

Лекция 10

# Грамматика Русского языка

# Использование Пролога в синтаксическом разборе

Порождающей грамматикой называется следующая четверка:

$$G = \langle VT, VN, S, P \rangle$$

где

**$VT$ ,  $VN$**  - соответственно, **терминальный** и **нетерминальный** словари,

**$S$**  - начальный символ,

**$P = \{\alpha i \rightarrow \beta i\}$**  - множество правил вывода,

*причем*

$\alpha i$  - цепочка, *содержащая нетерминальный символ*

$\beta i$  - произвольная цепочка из терминальных и нетерминальных символов.

Непосредственным порождением называется отношение:

$$\lambda \Rightarrow \mu,$$

где  $\lambda = \delta_1 \alpha_i \delta_2$

$$\mu = \delta_1 \beta_i \delta_2$$

и существует правило  $\alpha_i \rightarrow \beta_i$ .

Порождением называется отношение:

$$\gamma_1 \Rightarrow^* \gamma_n, \quad n = 1, 2, \dots,$$

если существует последовательность отношений:

$$\gamma_1 \Rightarrow \gamma_2 \Rightarrow \dots \Rightarrow \gamma_n.$$

**Языком**, порождаемым грамматикой  $G$ , называется:

$$L(G) = \{\gamma \mid S \Rightarrow^* \gamma\},$$

где  $\gamma$  - терминальная цепочка

(цепочка, состоящая только из элементов терминального словаря  $V_T$ ).

# Парсер предложений

*Большой кот вскочил в переполненный московский трамвай*

$VT = \{\text{слова из словаря: кот, пес, большой, в, на...}\}$

$VN = \{\text{ФР, Г\_С, Г\_Г, СУЩ, ГЛ, ПР, ПД}\}$

$S = \text{ФР}$

Р:

- $\text{ФР} \rightarrow \text{Г\_С Г\_Г}$
- $\text{Г\_С} \rightarrow \text{СУЩ} \mid \text{ПР Г\_С}$
- $\text{Г\_Г} \rightarrow \text{ГЛ} \mid \text{ГЛ ПД Г\_С}$
- $\text{СУЩ} \rightarrow \text{кот} \mid \text{пес} \mid \text{трамвай} \mid \text{тротуар} \mid \dots$
- $\text{ПР} \rightarrow \text{большой} \mid \text{огромный} \mid \text{черный} \mid \text{рыжий} \mid \text{лохматый} \mid \text{московский} \mid \text{переполненный} \mid \dots$
- $\text{ГЛ} \rightarrow \text{вскочил} \mid \text{выскочил} \mid \text{прогуливался} \mid \dots$
- $\text{ПД} \rightarrow \text{в} \mid \text{на} \mid \dots$

# Грамматика предиката

## Грамматика языка

- $\Phi P \rightarrow \Gamma_C \Gamma_\Gamma$
- $\Gamma_C \rightarrow \text{СУЩ} \mid \text{ПР } \Gamma_C$
- $\Gamma_\Gamma \rightarrow \text{ГЛ} \mid \text{ГЛ ПД } \Gamma_C$

```
an_ph(L, ph(X, Y)):-  
    append(L1, L2, L),  
    an_gs(L1, X),  
    an_gg(L2, Y).
```

# Предикат синтаксического разбора

```
test1(['большой','черный','кот','вскочил','в','переполненный','московский','трамвай']).  
test2(['кот','прогуливался']).  
test3(['большой','рыжий','лохматый','пес','выскочил','на','тротуар']).
```

## Словари:

```
s_list(['кот','пес','трамвай','тротуар']).  
p_list(['большой','лохматый','московский',  
        'переполненный','рыжий','черный']).  
g_list(['вскочил','выскочил','прогуливался']).  
pd_list(['в','на']).
```

```
an_ph(L,ph(X, Y)) :-  
    append(L1, L2, L),  
    an_gs(L1, X),  
    an_gg(L2, Y).
```

```
an_gs([X],gs(X)) :- an_s([X],s(X)).  
an_gs(S,gs(X,Y)) :- append(S1,S2,S),  
    an_p(S1,X),  
    an_gs(S2,Y).
```

```
an_gg([X],gg(X)) :- an_g([X],g(X)).  
an_gg(S,gg(X,Y,Z)) :- append(S1,S2,S), append(S4,S5,S2),  
    an_g(S1,X),  
    an_pd(S4,Y),  
    an_gs(S5,Z).
```

```
an_s([X],s(X)) :- s_list(L), member(X,L).  
an_p([X],p(X)) :- p_list(L), member(X,L).  
an_g([X],g(X)) :- g_list(L), member(X,L).  
an_pd([X],pd(X)) :- pd_list(L), member(X,L).
```

# Запросы

?-test1(L), an\_ph(L, Res).

?-test2(L), an\_ph(L, Res).

?-test3(L), an\_ph(L, Res).

