Часть 4. Автоматический синтаксический анализ

Алгоритмы реализации синтаксического анализа в терминах контекстно-свободной грамматики (КСГ)

Алгоритмы реализации анализа в терминах КСГ (Context-Free Grammar, CFG)

ПЛАН

- Алгоритмы:
 - Метод рекурсивного спуска (top-down parsing)
 - Восходящий анализ (bottom-up parsing)
 - Алгоритм Кока-Янгера-Касами (CKY Parsing)
 - Алгоритм Эрли (Earley parser)
- Методы улучшения качества:
 - Вероятностные контекстно-свободные грамматики (PCFG)
 - Лексикализованные контекстно-свободные грамматики
- "Борьба" с разрывными составляющими

Алгоритмы реализации анализа в терминах КСГ (Context-Free Grammar, CFG)

- Алгоритмы:
 - Метод рекурсивного спуска (top-down parsing)
 - Восходящий анализ (bottom-up parsing)
 - Алгоритм Кока-Янгера-Касами (CKY Parsing)
 - Алгоритм Эрли (Earley parser)

Алгоритмы реализации анализа в терминах КСГ (Context-Free Grammar, CFG)

- Что значит провести синтаксический анализ в терминах НС?
 - ✓ Поставить в соответствие предложению (цепочке терминальных символов) систему вложенных составляющих (расставить скобки)

Дано:

КС-грамматика, предложение

Найти:

множество правил из КС-грамматики, которые это предложение порождают

Алгоритмы синтаксического анализа в терминах CFG

- Нам дано 2 набора данных:
 - Цепочка терминальных символов
 - Начальный символ

Пример

 $S \rightarrow NP VP$

 $S \rightarrow Aux NP VP$

 $S \rightarrow VP$

NP → Pronoun

NP → Proper-Noun

NP → Det Nominal

Nominal → Noun

Nominal → Nominal Noun

Nominal → Nominal PP

 $VP \rightarrow Verb$

VP → Verb NP

VP → Verb NP PP

VP → Verb PP

 $VP \rightarrow VP PP$

PP → Preposition NP

Det → that I this I a

Noun → book I flight I meal I money

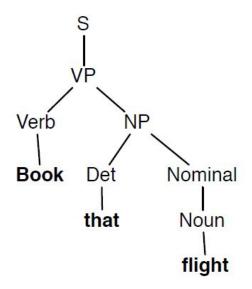
Verb → book I include I prefer

Pronoun → II she I me

Proper-Noun → Houston I TWA

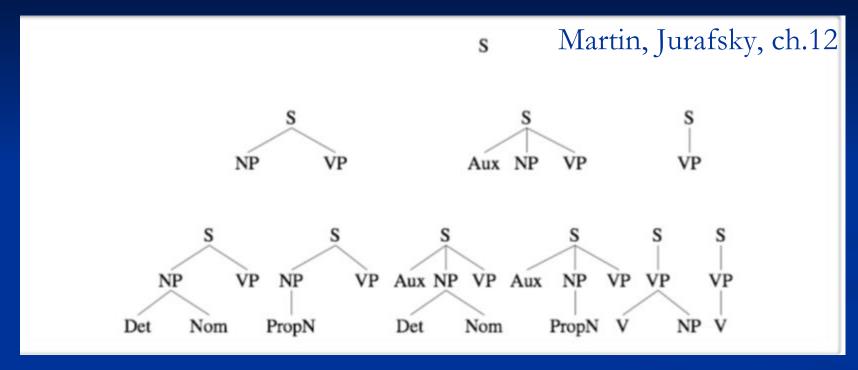
Aux → does

Preposition → from I to I on I near I through



Martin, Jurafsky, ch.12

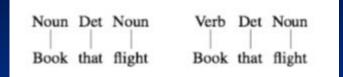
Метод рекурсивного спуска

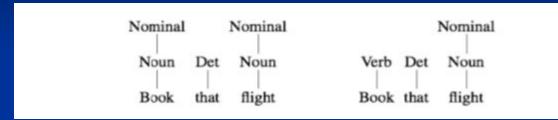


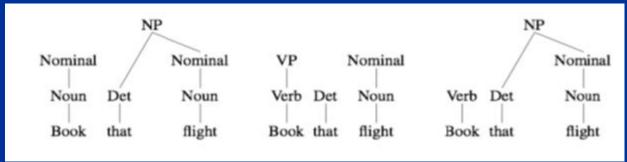
- "Разворачивать" узлы начинаем слева
- Разворачиваем каждое дерево пока не доберемся до частей речи
- отклоняем деревья, листья которых не совпадают с цепочкой частей речи анализируемого предложения

