DOI: 10.25728/mlsd.2022.0641

МЕТОДЫ РАСПОЗНАВАНИЯ АВТОМАТОВ ПО ИХ ГЕОМЕТРИЧЕСКИМ ОБРАЗАМ[1]

Епифанов А.С.
Институт проблем точной механики и управления РАН
Россия, г. Саратов ул. Рабочая д.24
epifanovas@list.ru

Аннотация: Рассматриваются законы функционирования дискретных динамических систем. В качестве математических моделей систем используются дискретные детерминированные автоматы, представленные с использованием аппарата геометрических образов. Разрабатываются методы распознавания автоматов по их геометрическим образам. Предложен метод сокращения диагностической информации, основанный на декомпозиции геометрических образов автоматов.

Ключевые слова: динамическая система, математическая модель, дискретный детерминированный автомат, геометрический образ автомата, распознавание автоматов, декомпозиция геометрических образов.

Введение

Одним из фундаментальных средств технического диагностирования является тестирование, теоретически разработанное как теория экспериментов с автоматами. В теории экспериментов с автоматами исходной базовой является расшифровка содержимого черного ящика. Начальными данными является информация о вариантах содержимого черного ящика. По общей схеме проведения эксперимента к черному ящику (к содержимому черного ящика) прикладываются воздействия, наблюдается реакция на эти воздействия и по этим наблюдениям строится логические выводы. В теории экспериментов с автоматами в задачах контроля задаются автомат и семейство автоматов и требуется определить, содержимым черного ящика является этот выделенный автомат или автомат из заданного семейства. В случае диагностирования (локализации неисправности технической системы по месту или функции) предполагается, что содержимым черного ящика является элемент заданного семейства автоматов и требуется определить какой это автомат. Работами Э.Мура [1], А.Гилла [2] и других авторов решены следующие задачи: найдены критерии существования решения задачи распознавания содержимого черного ящика; разработан принципиальный метод построения эксперимента по распознаванию автомата в заданном семействе автоматов, включая построение минимального по длине простого безусловного эксперимента (Э.Мур, А.Гилл). Однако, указанные результаты относятся только к одному средству получения контрольной и диагностической информации – тестированию сигналами, прикладываемыми к внешним входным узлам системы. В работе [3] показано, что тестирование, измерение значений физических параметров процессов в системе, визуальный осмотр, наблюдение звуковых и тактильных сигналов и т.д. формально может быть представлено как диагностическое воздействие на систему. Это позволило расширить тестовые автоматные методы контроля и диагностирования на общий случай средств получения диагностической информации сохраняя структуру эксперимента по распознаванию содержимого черного ящика. В этом случае входной сигнал интерпретируется не только как тестовое воздействие, но и как воздействие на систему любого средства получения контрольной и диагностической информации. Дальнейшим принципиально важным расширением подходов и методов технического диагностирования явилось представление законов функционирования автоматов геометрическими образами, т.е. числовыми математическими структурами в форме дискретных числовых графиков. Если автоматы, представленные для решения задач контроля и диагностирования их геометрическими образами, совмещать с аналитически заданными кривыми, то поиск и построение контрольных и диагностических экспериментов можно осуществлять на основе решения систем уравнений для геометрических кривых линий, заданных аналитически, а также на основе анализа числовых последовательностей.

[1] *Работа выполнена в рамках государственного задания Министерства науки и высшего образования Российской Федерации (тема № FFNM-2022-0010)*

