



Иерархия Хомского

Материал из Википедии — свободной энциклопедии

[\[править вики-текст\]](#)

Иерархия Хомского — классификация [формальных языков](#) и [формальных грамматик](#), согласно которой они делятся на 4 типа по их условной сложности. Предложена профессором [Массачусетского технологического института](#), лингвистом [Ноамом Хомским](#).

Содержание [\[убрать\]](#)

- 1 Классификация грамматик
 - 1.1 Тип 0 — неограниченные
 - 1.2 Тип 1 — контекстно-зависимые
 - 1.3 Тип 2 — контекстно-свободные
 - 1.4 Тип 3 — регулярные
- 2 Классификация языков
- 3 Примечания
- 4 Источники
- 5 Литература

Классификация грамматик [\[править вики-текст\]](#)

Согласно Хомскому, [формальные грамматики](#) делятся на четыре типа. Для отнесения грамматики к тому или иному типу необходимо соответствие *всех* её правил (продукций) некоторым схемам.

Тип 0 — неограниченные [\[править вики-текст\]](#)

Грамматика с фразовой структурой *G* — это [алгебраическая структура](#), упорядоченная четвёрка (V_T, V_N, P, S) , где^[1]:

- V_T — *алфавит* (множество) терминальных символов — *терминалов*,
- V_N — *алфавит* (множество) нетерминальных символов — *нетерминалов*,
- $V = V_T \cup V_N$ — *словарь* *G*, причём $V_T \cap V_N = \emptyset$
- P — *конечное множество продукций* (правил) грамматики, $P \subseteq V^+ \times V^*$
- S — *начальный символ* (*источник*).

Здесь V^* — множество всех строк над алфавитом *V*, а V^+ — множество непустых строк над алфавитом *V*.

К типу 0 по классификации Хомского относятся **неограниченные грамматики** — [грамматики с фразовой структурой](#), то есть все без исключения формальные грамматики. Правила можно записать в виде:

$$\alpha \rightarrow \beta,$$

где α — любая непустая цепочка, содержащая хотя бы один нетерминальный^[источник не указан 765 дней] символ, а β — любая цепочка символов из алфавита.

Практического применения в силу своей сложности такие грамматики не имеют.

Тип 1 — контекстно-зависимые [\[править вики-текст\]](#)

К этому типу относятся [контекстно-зависимые \(КЗ\) грамматики](#) и неукорачивающие грамматики. Для грамматики $G(V_T, V_N, P, S)$, $V = V_T \cup V_N$ все правила имеют вид^[2]:

- $\alpha A \beta \rightarrow \alpha \gamma \beta$, где $\alpha, \beta \in V^*$, $\gamma \in V^+$, $A \in V_N$. Такие грамматики относят к контекстно-зависимым.
- $\alpha \rightarrow \beta$, где $\alpha, \beta \in V^+$, $1 \leq |\alpha| \leq |\beta|$. Такие грамматики относят к неукорачивающим.

Эти классы грамматик эквивалентны. Могут использоваться при анализе текстов на естественных языках, однако при построении компиляторов практически не используются в силу своей сложности. Для контекстно-зависимых грамматик доказано утверждение: по некоторому алгоритму за конечное число шагов можно установить, принадлежит цепочка терминальных символов данному языку или нет.

Тип 2 — контекстно-свободные [\[править вики-текст\]](#)

К этому типу относятся [контекстно-свободные \(КС\) грамматики](#). Для грамматики

$G(V_T, V_N, P, S)$, $V = V_T \cup V_N$ все правила имеют вид:

- $A \rightarrow \beta$, где $\beta \in V^+$ (для неукорачивающих КС-грамматик) или $\beta \in V^*$ (для укорачивающих), $A \in V_N$. То

есть грамматика допускает появление в левой части правила только **нетерминального символа**.

КС-грамматики широко применяются для описания синтаксиса **компьютерных языков** (см. **синтаксический анализ**).

Тип 3 — регулярные [править вики-текст]

К третьему типу относятся **регулярные грамматики** (автоматные) — самые простые из формальных грамматик. Они являются контекстно-свободными, но с ограниченными возможностями.

Все регулярные грамматики могут быть разделены на два эквивалентных класса, которые для грамматики вида III будут иметь правила следующего вида:

- $A \rightarrow B\gamma$ или $A \rightarrow \gamma$, где $\gamma \in V_T^*$, $A, B \in V_N$ (для левостолбчатых грамматик).
- $A \rightarrow \gamma B$; или $A \rightarrow \gamma$, где $\gamma \in V_T^*$, $A, B \in V_N$ (для правостолбчатых грамматик).

Регулярные грамматики применяются для описания простейших конструкций: **идентификаторов**, **строк**, **констант**, а также **языков ассемблера**, **командных процессоров** и др.

Классификация языков [править вики-текст]

Формальные языки классифицируются в соответствии с типами грамматик, которыми они задаются. Однако, один и тот же язык может быть задан разными грамматиками, относящимися к разным типам. В таком случае, считается, что язык относится к наиболее простому из них. Так, язык, описанный грамматикой с фразовой структурой, контекстно-зависимой и контекстно-свободной грамматиками, будет контекстно-свободным.

Так же, как и для грамматик, сложность языка определяется его типом. Наиболее сложные — языки с фразовой структурой (сюда можно отнести естественные языки), далее — КЗ-языки, КС-языки и самые простые — регулярные языки.

Примечания [править вики-текст]

- ↑ Кук, Бейз, 1990, с. 258,264
- ↑ Кук, Бейз, 1990, с. 268

Источники [править вики-текст]

- А. Ю. Молчанов. *Системное программное обеспечение*. Санкт-Петербург. Питер, 2006

Литература [править вики-текст]

- Джон Хопкрофт, Раджив Мотвани, Джеффри Ульман* **ГЛАВА 5. Контекстно-свободные грамматики и языки** // Введение в теорию автоматов, языков и вычислений = Introduction to Automata Theory, Languages, and Computation. — М.: «Вильямс», 2002. — С. 528. — ISBN 0-201-44124-1.
- Робин Хантер* Основные концепции компиляторов = The Essence of Compilers. — М.: «Вильямс», 2002. — С. 256. — ISBN 5-8459-0360-2.
- Кук Д., Бейз Г.* Глава 8. Языки и грамматики // Компьютерная математика = Computer Mathematics. — М.: Наука. Физматлит, 1990. — 384 с. — ISBN 5-02-014216-6.

Категория: **Формальные языки**

Последнее изменение этой страницы: 20:35, 26 июня 2014.

Текст доступен по лицензии **Creative Commons Attribution-ShareAlike**; в отдельных случаях могут действовать дополнительные условия. Подробнее см. **Условия использования**.

Wikipedia® — зарегистрированный товарный знак некоммерческой организации **Wikimedia Foundation, Inc.**

Свяжитесь с нами

[Политика конфиденциальности](#) [Описание Википедии](#) [Отказ от ответственности](#) [Разработки](#) [Мобильная версия](#)