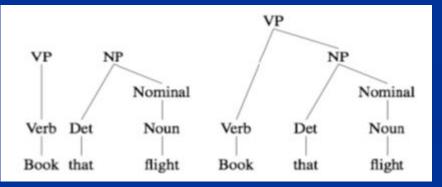
Восходящий анализ

Book that fly





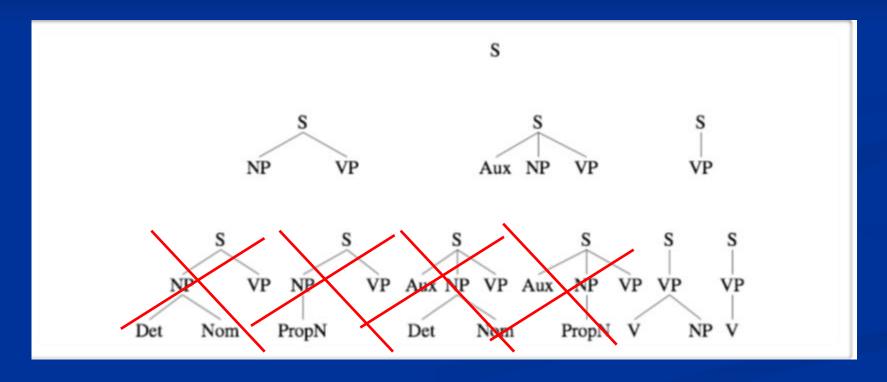




Martin, Jurafsky, ch.12

Стандартный алгоритм «сверху вниз». Фильтр «Снизу вверх»

- Проверка крайнего левого элемента в деривации



Алгоритм Кока-Янгера-Касами (CYK Parsing)

- Нормальная форма Хомского
- Допускаются только правила вида
- X -> B C, X -> w
- Петя подарил Маше книгу
- $VP \rightarrow VNPNP$

Типы «нарушающих» правил:

- (1) $X \rightarrow Bw$, (2) $X \rightarrow B$, (3) $X \rightarrow ABC$
- (1) $X \rightarrow Bw$: INF-VP -> to VP \longrightarrow INF-VP -> TO VP; TO-> to
- (2) $X \rightarrow B: NP \rightarrow Nominal, Nominal \rightarrow N$
- (3) X -> ABC: NP -> D N NP NP -> D Nominal;
 Nominal -> N NP

Алгоритм СҮК Исходная грамматика

```
S \to NP VP
S \to Aux NP VP
S \to VP
```

NP → Pronoun

NP → Proper-Noun

NP → Det Nominal

Nominal → Noun

Nominal → Nominal Noun

Nominal → Nominal PP

VP → Verb

 $VP \rightarrow Verb NP$

 $VP \rightarrow Verb NP PP$

 $VP \rightarrow Verb PP$

 $VP \rightarrow VP PP$

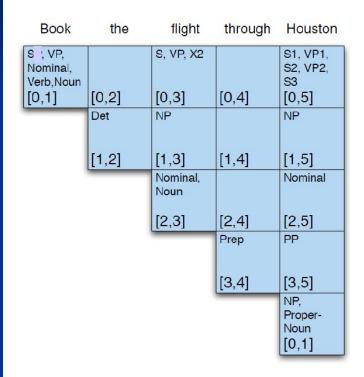
PP → Preposition NP

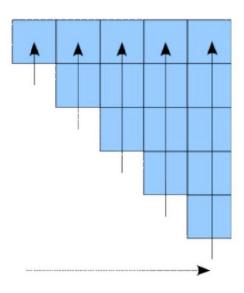
Приведение к нормальной форме Хомского

```
S \rightarrow NP VP
                                       S \rightarrow NP VP
S \rightarrow Aux NP VP
                                       S \rightarrow X1 VP
                                       X1 \rightarrow Aux NP
S \rightarrow VP
                                       S \rightarrow book \mid include \mid prefer
                                       S \rightarrow Verb NP
                                       S \rightarrow X2 PP
                                       S \rightarrow Verb PP
                                       S \rightarrow VPPP
                                       NP \rightarrow I \mid she \mid me
NP \rightarrow Pronoun
NP \rightarrow Proper-Noun
                                      NP \rightarrow TWA \mid Houston
                                      NP \rightarrow Det\ Nominal
NP \rightarrow Det Nominal
Nominal \rightarrow Noun
                                       Nominal \rightarrow book \mid flight \mid meal \mid money
Nominal → Nominal Noun
                                     Nominal \rightarrow Nominal Noun
Nominal \rightarrow Nominal PP
                                       Nominal \rightarrow Nominal PP
VP \rightarrow Verb
                                       VP \rightarrow book \mid include \mid prefer
VP \rightarrow Verb NP
                                       VP \rightarrow Verb NP
                                       VP \rightarrow X2 PP
VP \rightarrow Verb NP PP
                                       X2 \rightarrow Verb NP
                                       VP \rightarrow Verb PP
VP \rightarrow Verb PP
                                       VP \rightarrow VP PP
VP \rightarrow VP PP
PP \rightarrow Preposition NP
                                      PP \rightarrow Preposition NP
```

Figure 12.8 \mathcal{L}_1 Grammar and its conversion to CNF.

