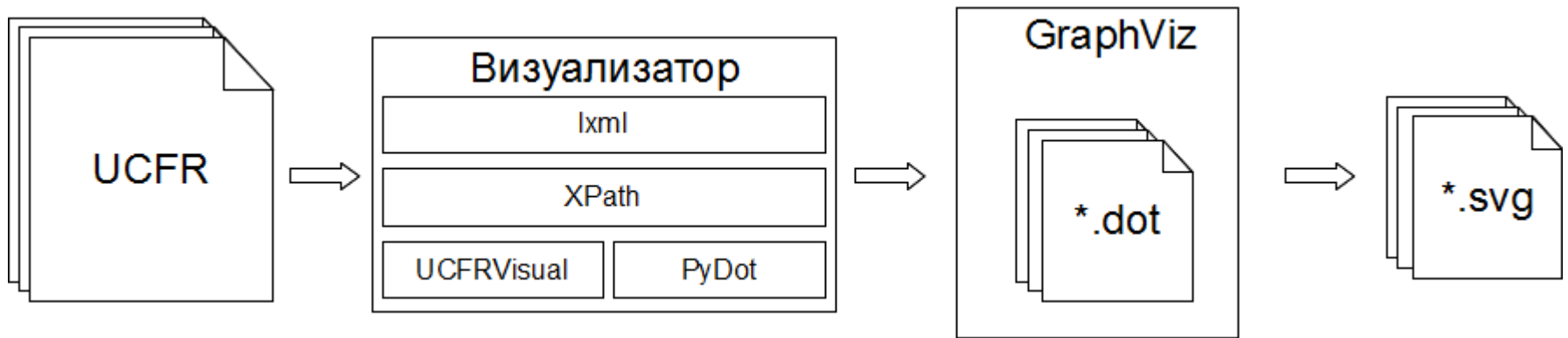
# Выбранные условия разработки

---

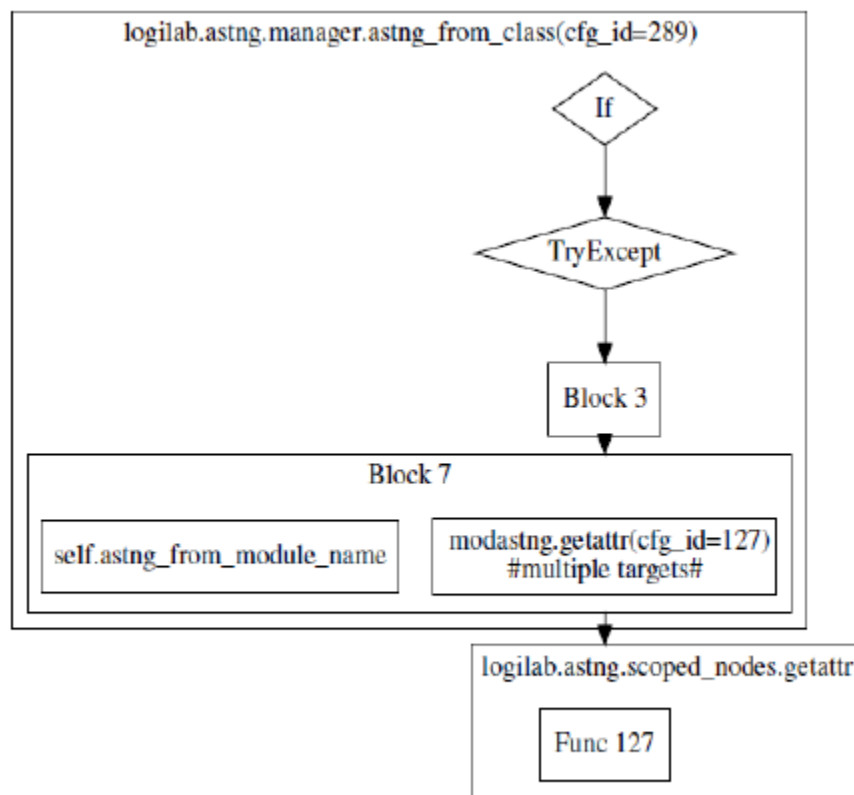
Для анализа и визуализации графа потока управления были разработаны две утилиты на языке Python. Для обработки XML используется распространенная библиотека lxml [2].

Для построения изображений используется утилита dot пакета Graphviz [3]. Библиотека PyDot [4] позволяет работать с текстовым форматом dot. В результате визуализации получается набор изображений в векторном формате svg.