1 Аппарат геометрических образов автоматных моделей систем

В работе [4] Твердохлебовым В.А. предложен и развит формальный аппарат замены символьных автоматных моделей в форме таблиц, графов, логических уравнений, числовыми структурами в форме геометрических фигур, числовых уравнений и последовательностей. Этот подход предназначен для поиска новых идей и методов организации технического диагностирования сложных систем. Преобразование символьной формы автоматной модели в числовую структуру (геометрический образ законов функционирования автомата) включает линейное упорядочивание автоматного отображения $\rho_s = \bigcup_{p\in X^*}\{(p,\lambda(s,p))\}$, для инициального автомата $A_s=(S, X, Y, \delta, \lambda, s)$, где S, X и Y – соответственно множества состояний, входных и выходных сигналов, а δ: S×X→S – функция переходов, λ: S×X→Y – функция выходов. Автоматное отображение ρ_s взаимнооднозначно преобразуется в автоматное отображение вида $\rho'_s = \bigcup_{p\in X^*}\{(p,\lambda'(s,p))\}$, где $\lambda'(s, p)$ – последний знак последовательности $\lambda(s, p)$. Из геометрического образа γ_s автомата A_s выделяется последовательность вторых координат точек геометрического образа, которая взаимнооднозначно соответствует полному геометрическому образу (при фиксированных порядке на множестве X^* и величине m=|X|). В результате законы функционирования автомата (то есть, фазовая картина) и конкретные процессы функционирования автомата (то есть, фазовые траектории) оказываются взаимнооднозначно определёнными последовательностью вторых координат точек геометрического образа. Это позволяет рассматривать произвольную последовательность элементов из конечного множества как последовательность вторых координат точек геометрического образа и, следовательно, как задание законов функционирования автомата.

Для этого множеством состояний автомата полагается подмножество множества $\{s_p\}_{p\in X^*}$ и функция δ определяется равенством: $\delta(s_p,x)=s_{px}$ для любых $p\in X^*$ и $x\in X$. Каждая имеющаяся в геометрическом образе автомата точка (px, y), где $p\in X^*$, $x\in X$ и $y\in Y$, определяет равенство $\lambda(s_p,x)=y$. Переход от символьных координат точек геометрического образа к числовым (в исходном случае положительным целочисленным) координатам позволяет размещать геометрические образы законов функционирования автомата на числовых кривых. Геометрический образ законов функционирования автомата с математической точки зрения является графиком с числовыми координатами точек, что позволяет принципиально неустранимую неполноту моделей больших систем устранять доопределением частично заданных графиков классических методами интерполяции и экстраполяции.

На рис.1 приведена схема геометрического образа дискретного детерминированного автомата с конечным или счетно-бесконечным множеством состояний.

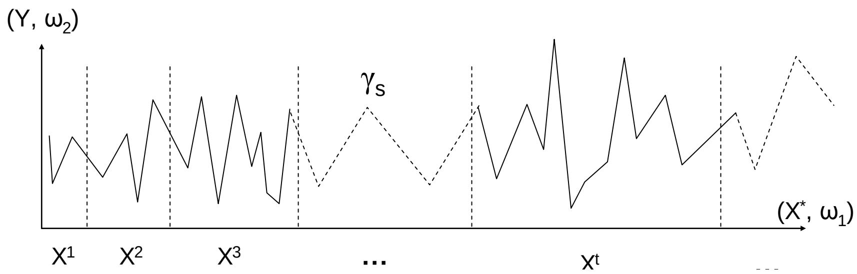

Рис.1. Геометрический образ γ_s как изображение автоматного отображения ρ_s с использованием порядков ω_1 и ω_2

Представление геометрического образа γ_s как числовой структуры позволяет использовать при постановках и в методах решения задач аппарат непрерывной математики: задание законов функционирования автоматов числовыми уравнениями, использование числовых процедур, интерполяцию и аппроксимацию частично заданных законов функционирования и т.п. Геометрический образ γ_s определяет полностью законы функционирования автомата A_s , то есть, всю фазовую картину связей входных последовательностей с выходными сигналами. Конкретные варианты процессов функционирования, то есть, фазовые траектории, имеют геометрические образы

$\gamma_s(p), p \in X^*$, в виде сечений γ_s по отдельным точкам. Геометрические образы также могут задаваться числовыми, а не символьными, уравнениями.

По одной из основных гипотез теории алгоритмов (которую в специальной литературе называют тезисом Черча-Тьюринга) в случае, если задача имеет решение (существует алгоритм решения), то существует машина Тьюринга, которая решает эту задачу. По данной гипотезе класс всех алгоритмов равномощен классу машин Тьюринга. В свою очередь машина Тьюринга является расширением конечного детерминированного автомата, класс преобразований которого ограничен числом состояний, входных и выходных сигналов. При снятии ограничения на конечность числа состояний автомата (что сделано Твердохлебовым В.А. за счет введения геометрических образов законов функционирования автоматов [4, 5]) машина Тьюринга может быть представлена как автомат с бесконечным числом состояний, считывающий и записывающий информацию с ячеек ленты.