Res = ph(gs(p('большой'), gs(p('черный'), gs(s('кот')))),  
gg(g('вскочил'), pd('в'), gs(p('переполненный'),  
gs(p('московский'), gs(s('трамвай')))))).

Res = ph(gs(s('кот')), gg(g('прогуливался'))).

Res = ph(gs(p('большой'), gs(p('рыжий'), gs(p('лохматый'),  
gs(s('пес'))))), gg(g('выскочил'), pd('на'), gs(s('тротуар')))).

?-test1(L), an\_ph(L, Res), an\_ph(L1, Res).

?-test2(L), an\_ph(L, Res), an\_ph(L1, Res).

?-test3(L), an\_ph(L, Res), an\_ph(L1, Res).



# Использование словарей

Использование словарей продемонстрируем на примере задачи преобразования **поверхностных** структур ограниченного естественного языка в **глубинные** структуры.

**Поверхностным** структурам русского языка:

*"Мой сосед купил видеокамеру"*

*"Видеокамера куплена моим соседом»"*

соответствует глубинная структура:

```
"predicate(action('купить'),  
            agent('мой сосед'),  
            object('видеокамера'))"
```

# Задача

*['он', 'любит', 'ее']* или *['она', 'любима', 'им']*  
должны быть преобразованы в одну и ту же глубинную структуру:  
**'любить'('он', 'она')**.

Каждая запись файла-словаря будет представлять собой  
структуру Пролога вида:

*<Каноническая форма>: <Контекст>: <Список словоформ>.*

*Например:*

**'любить':zalog('пассив'):['любим','любима','любимо','любимы']**.

# Предикат

```
:-op(200, xfy, ':').
```

```
%
```

*Поиск в словаре*

```
fid(X, XC, K, File):-
```

```
member(M,File), condition(X, XC, K, M).
```

```
condition(X, XC, K, XC:K:L):- member(X, L).
```

```
an_phrase(Y1, zalog('актив'),[X,Y,Z]) :-
```

```
    an_mest(X1, pad('именит'), X),
```

```
    an_glag(X2, zalog('актив'),Y),
```

```
    an_mest(X3, pad('винит'),Z),
```

```
    Y1=..[X2, X1, X3].
```

```
an_phrase(Y1, zalog('пассив'),[X,Y,Z]) :-
```

```
    an_mest(X1, pad('именит'),X),
```

```
    an_glag(X2, zalog('пассив'),Y),
```

```
    an_mest(X3, pad('творит'),Z),
```

```
    Y1=..[X2, X3, X1].
```

```
an_mest(XC, K,X) :- gen1(File1),fid(X, XC, K, File1).
```

```
an_glag(XC, K, Y) :- gen2(File2),fid(Y, XC, K, File2).
```

```
test1(['ты','любишь','ее']).
```

```
test2(['она','любима','вами']).
```

```
test3(['оно','куплено','им']).
```

# Словари

```
%-----  
%                               Словарь местоимений  
%-----
```

```
gen(File1):-File1=[  
'я':pad('именит'):['я'],  
'я':pad('винит'):['меня'],  
'я':pad('творит'):['мной', 'мною'],  
'ты':pad('именит'):['ты'],  
'ты':pad('винит'):['тебя'],  
'ты':pad('творит'):['тобой', 'тобою'],  
'он':pad('именит'):['он'],  
'он':pad('винит'):['его'],  
'он':pad('творит'):['им'],  
'она':pad('именит'):['она'],  
'она':pad('винит'):['ее'],  
'она':pad('творит'):['ей', 'ею'],  
'оно':pad('именит'):['оно'],  
'оно':pad('винит'):['его'],  
'оно':pad('творит'):['им'],  
'мы':pad('именит'):['мы'],  
'мы':pad('винит'):['нас'],  
'мы':pad('творит'):['нами'],  
'вы':pad('именит'):['вы'],  
'вы':pad('винит'):['вас'],  
'вы':pad('творит'):['вами'],  
'они':pad('именит'):['они'],  
'они':pad('винит'):['их'],  
'они':pad('творит'):['ими'] ]
```

%-----

%  
Словарь глаголов

%-----

gen2(File2):-File2=

['любить':zalog('актив'): ['люблю','любишь','любит','любим','любите','любят'],

'любить':zalog('пассив'): ['любим','любима','любимо','любимы'],

'купить':zalog('актив'): ['купил','купила','купило','купили'].

'купить':zalog('пассив'): ['куплен','куплена','куплено','куплены'] ]

%-----

?-test1(L), an\_phrase(Res, K, L).

?-test2(L), an\_phrase(Res, K, L).

?-test3(L), an\_phrase(Res, K, L).

Res = 'любить'('ты', 'она') K = zalog('актив').

Res = 'любить'('вы', 'она') K = zalog('пассив').

Res = 'купить'('он', 'оно') K = zalog('пассив').