

# Вывод типов для языка R в интегрированной среде разработки IntelliJ IDEA

Бакрадзе Лиана Георгиевна  
группа 544

Научный руководитель: к.ф.-м.н. Булычев Д.Ю.

Рецензент: Тузова Е.А., разработчик JetBrains

17 июня 2015 г.

# Введение

- ▶ Язык R — это язык с динамической типизацией, популярный в сфере обработки и анализа данных
- ▶ Разработка системы типов для динамического языка — сложная задача
- ▶ Не было предложено системы типов для языка R
- ▶ В JetBrains разрабатывается плагин для языка R для IntelliJ IDEA

## Цель

С помощью введения системы типов улучшить поддержку для языка R в плагине для IntelliJ IDEA

# Задачи

- ▶ Предложить систему типов для языка R
- ▶ Реализовать подсистему, производящую анализ исходного кода без его выполнения с целью получения информации о типах
- ▶ Реализовать встроенный статический анализатор, выявляющий несоответствие реальных типов ожидаемым
- ▶ Протестировать статический анализатор на реальных проектах на языке R

## Тип функций

```
x ← vector("integr", 3)
y ← vector("character", 10)
```

*function* : (*Params*, *paramTypes*, *return*, *rules*)

*paramTypes* : *Params* → *Types*

*return* ∈ *Types*

*rules* ⊆ (*type*, *value*, *return*)

## Система типов для языка R

- ▶ Гомогенные коллекции
- ▶ Гетерогенные коллекции
- ▶ S3 и S4 классы
- ▶ Функции
- ▶ Тип-объединение (*union type*)
- ▶ Универсальный тип (*top*)
- ▶ Пустой тип (*bottom*)

## Вывод типов

- ▶ Генерация модулей-заглушек (*skeletons, type definitions*)
- ▶ Предложены аннотации типов для языка R
- ▶ Сгенерированы аннотации для 45% функций библиотеки *base*

```
vector <- function (mode = "logical", length = 0L)  
.Internal(vector(mode, length))
```

## Вывод типов

```
## @type length : numeric
## @type mode : character
## @rule (mode="logical")->logical
## @rule (mode="numeric")->numeric
## @rule (mode="integer")->integer
## @rule (mode="double")->numeric
## @rule (mode="complex")->complex
## @rule (mode="character")->character
## @rule (mode="raw")->raw
## @rule (mode=string) -> error(Wrong mode)
vector <- function (mode = "logical", length = 0L)
.Internal(vector(mode, length))
```



## Вывод типов

- ▶ Анализ потока управления
- ▶ Кэширование
- ▶ Встроенный статический анализатор — инспекция

# Тестирование

Название проекта	Количество строк кода	Ошибки	Ложные срабатывания
htr	4375	0	1
readr	1440	0	0
reconstructr	615	0	0
rprime	1137	3	0
selfie	673	0	0
signals	5487	2	1

## Результаты

1. Предложена система типов для языка R
2. Реализована подсистема, производящая анализ исходного кода без его выполнения с целью получения информации о типах
3. Реализован встроенный статический анализатор, выявляющий несоответствие реальных типов ожидаемым
4. Статический анализатор протестирован на реальных проектах на языке R
5. Часть реализованной функциональности переиспользована в других компонентах плагина
6. Реализованная функциональность войдёт в первую версию продукта