Среди различных подходов к анализу процессов, алгоритмов, законов функционирования автоматов и реализаций этих законов для исследования выбран подход, при котором используется геометрическое представление поведения автоматов. Обоснованием такого выбора является то, что задание законов функционирования автоматов геометрическим образом однозначно определяется последовательностью вторых координат точек геометрического образа, в представлении алгоритмов схемами Янова алгоритмы также определяются линейными структурами, т.е. последовательностями, реализациям алгоритмов для конкретных исходных данных соответствуют последовательность операторов действий и последовательность, состоящая из начальных, промежуточных и заключительных данных.

2 Синтез автоматов по числовым последовательностям и распознавание автоматов по их геометрическим образам

Для изучения связей свойств законов функционирования автоматов со свойствами новых, задающих автоматы, числовых непрерывных структур проведено исследование более 100000 фундаментальных математических последовательностей из банка целочисленных последовательностей [6]. Для каждой последовательности рассмотрены начальные отрезки длин до 5 млн. элементов, по ним построены геометрические образы законов функционирования автоматов и автоматы минимизированы. При таком исследовании преобразования целочисленных последовательностей в законы функционирования рассматривались различные значения мощности входного алфавита и анализировалась специфика влияния изменения мощности входного алфавита. В данной работе предлагается метод распознавания автомата в заданном конечном семействе автоматов, основанный на выделении из геометрических образов набора характеристических последовательностей. Законы функционирования сложной системы при неисправностях из множества учитываемых неисправностей и в работоспособном состоянии представляются геометрическими образами автоматов. Поиск диагностических последовательностей при таком способе задания законов функционирования сводится к нахождению таких интервалов на оси абсцисс, в которых геометрические образы, соответствующие различным неисправностям, не совпадают. В работе [4] показано, что при зафиксированном числе входных сигналов автомата и порядке на множестве входных слов геометрическому образу взаимно-однозначно соответствует последовательность вторых координат точек геометрического образа. Ввиду этого эффективный поиск диагностических воздействий возможен на основе анализа числовых последовательностей. Каждой неисправности из множества рассматриваемых и учитываемых неисправностей сопоставляется числовая последовательность (последовательность вторых координат точек геометрического образа автомата) и задача распознавания неисправностей сводится к задаче поиска таких номеров элементов в последовательностях, значения которых различны в каждой из рассматриваемых последовательностей.

В работе [4] В.А.Твердохлебовым предложены и разработаны методы синтеза автомата по последовательностям и геометрическим кривым. Последовательность элементов из конечного множества, совмещенная с линейным порядком на множестве входных слов, определяет законы функционирования дискретной детерминированной динамической системы (автомата). В.А.Твердохлебовым предложен новый тип автомата - $(H, \mathrm{m}, d(H))$-автомат. Законы функционирования данного типа автомата задаются числовой последовательностью Н , которая полагается последовательностью вторых координат точек геометрического образа. Рассматривается начальный отрезок длины $d(H)$ последовательности H . Величина m - мощность входного алфавита автомата, количество выходных сигналов определяется спецификой начального отрезка последовательности Н длины $d(H)$ (число различных значений элементов в начальном отрезке длины $d(H)$).

