

Двухуровневая морфология.

Традиционные грамматики состояли из упорядоченной последовательности правил преобразования, которые превращали фонологическое представление слова в поверхностное представление через несколько промежуточных уровней. Например, правило выглядит следующим образом: $a \rightarrow b / c_d$, что означает “а переходит в b, если расположена между c и d. Однако, у этого подхода есть существенный недостаток: он хорошо применим только в качестве одностороннего преобразования из лексической формы в поверхностную. Применяя этот подход для генерации лексической формы, сталкиваемся с проблемой избыточности производимых форм.

Kimmo Koskenpää предложил использовать систему двухуровневых правил, которые применялись не последовательно, а параллельно. Такие правила представляют собой ограничения на соответствия между лексической и поверхностной формой и условия, при которых это правило может, должно выполняться или запрещено. Двухуровневая морфология базируется на принципах:

- 1) Правила являются символьными ограничениями, применяемыми параллельно
- 2) Ограничения могут касаться лексического контекста, поверхностного или обоих
- 3) Словарный поиск и морфологический анализ работают вместе

Помимо прочего, такие правила сопоставляют строки одинаковой длины, появляется пустой символ “0”.

Однако, это не решает проблему избыточности. Поэтому предлагается использовать структуру данных “бор” для эффективного поиска по словарю и перехода только по существующим дугам. Таким образом, исключается генерация несуществующих в языке слов.

В итоге получаем систему, состоящую из правил и словаря, зависящих от языка в комплекте с рантайм машиной, применимой ко всем языкам.