Распознавание





```
S \rightarrow NP VP 
 S \rightarrow X1 VP
```

 $XI \rightarrow Aux NP$

 $S \rightarrow book \mid include \mid prefer$

 $S \rightarrow Verb NP$

 $S \rightarrow X2 PP$

 $S \rightarrow Verb PP$

 $S \rightarrow VPPP$

 $NP \rightarrow I \mid she \mid me$

 $NP \rightarrow TWA \mid Houston$

 $NP \rightarrow Det\ Nominal$

 $Nominal \rightarrow book \mid flight \mid meal \mid money$

Nominal → *Nominal Noun*

 $Nominal \rightarrow Nominal PP$

 $VP \rightarrow book \mid include \mid prefer$

 $VP \rightarrow Verb NP$

 $VP \rightarrow X2 PP$

 $X2 \rightarrow Verb NP$

 $VP \rightarrow Verb PP$

 $VP \rightarrow VP PP$

 $PP \rightarrow Preposition NP$

```
S \rightarrow NP VP
S \rightarrow X1 VP
XI \rightarrow Aux NP
S \rightarrow book \mid include \mid prefer
S \rightarrow Verb NP
S \rightarrow X2 PP
S \rightarrow Verb PP
S \rightarrow VPPP
NP \rightarrow I \mid she \mid me
NP \rightarrow TWA \mid Houston
NP \rightarrow Det\ Nominal
Nominal \rightarrow book \mid flight \mid meal \mid money
Nominal \rightarrow Nominal Noun
Nominal \rightarrow Nominal PP
VP \rightarrow book \mid include \mid prefer
VP \rightarrow Verb NP
VP \rightarrow X2 PP
X2 \rightarrow Verb NP
VP \rightarrow Verb PP
VP \rightarrow VP PP
PP \rightarrow Preposition NP
```

Синтаксический разбор

) 18 Paparis (415) - No. 818 (1964) (616) (1800) (616) (1900)				9-19	
$S \rightarrow NP VP$	Book	the	flight	through	Houston
$S \rightarrow XIVP$					
$XI \rightarrow Aux NP$					
$S \rightarrow book \mid include \mid prefer$					
$S \rightarrow Verb NP$	[0,1]	[0,2]	[0,3]	[0,4]	[0,5]
$S \rightarrow X2 PP$	[0,1]	[0,2]	[0,0]	[0,7]	[0,0]
$S \rightarrow Verb PP$					
$S \rightarrow VPPP$					
$NP \rightarrow I \mid she \mid me$		[1,2]	[1,3]	[1,4]	[1,5]
NP → TWA Houston		[.,-]	[.,-]	[.,,.]	[.,-]
NP → Det Nominal					
Nominal → book flight meal money					
Nominal → Nominal Noun			[2,3]	[2,4]	[2,5]
Nominal → Nominal PP					PP
$VP \rightarrow book \mid include \mid prefer$					PP
$VP \rightarrow Verb NP$					200
$VP \rightarrow X2 PP$				[3,4]	[3,5]
$X2 \rightarrow Verb NP$					
$VP \rightarrow Verb PP$					
$VP \rightarrow VP PP$					
PP → Preposition NP					[0,1]

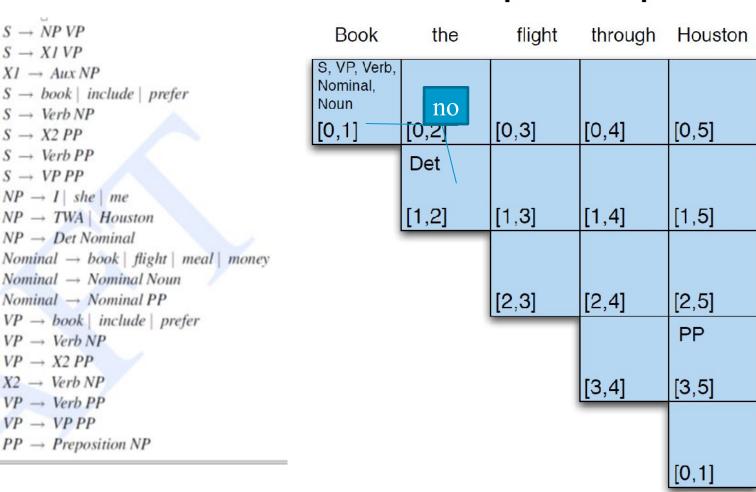
```
S \rightarrow NP VP
S \rightarrow X1 VP
XI \rightarrow Aux NP
S \rightarrow book \mid include \mid prefer
S \rightarrow Verb NP
S \rightarrow X2 PP
S \rightarrow Verb PP
S \rightarrow VPPP
NP \rightarrow I \mid she \mid me
NP \rightarrow TWA \mid Houston
NP \rightarrow Det\ Nominal
Nominal \rightarrow book \mid flight \mid meal \mid money
Nominal \rightarrow Nominal Noun
Nominal \rightarrow Nominal PP
VP \rightarrow book \mid include \mid prefer
VP \rightarrow Verb NP
VP \rightarrow X2 PP
X2 \rightarrow Verb NP
VP \rightarrow Verb PP
VP \rightarrow VP PP
PP \rightarrow Preposition NP
```

