

# Задание к лекции 1

Ольга Сопильняк

1 октября 2018 г.

## 1 Lauri Karttunen, Kenneth R. Beesley. Twenty-five years of finite-state morphology. 2005

Основная идея модели двухуровневой морфологии — использовать вместо последовательных правил вывода в контекстно-зависимой грамматике систему утверждений, которые ограничивает каждую лексему или поверхностное представление контекстом, где их можно или нельзя использовать. Таким образом, правила можно применять не последовательно, а параллельно.

Базовые принципы двухуровневой морфологии:

- Правила применяются посимвольно
- Ограничения могут затрагивать лексический или поверхностный контекст
- И лексический, и морфологический анализ применяются вместе

В такой нотации можно описывать ограничения, которые задают какое-то правило соответствия символа и его поверхностного представления и запрещают все остальные реализации этого символа. Например, правило  $p:q \Leftrightarrow \_ m$  значит « $p$  соответствует на  $q$  тогда и только тогда, когда оно идет перед  $m$ ».

В правилах также могут использоваться  $\varepsilon$ -символы. В отличие от контекстно-зависимых грамматик, задающих отображение из слова в слово или из части слова в слово, нулевые символы позволяют считать, что правила представляют собой отношения между словами равной длины,

поэтому они замкнуты относительно пересечения и могут применяться параллельно.

При построении компилятора возникли две основные проблемы: когда заменяемый символ находился в контекстной части правила и когда в правилах содержалось двойное отрицание. Обе проблемы были решены.

После публикации работы последовало много реализаций компилятора двухуровневой морфологии, одна из них позволила составлять языковую модель на основе правил с помощью лексического конечного преобразователя (трансдюсера). В языковой модели содержались леммы, лексические формы слова, фонетические формы слова и соответствие между ними.

## 2 Лингвистический анализ

### 2.1 Разновидность латинского склонения

1)  $\# \rightarrow n / [aeiou]_[aeiou]$

Пример: /Cicero#is/  $\rightarrow$  /Ciceronis/

2) Для морфем #is, #i, #e:

$\# \rightarrow n / [aeiou]_$

Пример: /#is/  $\rightarrow$  /nis/

3) Неизменяемые части:

Cicero|, Cicero|nis, Cicero|ni, Cicero|ne

Nero|, Nero|nis, Nero|ni, Nero|ne

### 2.2 Формы японских глаголов

Правила для суффиксов:

$t \rightarrow d / \_oi\#$

$t \rightarrow d / \_n\#$

Правила для корней:

$[kg] \rightarrow i / \_\#[td]$

$[bm] \rightarrow n / \_\#[td]$

$[wr] \rightarrow t / \_\#[td]$

$[w] \rightarrow \varepsilon / \_\#[aueo]$