В данной работе содержатся результаты по построению и анализу конечных детерминированных автоматов, законы функционирования которых определены начальными отрезками геометрических образов и выбором числа входных сигналов автомата. Для этого из банка последовательностей [6] извлечено более 100000 последовательностей длиной до 1000000 знаков, в том числе последовательности, задающие приближения следующих математических величин: π, e, $\varphi = \frac{1+\sqrt{5}}{2}$ (т.н. золотое сечение), $\sqrt{2}$, $\sqrt[3]{2}$, $\ln(2), \ln(10)$, $\zeta(3) = \sum_{x=1}^{\infty} \frac{1}{x^3}$, константы Каталана $C = \sum_{n=0}^{\infty} \frac{(-1^n)}{(2n+1)^2}$, константы Эйлера $\gamma = \lim_{n\to\infty}(1+\frac{1}{2}+\frac{1}{3}+...+\frac{1}{n}-\ln(n))$, χ_F (характеристическая последовательность для чисел Фибоначчи) и χ_P (характеристическая последовательность для простых чисел). Каждая последовательность представлена набором начальных отрезков, имеющих длины d_k знаков, где d_k=10000·k, k=1,2,...,100. Полученное множество из 1200 последовательностей рассматривается как множество начальных отрезков последовательностей вторых координат точек геометрических образов законов функционирования автоматов. Соответствующие последовательности первых координат точек геометрического образа определялись вариантами выбора числа входных сигналов автомата и линейным порядком ω_1 на множестве входных последовательностей. Возможны и другие варианты линейных порядков на X* (см., например, [4]). В данной работе исследование законов функционирования автоматов проводится с использованием порядка ω_1 . Рассматриваются множества входных сигналов, содержащие 2, 5, 10, 20, 50, 100, 200, 500, 1000 и 2000 элементов. Детальное и подробное описание метода синтеза автомата по последовательности, а также метода проверки эквивалентности состояний автомата по геометрическому образу содержится в монографии [4]. Если в качестве задания законов функционирования автомата рассматривается последовательность ξ = <u(1), u(2), … ,u(t),…> , то в ее интерпретацию включается разбиение последовательности ξ на отрезки длины m , где m = |X|. Каждый отрезок при таком разбиении определяет функционирование автомата для конкретного состояния памяти. Если множеством состояний автомата полагать множество $S=\{s_p\}_{p\in X^*}$, а функцию переходов δ определить правилами $s_0 = s_\varepsilon$ и $\delta(s_p,x) = s_{px}$, то функция δ оказывается стандартной для всех автоматов с множеством входных сигналов X . Специфика автоматов проявляется в том, что на бесконечном (счетном) множестве состояний для каждого автомата выделяются классы эквивалентных состояний. Существенным при синтезе законов функционирования автомата является способ доопределения функции переходов δ автомата. Возможно циклическое доопределение, доопределение в начальное состояние, генерация состояния псевдослучайным образом (из множества возможных состояний). В случае, когда $\frac{d(H)}{|X|} = \left[\frac{d(H)}{|X|}\right]$, где |X| - мощность входного алфавита автомата, а d(H)- длина начального отрезка последовательности H (по которой строятся законы функционирования автомата), доопределение требуется и для функции выходов λ . В данной работе исследуются все указанные способы доопределения функции переходов, а значение мощности входного алфавита и длины начальных отрезков последовательностей выбраны таким образом, что $\frac{d(H)}{|X|} = \left[\frac{d(H)}{|X|}\right]$, поэтому доопределение функции λ не требуется.

Рассматриваемому множеству из 1200 последовательностей, при трех различных значениях мощности входного алфавита и трех способах доопределения функции переходов сопоставляется класс автоматов, состоящий из 36000 элементов. Далее осуществляется минимизация автоматов из построенного класса и разбиение на подклассы автоматов по числу состояний в автомате после минимизации.

Математические модели систем являются базовой информацией для решения задач разработки схем, анализа и технического диагностирования, оптимизации структур и законов функционирования интегральных схем. В работе [4] показано, что техническое диагностирование систем, которые характеризуются как большие или сложные системы, осуществляется в условиях существенных ограничений на математические модели и средства диагностирования. Сложная система как объект диагностирования не допускает достаточно полного и точного интуитивного обозрения и формального представления традиционными средствами: таблицами, графами, логическими уравнениями. Кроме того, возможности средств диагностирования в каждом используемом интервале времени ограничены наблюдением только части структуры объекта и наблюдением только некоторых функций объекта. Анализ работоспособности и локализация дефектов могут потребоваться на интервале времени, как угодно удаленном от начала функционирования объекта. Техническое диагностирование сложных не может быть осуществлено однородными средствами диагностирования. Только совмещение тестирования, измерения физических параметров, анализ процессов "решения" объектами диагностических задач, визуальный осмотр и сигнализация и т.п. должно образовывать средства диагностирования.

Метод технического диагностирования сложных систем с использованием аппарата геометрических образов включает построение математических моделей средств технического диагностирования в форме связи диагностических взаимодействий с реакциями на них объекта диагностирования; построение математических моделей в форме геометрических образов для объекта диагностирования, включающее разработку геометрических образов на основе интерполяции и экстраполяции; разработку стратегии проведения диагностического эксперимента на основе анализа геометрических образов и реализацию диагностического эксперимента в соответствии с разработанной стратегией.

