

# Улучшенная оптимизация совместного сопоставления с образцом в компиляторе Рефал-5λ

Павел Савельев

МГТУ им. Н. Э. Баумана

Совместное рабочее совещание  
ИПС имени А. К. Айламазяна РАН

и

МГТУ имени Н. Э. Баумана

5 июня 2018 г.

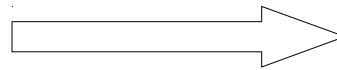
# Вступление

- Предложения в одной и той же функции зачастую имеют схожую структуру
- Функции языка не могут принимать более одного аргумента
- Из аргументов формируется единый объект, который разбивается на части уже непосредственно в теле функции

# Актуальность проблемы

- **Форматы функций**  
формально - 1 параметр, фактически - несколько
- **Неявные параметры для вложенных функций**

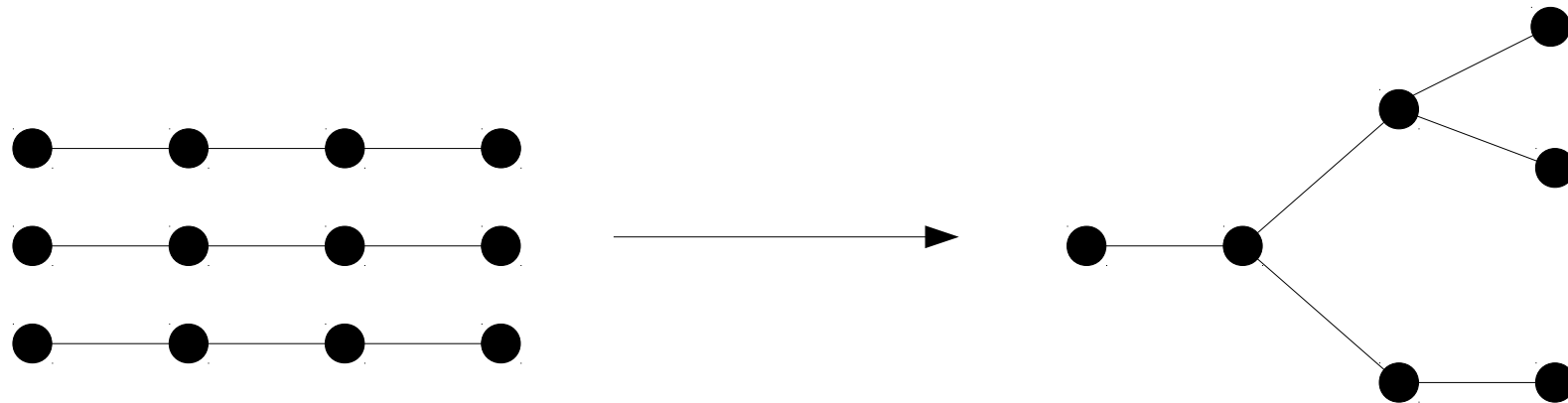
```
{  
  1 = e.X;  
  2 = e.Y;  
}
```



```
{  
  (e.X) (e.Y) 1 = e.X;  
  (e.X) (e.Y) 2 = e.Y;  
}
```

# Подходы

- Построение дерева
  - Строим последовательность команд сопоставления для каждого из предложений
  - Выделяем общие префиксы



# Подходы

- Определение образца общего вида, анализируя внешний вид отдельных образцов

# Подходы

## а) построение ЛСО

- обобщение происходит терм за термом на каждом скобочном уровне
- выбираем самые левые и самые правые термы
- обобщаем те из них, обобщение которых приводит к наименьшей потере информации

# Подходы

b) Построение ГСО

Вычисляем обобщение для всех образцов

c) Построение ГСО + группировка

# Теория

- Жесткие выражения
- Уточнение и обобщение
- Сложность выражения

$$C(P) = n_t + 2 \times n_s + 3 \times n_x + 3 \times n_{()} - n_e + 1$$

- Минимальное уточнение

$$e \rightarrow \varepsilon$$

$$e \rightarrow te$$

$$e \rightarrow et$$

$$t \rightarrow (e)$$

$$t \rightarrow s$$

$$s \rightarrow X$$

# Теория

- ЛСО — локальное сложнейшее обобщение
- ГСО — глобальное сложнейшее обобщение

Исходные образцы	ЛСО	Сложность	ГСО
s (e) A t s	e t t t s	5	
s (e) t s	s e t t s	6	
	s (e) e t s	7	s (e) e t s
	s (e) t e s	7	s (e) t e s
	s (e) t t e	6	

$$C(P) = n_t + 2 \times n_s + 3 \times n_x + 3 \times n_{()} - n_e + 1$$

# Алгоритм построения ГСО - улучшенный

- Основные отличия от исходного алгоритма
  - Отказываемся от хранения подстановок
  - Отказываемся от промежуточного быстрого обобщения
  - Добавляем промежуточное представление образцов, обладающее ассоциативностью