Синтаксический разбор

$S \rightarrow NP VP$	Book	the	flight	through	Houston
$S \rightarrow XIVP$	S, VP, Verb,		1		
$XI \rightarrow Aux NP$	Nominal,				
$S \rightarrow book \mid include \mid prefer$	Noun				
$S \rightarrow Verb NP$	[0,1]	[0,2]	[0,3]	[0,4]	[0,5]
$S \rightarrow X2 PP$. , ,		. , .
$S \rightarrow Verb PP$					
$S \rightarrow VPPP$					
$NP \rightarrow I \mid she \mid me$		[1,2]	[1,3]	[1,4]	[1,5]
NP → TWA Houston					
NP → Det Nominal					
Nominal → book flight meal money				and the second	
Nominal → Nominal Noun			[2,3]	[2,4]	[2,5]
Nominal → Nominal PP					PP
$VP \rightarrow book \mid include \mid prefer$					1.1
$VP \rightarrow Verb NP$					
$VP \rightarrow X2 PP$				[3,4]	[3,5]
$X2 \rightarrow Verb NP$					
$VP \rightarrow Verb PP$					
$VP \rightarrow VP PP$					FO 43
PP → Preposition NP					[0,1]
2					

```
S \rightarrow NP VP
S \rightarrow X1 VP
XI \rightarrow Aux NP
S \rightarrow book \mid include \mid prefer
S \rightarrow Verb NP
S \rightarrow X2 PP
S \rightarrow Verb PP
S \rightarrow VPPP
NP \rightarrow I \mid she \mid me
NP \rightarrow TWA \mid Houston
NP \rightarrow Det\ Nominal
Nominal \rightarrow book \mid flight \mid meal \mid money
Nominal \rightarrow Nominal Noun
Nominal \rightarrow Nominal PP
VP \rightarrow book \mid include \mid prefer
VP \rightarrow Verb NP
VP \rightarrow X2 PP
X2 \rightarrow Verb NP
VP \rightarrow Verb PP
VP \rightarrow VP PP
PP \rightarrow Preposition NP
```

Синтаксический разбор

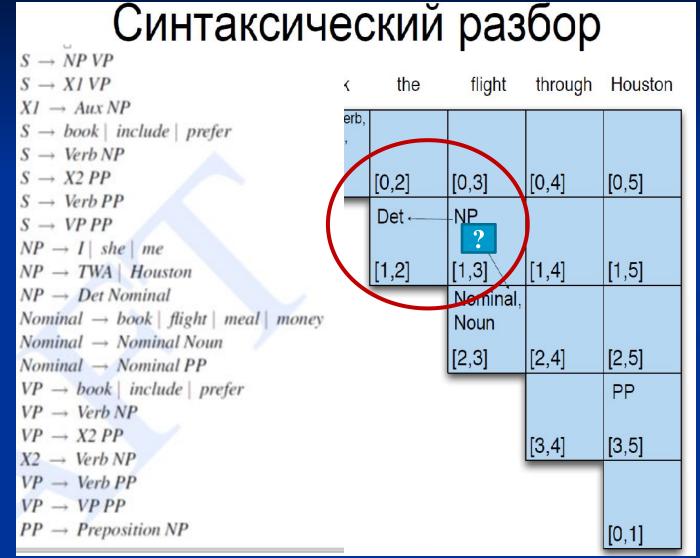


```
S \rightarrow NP VP
S \rightarrow X1 VP
XI \rightarrow Aux NP
S \rightarrow book \mid include \mid prefer
S \rightarrow Verb NP
S \rightarrow X2 PP
S \rightarrow Verb PP
S \rightarrow VPPP
NP \rightarrow I \mid she \mid me
NP \rightarrow TWA \mid Houston
NP \rightarrow Det\ Nominal
Nominal \rightarrow book \mid flight \mid meal \mid money
Nominal \rightarrow Nominal Noun
Nominal \rightarrow Nominal PP
VP \rightarrow book \mid include \mid prefer
VP \rightarrow Verb NP
VP \rightarrow X2 PP
X2 \rightarrow Verb NP
VP \rightarrow Verb PP
VP \rightarrow VP PP
PP \rightarrow Preposition NP
```

