

Дисциплина «Проектирование баз данных»



Маркова Ирина Васильевна,
начальник управления
информатизации
markova@mit.ru



Синтаксический анализ и деревья разбора

Задача синтаксического анализа – преобразовании текста, написанного на языке, подобном SQL, в дерево разбора. Каждая вершина дерева разбора представляет собой одну из следующих сущностей :

- ✓ атом – лексические единицы (служебные слова, наименования атрибутов, константы и т. д.);
- ✓ синтаксическая категории – символьные обозначения сегментов частей запроса выполняющих одинаковые функции.

Если вершина дерева разбора является атомом, она не может иметь подвершин. Если же вершина представляет синтаксическую категорию, дочерние вершины описываются посредством одного из правил грамматики языка.



Грамматика простого подмножества языка SQL (основные категории)

Запросы

```
<Query> ::= <SFW>  
<Query> ::= (<QUERY> и др.)
```

Форма «Select-From-Where»

```
<SFW> ::= SELECT < Sellist > From <Fromlist> Where <Condition>
```

Список Select

```
<Sellist> ::= <Attribute>  
<Sellist> ::= <Attribute>, <Sellist>
```



Грамматика простого подмножества языка SQL (основные категории)

Условия

```
<Condition> ::= <Condition> and <Condition>;  
<Condition> ::= <Tuple> in <Query>;  
<Condition> ::= <Attribute> = <Attribute>;  
<Condition> ::= <Attribute> like <Pattern> и др.
```

Список From

```
<Fromlist> ::= <Relation>;  
<Fromlist> ::= <Relation>, <Fromlist>.
```

Базовые синтаксические категории

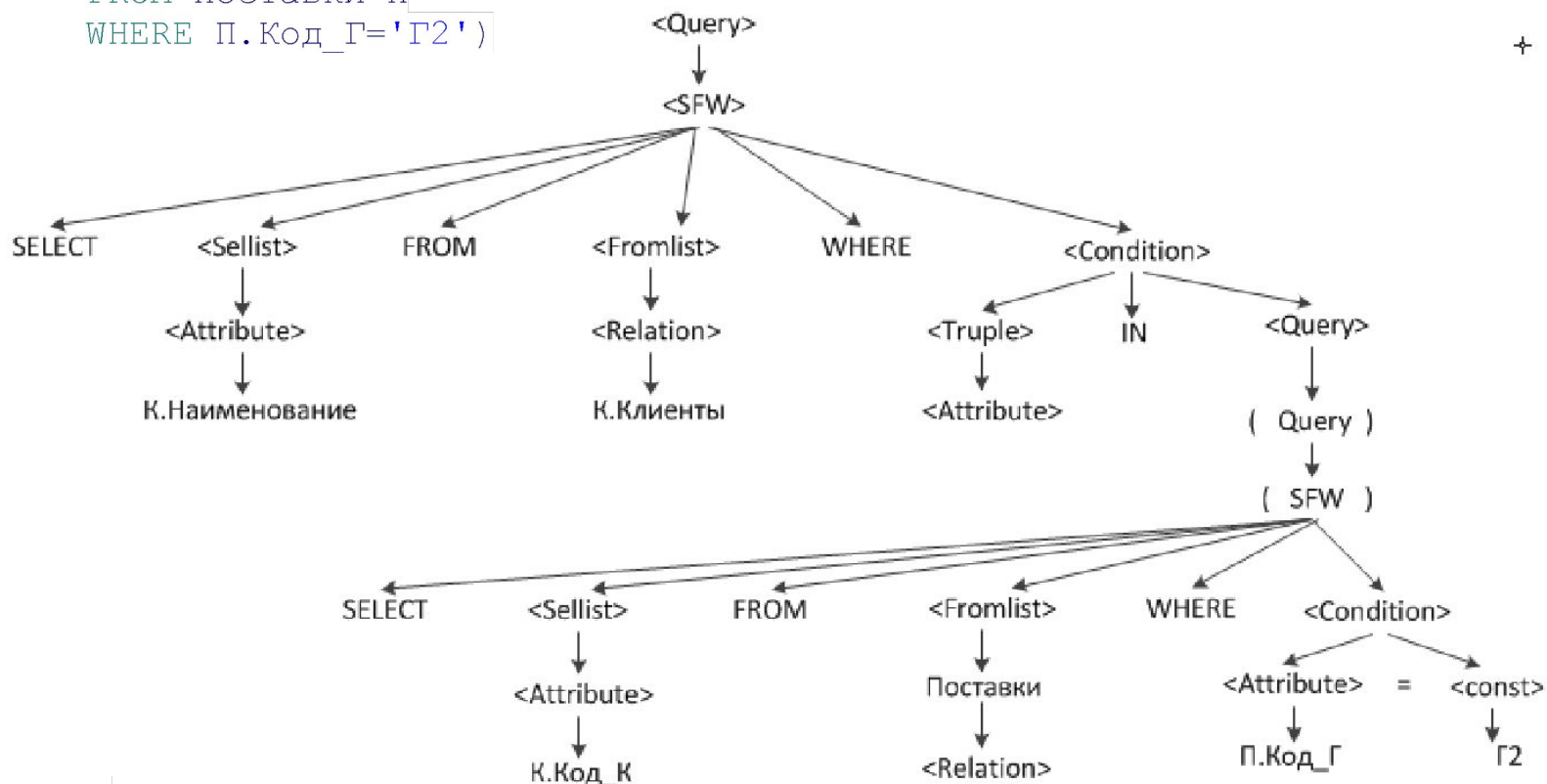
```
<Attribute>;  
<Relation>;  
<Pattern>
```

Базовые синтаксические категории имеют специальные значения и отличаются от остальных тем, что определяются не грамматическими правилами, а правилами для соответствия атомов.



Примеры

```
SELECT К.Наименование
FROM Клиенты К
WHERE К.Код_К IN
    (SELECT П.Код_К
     FROM Поставки П
     WHERE П.Код_Г='Г2')
```



```
SELECT К.Наименование
FROM Клиенты К, Поставки П
WHERE К.Код_К = П.Код_К AND П.Код_Г='Г2'
```



Функции препроцессора

Препроцессор – специализированный компонент компилятора запросов выполняющий следующие операции:

- ✓ замена ссылок на дерево разбора текстовой конструкции описывающей эту ссылку (виртуальная таблица - дерево)
- ✓ проведение синтаксического разбора
- ✓ проведение семантического контроля
- ✓ контроль употребления имен отношений
- ✓ контроль употребления использования имен атрибутов и их разрешения
- ✓ контроль употребления типов

Дерево разбора прошедшее все стадии разбора называется действительным (valid).

Дерево, модифицированное за счет возможной вставки поддеревьев виртуальных таблиц и разрешения имен, поддается генерации логического плана запросов.