

Разработка проблемно-специфичного языка для управления физическим уровнем баз данных

Выполнила: Ольга Силина, 544 группа
Научный руководитель: к.ф.-м.н., доц. Д.В. Барашев
Рецензент: ст. преп. Д.В. Луцив

Санкт-Петербургский Государственный Университет

2010

Цели работы

- ▶ Решение проблемы хранения сложных данных
 - ▶ Отображение в реляционную или объектную модель и использование соответствующей СУБД не всегда работает
 - ▶ Не всегда можно сделать эффективное отображение
 - ▶ Иногда хочется вручную контролировать физическое представление данных
- ▶ Прототипирование новых структур хранения данных

Постановка задачи

- ▶ Создание проблемно-ориентированного языка для управления физическим уровнем баз данных
- ▶ Язык должен быть простым, но выразительным
- ▶ Разработка структуры и синтаксиса языка, создание синтаксического анализатора и транслятора в Java

Абстракции, реализованные в языке:

- ▶ поблочная буферизованная итерация
- ▶ управление буферами в оперативной памяти
- ▶ буферизованная запись в файл
- ▶ сравнение структурированных данных
- ▶ работа с файлами и временными файлами

Программная реализация

- ▶ Haskell
- ▶ генератор синтаксических анализаторов Happy
- ▶ генератор лексических анализаторов Alex
- ▶ транслятор в язык Java

Фрагмент кода

```
list tempIterators
for (buffer :: inputFile by Data(BUF_SIZE)) {
  sort(buffer, Data, dataComp)
  file tempfile in temp
  auto tempfile <- buffer
  tempIterators += iterator of Data for tempfile
}
writing Data to outputFile {
  do {
    min = findMinAndPull(tempIterators, dataComp)
    if (min != null) {
      auto outputFile <- min
    }
  } while (true)
}
```

- ▶ Разработана структура и синтаксис языка
- ▶ Разработана реализация лексического и синтаксического анализаторов
- ▶ Разработан транслятор в язык Java
- ▶ В качестве примера реализован алгоритм многоканальной сортировки слиянием и простая индексная структура