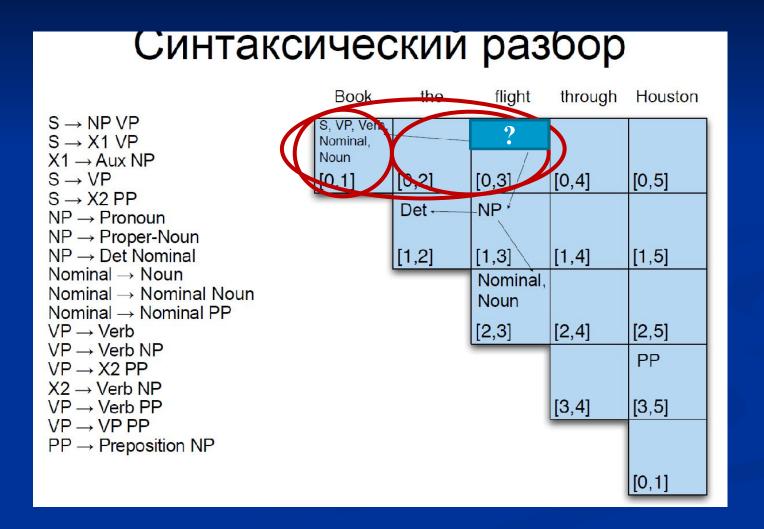
Синтаксический разбор

₩.						
$S \to NP VP S \to XI VP$	Book	the	flight	through	Houston	
$XI \rightarrow Aux NP$ $S \rightarrow book \mid include \mid prefer$ $S \rightarrow Verb NP$, VP, Verb, lominal, loun					
$S \rightarrow X2 PP$	0,1]	[0,2]	[0,3]	[0,4]	[0,5]	
$S \rightarrow Verb PP S \rightarrow VP PP$		Det				
$NP \rightarrow I \mid she \mid me$ $NP \rightarrow TWA \mid Houston$		[1,2]	[1,3]	[1,4]	[1,5]	
NP → Det Nominal Nominal → book flight meal money			Nominal, Noun			
Nominal → Nominal Noun Nominal → Nominal PP			[2,3]	[2,4]	[2,5]	
$VP \rightarrow book \mid include \mid prefer$ $VP \rightarrow Verb NP$					PP	
$VP \rightarrow X2 PP$				[3,4]	[3,5]	
$X2 \rightarrow Verb NP$ $VP \rightarrow Verb PP$						
$VP \rightarrow VP PP$ $PP \rightarrow Preposition NP$					[0,1]	

```
S \rightarrow NP VP
S \rightarrow X1 VP
XI \rightarrow Aux NP
S \rightarrow book \mid include \mid prefer
S \rightarrow Verb NP
S \rightarrow X2 PP
S \rightarrow Verb PP
S \rightarrow VPPP
NP \rightarrow I \mid she \mid me
NP \rightarrow TWA \mid Houston
NP \rightarrow Det\ Nominal
Nominal \rightarrow book \mid flight \mid meal \mid money
Nominal \rightarrow Nominal Noun
Nominal \rightarrow Nominal PP
VP \rightarrow book \mid include \mid prefer
VP \rightarrow Verb NP
VP \rightarrow X2 PP
X2 \rightarrow Verb NP
VP \rightarrow Verb PP
VP \rightarrow VP PP
PP \rightarrow Preposition NP
```



```
S \rightarrow NP VP
S \rightarrow X1 VP
XI \rightarrow Aux NP
S \rightarrow book \mid include \mid prefer
S \rightarrow Verb NP
S \rightarrow X2 PP
S \rightarrow Verb PP
S \rightarrow VPPP
NP \rightarrow I \mid she \mid me
NP \rightarrow TWA \mid Houston
NP \rightarrow Det\ Nominal
Nominal \rightarrow book \mid flight \mid meal \mid money
Nominal \rightarrow Nominal Noun
Nominal \rightarrow Nominal PP
VP \rightarrow book \mid include \mid prefer
VP \rightarrow Verb NP
VP \rightarrow X2 PP
X2 \rightarrow Verb NP
VP \rightarrow Verb PP
VP \rightarrow VP PP
PP \rightarrow Preposition NP
```

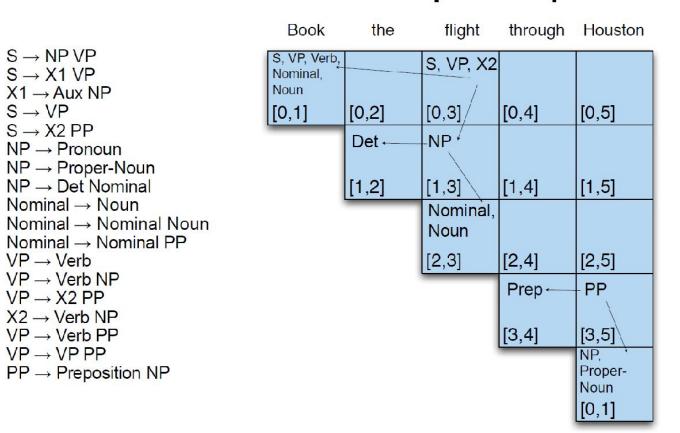


```
S \rightarrow NP VP
S \rightarrow X1 VP
XI \rightarrow Aux NP
S \rightarrow book \mid include \mid prefer
S \rightarrow Verb NP
S \rightarrow X2 PP
S \rightarrow Verb PP
S \rightarrow VPPP
NP \rightarrow I \mid she \mid me
NP \rightarrow TWA \mid Houston
NP \rightarrow Det\ Nominal
Nominal \rightarrow book \mid flight \mid meal \mid money
Nominal \rightarrow Nominal Noun
Nominal \rightarrow Nominal PP
VP \rightarrow book \mid include \mid prefer
VP \rightarrow Verb NP
VP \rightarrow X2 PP
X2 \rightarrow Verb NP
VP \rightarrow Verb PP
VP \rightarrow VP PP
PP \rightarrow Preposition NP
```