# Алгоритм построения ГСО - улучшенный

- Этапы
  - Построение изображений
  - Обобщение
  - Коллапс

# Алгоритм построения ГСО - улучшенный

- Построение изображений
  - Скобочный терм  $(t_1 \dots t_K)$  класса  $K$  изображается как  $(K : t'_1 \dots t'_K)$ , где  $t'_i$  — изображение соответствующего терма
  - Скобочный терм  $(t_{L1} \dots t_{LM} e . X t_{RN} \dots t_{R1})$  изображается как  $(MN : (M+N), (t'_{L1} \dots t'_{LM}), (t'_{RN} \dots t'_{R1}))$
  - Символы, s- и t- переменные изображаются собой. Индексы отбрасываются

# Алгоритм построения ГСО - улучшенный

- Обобщение

- Термы  $(K : terms_1)$  и  $(K : terms_2)$  равной длины обобщаются в  $(K : terms')$
- Термы  $(MN : K_1, l_1, r_1)$  и  $(MN : K_2, l_2, r_2)$  обобщаются в  $(MN : \min(K_1, K_2), l', r')$ ;  $l' : (l_1 r_1) u (l_2 r_2)$  обобщаются слева, а  $r'$  соответственно справа
- Термы  $(MN : K_1, l_1, r_1)$  и  $(K : terms_1)$  обобщаются как  $(MN : K_1, l_1, r_1)$  и  $(MN : |terms_1|, terms_1, terms_1)$

# Алгоритм построения ГСО - улучшенный

- Коллапс
  - Терм  $(K:t_1\dots t_K)$  коллапсирует в  $(t_1\dots t_K)$
  - Терм  $(MN:K, (t'_{L1}\dots t'_{LM}), (t'_{RN}\dots t'_{R1}))$  соответствует семейству ЛСО. После коллапса вложенных термов строим образцы вида  $(t_{L1}\dots t_{LM} e.X t_{RN}\dots t_{R1})$  и среди этих образцов выбираем обладающий наибольшей сложностью
- Результат — ГСО для заданного набора жестких выражений

# Пример

## 1. Исходные образцы

s (e) A t s  
s (e) t s

(K: 5, s (e) A t s)  
(K: 4, s (e) t s)

## 2. Построение изображений

Примечание:  
(e) → (MN: 0, ( ), ( ))

## 3. Обобщение

слева:  
s (e) A t s  
s (e) t s  
получим  
s (e) t t

(MN: 4,	s (e) t t,	t t t s)	
0+4:		e t t t s,	СЛОЖН. 5
1+3:	s	e t t s,	СЛОЖН. 6
2+2:	s (e)	e t s,	СЛОЖН. <b>7</b>
3+1:	s (e) t	e s,	СЛОЖН. <b>7</b>
4+0:	s (e) t t	e,	СЛОЖН. 6

справа:  
s (e) A t s  
s (e) t s  
получим  
t t t s

Получим ГСО: {s(e)e t s, s(e)t e s}

# Построение ГСО + группировка

- Последовательно обобщаем подстановки до  $e$ - и  $t$ - переменных
- Формируем группы
  - Недостаточное различие в подстановках — получаем всего одну группу
  - Для каждой группы запоминаем количество образцов
- Для каждой группы вычисляем ГСО

# ИТОГОВЫЙ ВИД функции до правок

Функция:

```
Общие команды всех предложений  
Защищенный блок {  
    Команды первого предложения  
}
```

...

```
Защищенный блок {  
    Команды k-го предложения  
}
```

...

```
Команды N-го предложения
```

# ИТОГОВЫЙ вид функции после правок

Функция:

Общие команды всех предложений

Защищенный блок {

    Общие команды первой группы

    Защищенный блок {

        Команды первого предложения

    }

    ...

    Команды k-го предложения

  }

Защищенный блок {

    Общие команды второй группы

    ...

  }

  ...

Команды N-го предложения

Спасибо за внимание

```
1 $ENTRY F {
2   Abc   ((t.1) e.2 s.3) (t.4 e.5) = 1;
3   (e.X) ((t.1) e.2 s.3)           = 2;
4 }
5
6 $ENTRY F {
7   Abc   ((t.1) e.2 s.3) ( t.4 e.   ) = 1;
8         ( e.X           ) ((t.1) e.2 s.3) = 2;
9 }
```