

Преподаватель курса  
**ФЕДОРЧЕНКО Людмила Николаевна**  
(доцент, снс СПИИРАН)

E-mail: [Inf@iias.spb.su](mailto:Inf@iias.spb.su)

Tel. +7-921-183-8657

# ТЕОРИЯ ФОРМАЛЬНЫХ ЯЗЫКОВ И ТРАНСЛЯЦИЙ

(мат-мех, 341, 344 группы)

2019

## Список литературы

1. Hopcroft, John E.; Ullman, Jeffrey D. (1968). *Formal Languages and their Relation to Automata*. Addison-Wesley.
2. Мартыненко Б.К. *Языки и трансляции*. СПб., 2004 (2-изд).
3. Ахо А. В., Сети Р., Ульман Д. Д., *Компиляторы: принципы, технологии и инструменты*. М.: Вильямс, 2001.
4. Ахо А. В., Ульман Дж., *Теория синтаксического анализа, перевода и компиляции*. Т.1. Синтаксический анализ, 612 с. Т.2. Компиляция, 487 с., М.:Мир,1978.
5. Dick Grune, *Parsing Techniques — A Practical Guides*, [ftp://ftp.cs.vu.nl/pub/dick/PTAPG\\_1st\\_Edition/BookBody.pdf](ftp://ftp.cs.vu.nl/pub/dick/PTAPG_1st_Edition/BookBody.pdf)
6. Ludmila Fedorchenko, Sergey Baranov, *Equivalent Transformations and Regularization in Context-Free Grammars*, // Bulgarian Academy of Sciences/ Cybernetics and Information Technologies (CIT) Volume 14, No 4, pp. 29-44. Sofia 2014.

## *Введение*

В современных информационных технологиях ***синтаксические методы*** играют существенную роль.

На использовании синтаксических методов основаны:

- трансляторы языков программирования (компиляторы, интерпретаторы, конверторы, кросс-трансляторы, и т. п.),
- синтаксические редакторы,
- машинный перевод,
- различные средства обработки текстовой информации и т. п.

## *Введение*

- Моделированию определённых объектов или явлений посвящены большие и значительные части теории формальных языков.

Модель может быть выражена или идентифицирована с помощью языка. Определённые задачи моделирования дали начало определённым видам языков.

## *Введение*

Типичным примером этого являются *L*-системы, введённые Аристидом Линденмайером (Aristid Lindenmayer) в конце 1960-х, предназначенные в качестве модели развития *в биологии*.

Этот и другие типы моделирования ситуаций, от молекулярной генетики и семиотики до искусственного интеллекта и искусственной жизни, представлены в [3].

## *Введение*

*Теория формальных языков и трансляций  
составляет теоретический фундамент  
этих методов.*

## *Введение*

Курс является введением в *формальную теорию языков, автоматов и трансляций*.

Как во всякой теории мы будем иметь дело не с языками, будь то естественные или искусственные, как, например, языки программирования, а с моделями, более или менее адекватно отражающими их свойства.



*Noam Chomsky. Аврам Ноам Хомский 7.12.1928*

Родился в Филадельфии Родился в Филадельфии, штат Пенсильвания Родился в Филадельфии, штат Пенсильвания, США Родился в Филадельфии, штат Пенсильвания, США) — американский Родился в Филадельфии, штат Пенсильвания, США) — американский



*Noam Chomsky. 7.12.1928*

Будут изучаться четыре модели языков, построенных Н. Хомским в середине 50-х годов прошлого века, на основе его понятия *формальной грамматики*.

## *Введение*

Публикации его работ на эту тему в свое время наделали много шума среди лингвистов и математиков, ищущих пути использования электронных вычислительных машин для целей перевода с естественных языков.

Работы Н. Хомского положили начало *математической лингвистики*, а понятие формальной грамматики легло в основу *математической теории формальных языков*.

## *Введение*

Что мы подразумеваем под термином *язык*?  
Энциклопедическое определение языка как  
”важнейшего средства общения, обмена мыслями и  
взаимного понимания в человеческом обществе” \*)

*не годится* для построения математической  
теории языков, ибо оно говорит о том, для  
чего язык служит, а не о том, чем он является  
сам по себе.

\*) Энциклопедический словарь. М.: Большая советская  
энциклопедия, 1955. С. 716.

## *Введение*

Поэтому мы определим язык абстрактно — как *математический объект*. Это даст нам возможность делать *строгие утверждения* о языках, а вернее, о математических моделях, которые более или менее адекватно отражают свойства реальных языков, естественных или искусственных, как, например, языки программирования.

## *Введение*

В формальной теории *язык* определяется как множество конечных цепочек символов из фиксированного *алфавита*, которые называются *предложениями*. Эти множества могут быть как конечными, так и бесконечными. Такое определение языка абстрагируется от понятия смысла предложения.

## *Введение*

Нас будет интересовать лишь строение предложений языка, в частности, какие предложения входят в данный язык, а какие не входят. Поэтому-то языки и называются *формальными*.

Это естественно приводит нас к вопросу о том, как *представлять языки* – объекты не всегда конечные. И существует ли вообще возможность представить любой язык конечным образом?

**Грамматики**, которые рассматриваются в теории, являются одним из способов конечного описания языков. Следуя Н. Хомскому, мы определим понятие формальной грамматики и способ её интерпретации, связывающий её с языком, через понятие **вывода**.

В лекциях мы рассмотрим ещё три типа грамматик, отличающихся ограничениями на вид их правил. Это грамматики типа 1 или **контекстно зависимые**, типа 2 или **контекстно-свободные** и типа 3 или **регулярные**.