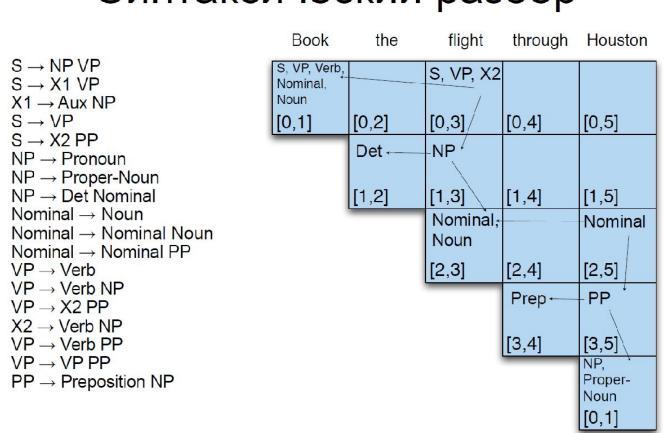
```
S \rightarrow NP VP
S \rightarrow X1 VP
XI \rightarrow Aux NP
S \rightarrow book \mid include \mid prefer
S \rightarrow Verb NP
S \rightarrow X2 PP
S \rightarrow Verb PP
S \rightarrow VPPP
NP \rightarrow I \mid she \mid me
NP \rightarrow TWA \mid Houston
NP \rightarrow Det\ Nominal
Nominal \rightarrow book \mid flight \mid meal \mid money
Nominal \rightarrow Nominal Noun
Nominal \rightarrow Nominal PP
VP \rightarrow book \mid include \mid prefer
VP \rightarrow Verb NP
VP \rightarrow X2 PP
X2 \rightarrow Verb NP
VP \rightarrow Verb PP
VP \rightarrow VP PP
PP \rightarrow Preposition NP
```

Синтаксический разбор

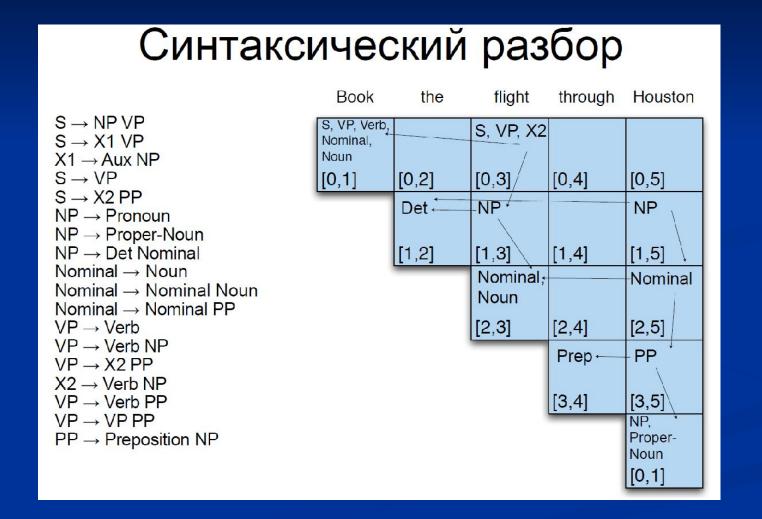


```
S \rightarrow NP VP
S \rightarrow X1 VP
XI \rightarrow Aux NP
S \rightarrow book \mid include \mid prefer
S \rightarrow Verb NP
S \rightarrow X2 PP
S \rightarrow Verb PP
S \rightarrow VPPP
NP \rightarrow I \mid she \mid me
NP \rightarrow TWA \mid Houston
NP \rightarrow Det\ Nominal
Nominal \rightarrow book \mid flight \mid meal \mid money
Nominal \rightarrow Nominal Noun
Nominal \rightarrow Nominal PP
VP \rightarrow book \mid include \mid prefer
VP \rightarrow Verb NP
VP \rightarrow X2 PP
X2 \rightarrow Verb NP
VP \rightarrow Verb PP
VP \rightarrow VP PP
PP \rightarrow Preposition NP
```