# Схема функционирования системы утилит для визуализации эквивалентного представления *UCFR*

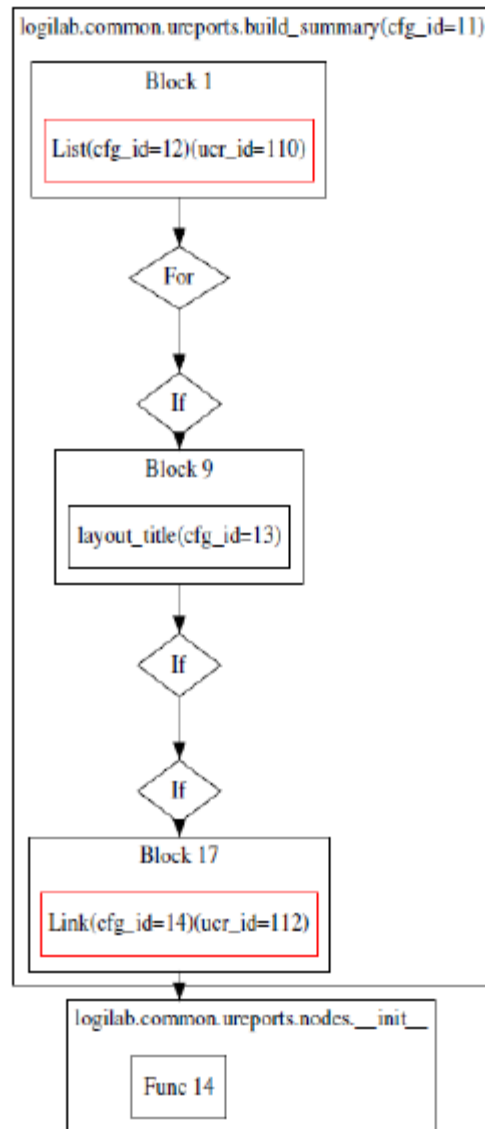


# Визуализация трассы в узле программного модуля logilab.astng.manager (Python)



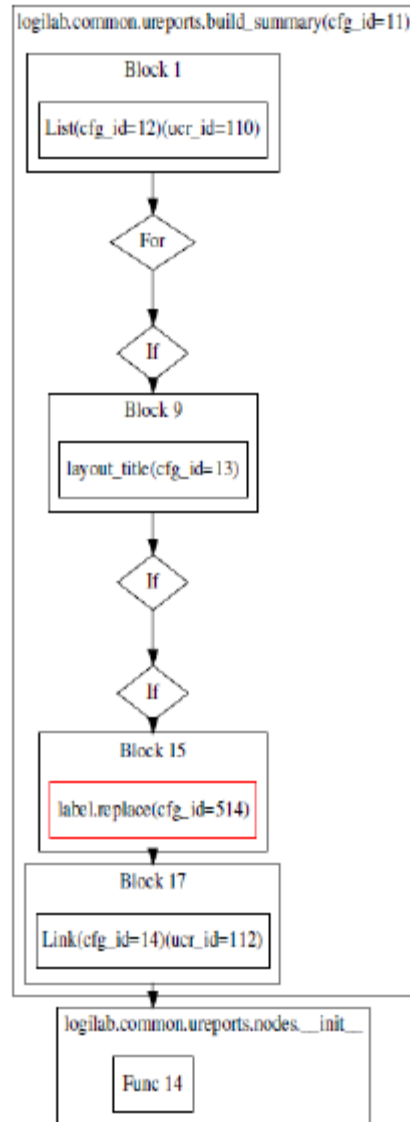
# Поиск создаваемых вдоль трассы объектов файла `logilab.common.ureports`

Обнаружено  
создание  
объектов  
классов `List` и  
`Link`



# Поиск вызовов методов, возникающих вдоль трассы в файле logilab.common.ureports

Обнаружен  
вызов метода  
replace для  
объекта label,  
находящийся на  
заданной трассе



# Результаты опытной разработки

- Опробован формат эквивалентного представления исходных текстов программ для группы языков, предназначенный для анализа связности потока управления
- формат эквивалентного представления исходных текстов программ для группы языков использован для анализа связности путем визуализации средствами ПО с открытым исходным текстом

# Список литературы

---

1. XML Schema [Электронный ресурс]. – Режим доступа: [www.w3.org/XML/Schema](http://www.w3.org/XML/Schema), свободный (дата обращения 14.01.2017).
2. lxml - XML and HTML with Python [Электронный ресурс]. – Режим доступа: <http://lxml.de/>, свободный (дата обращения 14.01.2017).
3. Spinellis D. Drawing Tools / D. Spinellis // IEEE Software. – 2009 . – №26. – CA, USA: IEEE Computer Society Press. – С. 12-13.
4. PyDot, a Python interface to Graphviz's Dot language [Электронный ресурс]. – Режим доступа: <https://code.google.com/p/pydot/>, свободный (дата обращения 14.07.2015)