# *Введение*

Свойства классов языков, порождаемых грамматиками соответствующих типов, вместе с соответствующими типами *распознавателей* – абстрактных устройств способных определять принадлежность цепочки символов данному языку, обсуждаются в последующих лекциях первой части курса.

Поскольку ограничения на вид правил грамматик усиливаются нарастающим порядком от типа 0 (исходное определение) к типу 3, то рассматривать мы будем, начиная с простейшего класса — **регулярных множеств и конечных автоматов**, затем **КС-языков** и **магазинных автоматов**. Именно эти классы языков и распознавателей имеют большое практическое применение.



*Alan Turing*

*23.06.1912-7.06.1954*

*Alan Mathison Turing* (born London, England—died Wilmslow, Cheshire), British mathematician and logician, who made major contributions to mathematics, cryptanalysis, logic, philosophy, and biology and to the new areas later named computer science, cognitive science, artificial intelligence.



*Donald Knuth*

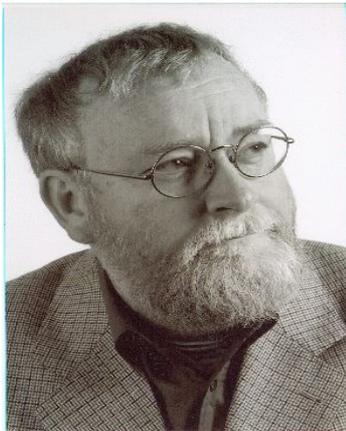
Born January 10, 1938) is a [computer scientist](#) Born January 10, 1938) is a computer scientist and [Professor](#) Born January 10, 1938) is a computer scientist and Professor at [Stanford University](#).

He is the author of multi-volume work [\*The Art of Computer Programming\*](#).

Ввёл понятие атрибутивной грамматики<sup>20</sup>

Во II части курса в КС-языки вводится понятие *семантики* посредством схем синтаксически управляемой трансляции и рассматриваются два типа трансляторов: на основе  $k$ -предсказывающего алгоритма анализа — для LL( $k$ )-грамматик и анализатора Д. Кнута — для LR( $k$ )-грамматик.

Эти подклассы КС-грамматик широко используются на практике (CDL, YACC) и потому включены в предлагаемый курс.



*Cornelis H.A.Koster (13 .07.1943)*

*He is a professor in the [Department of Informatics](#) of the [University of Nijmegen](#) in the Netherlands.  
His research interests are many:*

- syntactic methods, two-level grammars, affix grammars*
- parsing techniques and grammar transformations*
- compiler construction, translation*
- programming languages, algorithmic languages, programming concepts*



*Лавров Святослав Сергеевич*

Родился 12 марта 1923 года в Петрограде,  
Возглавлял кафедру «Математического обеспечения  
ЭВМ» более 20 лет, (ныне кафедра информатики).

В настоящее время применение технологических средств, основанных на синтаксических методах, стало обыденным делом. Однако их грамотное и эффективное применение требует от пользователя знания основ математической теории, на которой они базируются.

# Структура и содержание учебной дисциплины

## Введение

- Алфавиты и языки.
- Процедуры и алгоритмы.
- Представление языков.
- Порождающая и распознающая процедуры.
- Грамматики Хомского.
- Иерархия языков и грамматик.
- Рекурсивно-перечислимые и рекурсивные языки.
- **Автоматы**
- Машина Тьюринга.
- Язык, распознаваемый машиной Тьюринга.
- Методы построения машин Тьюринга. Модификации машины Тьюринга.
- Недетерминированная машина Тьюринга.
- Конечный автомат. Эквивалентность детерминированных и недетерминированных конечных автоматов.
- Магазинный автомат.
- Линейно-ограниченный автомат.
- Универсальная машина Тьюринга. Проблема остановки.

# Структура и содержание учебной дисциплины

## *«Конечные автоматы и регулярные языки»*

- Регулярные грамматики.
- Моделирование детерминированного и недетерминированного конечного автомата.
- Разбиение регулярного языка на классы эквивалентности правоинвариантного отношения конечного индекса. Оптимизация детерминированного конечного автомата. Эквивалентность конечных автоматов.
- Адекватность класса регулярных языков и класса конечных автоматов.
- Операции над регулярными множествами: объединение, дополнение, пересечение, замыкание. Регулярность конечных множеств.
- Теорема Клини. Регулярные выражения.
- Построение недетерминированного конечного автомата по регулярному выражению.
- Построение детерминированного конечного автомата по регулярному выражению.
- Алгоритмы определения пустоты и бесконечности регулярного языка.

# Структура и содержание учебной дисциплины

## «Контекстно-свободные грамматики (КС-грамматики)»

- Контекстно-свободные грамматики. Дерево вывода.
- Алгоритм определения пустоты контекстно-свободного языка.
- Алгоритмы удаления непродуктивных и недостижимых нетерминалов.
- Левосторонний и правосторонний выводы. Приведение произвольного вывода к левостороннему виду.
- Удаление цепных правил.
- Преобразование грамматики с пустыми правилами.
- Приведение контекстно-свободной грамматики к нормальной форме Хомского.

## «КС-грамматики и магазинные автоматы»

- Нормальная форма Грейбах.
- Лемма о разрастании (pumping lemma) для контекстно-свободных языков.
- Алгоритм определения бесконечности контекстно-свободного языка.
- Свойство самовставленности грамматик.
- Адекватность классов контекстно-свободных языков и недетерминированных магазинных автоматов.

# Структура и содержание учебной дисциплины

- **Свойства алгоритмов анализа при обнаружении ошибок во входном предложении.**
- **Двухуровневые грамматики**
- **Двухуровневые грамматики: грамматики ван Вейнгаардена и аффиксные грамматики Костера.**