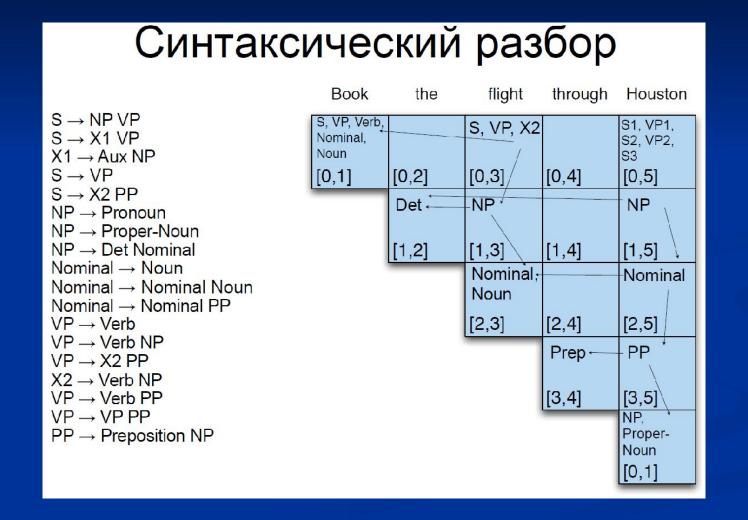
Синтаксический разбор



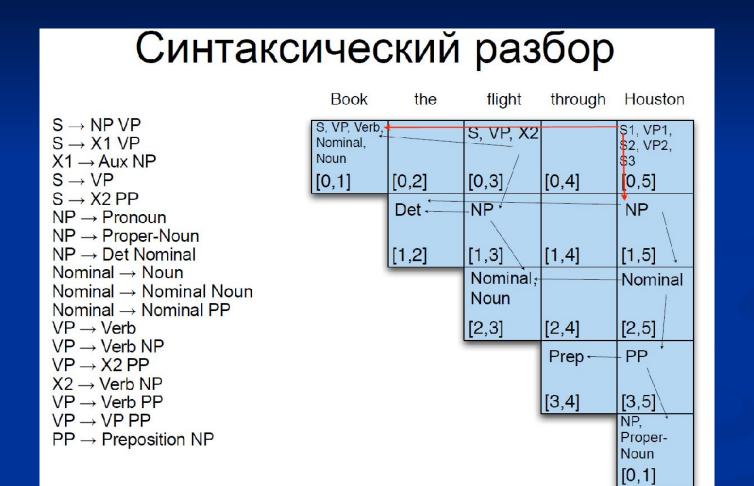
```
S \rightarrow NP VP
S \rightarrow X1 VP
XI \rightarrow Aux NP
S \rightarrow book \mid include \mid prefer
S \rightarrow Verb NP
S \rightarrow X2 PP
S \rightarrow Verb PP
S \rightarrow VPPP
NP \rightarrow I \mid she \mid me
NP \rightarrow TWA \mid Houston
NP \rightarrow Det\ Nominal
Nominal \rightarrow book \mid flight \mid meal \mid money
Nominal \rightarrow Nominal Noun
Nominal \rightarrow Nominal PP
VP \rightarrow book \mid include \mid prefer
VP \rightarrow Verb NP
VP \rightarrow X2 PP
X2 \rightarrow Verb NP
VP \rightarrow Verb PP
VP \rightarrow VP PP
PP \rightarrow Preposition NP
```



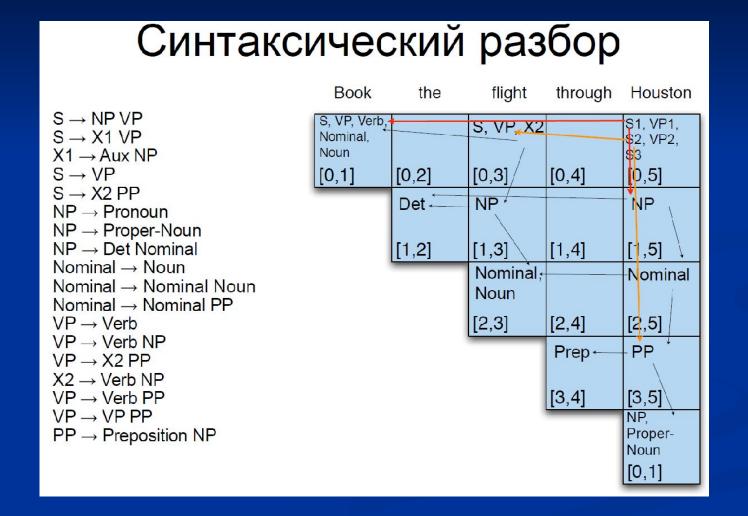
```
S \rightarrow NP VP
S \rightarrow X1 VP
XI \rightarrow Aux NP
S \rightarrow book \mid include \mid prefer
S \rightarrow Verb NP
S \rightarrow X2 PP
S \rightarrow Verb PP
S \rightarrow VPPP
NP \rightarrow I \mid she \mid me
NP \rightarrow TWA \mid Houston
NP \rightarrow Det\ Nominal
Nominal \rightarrow book \mid flight \mid meal \mid money
Nominal \rightarrow Nominal Noun
Nominal \rightarrow Nominal PP
VP \rightarrow book \mid include \mid prefer
VP \rightarrow Verb NP
VP \rightarrow X2 PP
X2 \rightarrow Verb NP
VP \rightarrow Verb PP
VP \rightarrow VP PP
PP \rightarrow Preposition NP
```



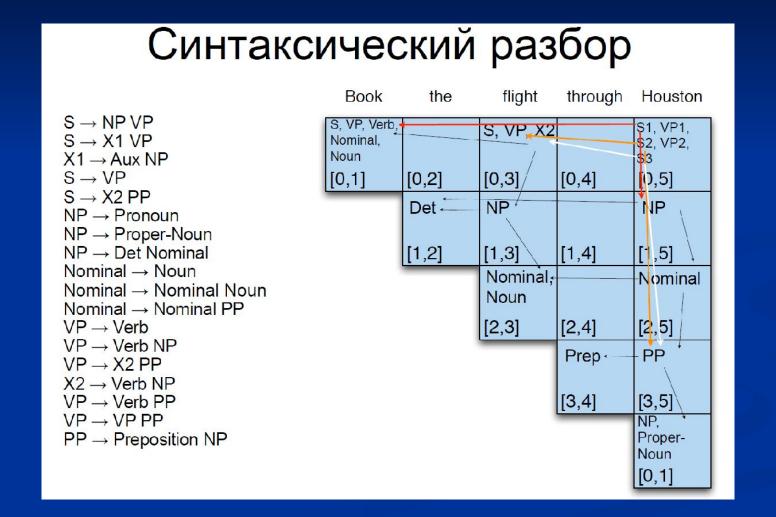
```
S \rightarrow NP VP
S \rightarrow X1 VP
XI \rightarrow Aux NP
S \rightarrow book \mid include \mid prefer
S \rightarrow Verb NP
S \rightarrow X2 PP
S \rightarrow Verb PP
S \rightarrow VPPP
NP \rightarrow I \mid she \mid me
NP \rightarrow TWA \mid Houston
NP \rightarrow Det\ Nominal
Nominal \rightarrow book \mid flight \mid meal \mid money
Nominal \rightarrow Nominal Noun
Nominal \rightarrow Nominal PP
VP \rightarrow book \mid include \mid prefer
VP \rightarrow Verb NP
VP \rightarrow X2 PP
X2 \rightarrow Verb NP
VP \rightarrow Verb PP
VP \rightarrow VP PP
PP \rightarrow Preposition NP
```



```
S \rightarrow NP VP
S \rightarrow X1 VP
XI \rightarrow Aux NP
S \rightarrow book \mid include \mid prefer
S \rightarrow Verb NP
S \rightarrow X2 PP
S \rightarrow Verb PP
S \rightarrow VPPP
NP \rightarrow I \mid she \mid me
NP \rightarrow TWA \mid Houston
NP \rightarrow Det\ Nominal
Nominal \rightarrow book \mid flight \mid meal \mid money
Nominal \rightarrow Nominal Noun
Nominal \rightarrow Nominal PP
VP \rightarrow book \mid include \mid prefer
VP \rightarrow Verb NP
VP \rightarrow X2 PP
X2 \rightarrow Verb NP
VP \rightarrow Verb PP
VP \rightarrow VP PP
PP \rightarrow Preposition NP
```



```
S \rightarrow NP VP
S \rightarrow X1 VP
XI \rightarrow Aux NP
S \rightarrow book \mid include \mid prefer
S \rightarrow Verb NP
S \rightarrow X2 PP
S \rightarrow Verb PP
S \rightarrow VPPP
NP \rightarrow I \mid she \mid me
NP \rightarrow TWA \mid Houston
NP \rightarrow Det\ Nominal
Nominal \rightarrow book \mid flight \mid meal \mid money
Nominal \rightarrow Nominal Noun
Nominal \rightarrow Nominal PP
VP \rightarrow book \mid include \mid prefer
VP \rightarrow Verb NP
VP \rightarrow X2 PP
X2 \rightarrow Verb NP
VP \rightarrow Verb PP
VP \rightarrow VP PP
PP \rightarrow Preposition NP
```



```
S \rightarrow NP VP
S \rightarrow X1 VP
XI \rightarrow Aux NP
S \rightarrow book \mid include \mid prefer
S \rightarrow Verb NP
S \rightarrow X2 PP
S \rightarrow Verb PP
S \rightarrow VPPP
NP \rightarrow I \mid she \mid me
NP \rightarrow TWA \mid Houston
NP \rightarrow Det\ Nominal
Nominal \rightarrow book \mid flight \mid meal \mid money
Nominal \rightarrow Nominal Noun
Nominal \rightarrow Nominal PP
VP \rightarrow book \mid include \mid prefer
VP \rightarrow Verb NP
VP \rightarrow X2 PP
X2 \rightarrow Verb NP
VP \rightarrow Verb PP
VP \rightarrow VP PP
PP \rightarrow Preposition NP
```

- Шаг 0. Преобразовать грамматику к нормальной форме
- Алгоритм (динамическое программирование)

```
function CKY-Parse(words, grammar) returns table

for j \leftarrow from 1 to Length(words) do

table[j-1,j] \leftarrow \{A \mid A \rightarrow words[j] \in grammar\}

for i \leftarrow from j-2 downto 0 do

for k \leftarrow i+1 to j-1 do

table[i,j] \leftarrow table[i,j] \cup

\{A \mid A \rightarrow BC \in grammar, B \in table[i,k],
C \in table[k,j]\}
```