3 Метод сокращения диагностической информации, основанный на декомпозиции геометрических образов автоматов

Использование аппарата геометрических образов для задания законов функционирования дискретных детерминированных динамических систем позволяет эффективно осуществлять поиск диагностических последовательностей. Для этого каждой неисправности из множества рассматриваемых и учитываемых неисправностей и работоспособному состоянию системы сопоставляется математическая модель в виде геометрического образа законов функционирования автомата. Точки геометрических образов предполагаются расположенными на кривых, заданных аналитически. При таком способе задания диагностической модели поиск диагностических последовательностей сводится к нахождению таких интервалов на оси абсцисс, в которых геометрические кривые, представляющие математические модели неисправностей, не имеют общих точек. Эффективный поиск таких интервалов может быть осуществлен на основе решения системы неравенств (или системы равенств).

В работе [4] показано, что при зафиксированном числе входных сигналов автомата и порядке на множестве входных слов геометрическому образу взаимно-однозначно соответствует последовательность вторых координат точек геометрического образа. Ввиду этого эффективный поиск диагностических воздействий возможен на основе анализа числовых последовательностей. Каждой неисправности из множества рассматриваемых и учитываемых неисправностей сопоставляется числовая последовательность (последовательность вторых координат точек геометрического образа автомата) и задача распознавания неисправностей сводится к задаче поиска таких номеров элементов в последовательностях, значения которых различны в каждой из рассматриваемых последовательностей. При большом числе последовательностей и большой длине самих последовательностей данная задача имеет сложное решение и требует больших объемов вычислительных ресурсов. Поэтому предлагается осуществлять распознавание исходных (числовых) последовательностей на основе анализа характеристических (бинарных) последовательностей, отражающих расположение значений элементов в исходных последовательностях.

Метод диагностирования на основе декомпозиции геометрических образов автоматов включает следующие этапы:

1. Построение математической модели системы в работоспособном состоянии и математических моделей законов функционирования системы при неисправностях в форме частично заданных геометрических образов автоматов;
2. Доопределение частично заданных геометрических образов до полных на основе использования классических методов интерполяции;
3. Извлечение из полученных на основе интерполяции полных геометрических образов числовых последовательностей (последовательностей вторых координат точек);
4. Декомпозиция последовательностей в набор характеристических последовательностей без потери информации;
5. Анализ полученного множества характеристических последовательностей с целью выявления такого минимального набора характеристических последовательностей, который покрывает все неисправности из множества рассматриваемых и учитываемых неисправностей системы.

Для иллюстрации описанного метода диагностирования с сокращением общности рассуждений допустим, что множество рассматриваемых и учитываемых неисправностей системы состоит из 999 неисправностей и в результате реализации этапов 1-3 предлагаемого метода диагностирования получено следующее семейство числовых последовательностей $\{H_0, H_1, H_2, \ldots, H_{999}\}$, где H_0 - последовательность вторых координат точек геометрического образа автоматной модели работоспособной системы, а $H_1, H_2, \ldots, H_{999}$ - последовательности вторых координат точек геометрических образов автоматных моделей системы при неисправностях. С сокращением общности рассуждений в качестве последовательностей $H_0, H_1, H_2, \ldots, H_{999}$ рассматриваются цифровые последовательности длины 1000 знаков, выделенные из начального отрезка числа π длиной 1000000 знаков по следующему правилу: первый знак последовательности H_i ($0 \le i \le 999$) имеет номер $(i \cdot 1000)+1$ в числе π, а последний знак последовательности H_i - $(i+1)\cdot 1000$, т.е. начальный отрезок числа π длиной 1000000 знаков последовательно делится на 1000 подпоследовательностей одинаковой длины 1000 знаков каждая. В результате реализации этапа 4 построено множество характеристических последовательностей, состоящее из 10000 последовательностей длины 1000. Проведенный вычислительный эксперимент, реализующий этап 5 предложенного метода диагностирования, выявил следующие специфические свойства рассматриваемого множества характеристических последовательностей. Распознавание всех 1000 последовательностей $H_0, H_1, H_2, \ldots, H_{999}$ (в случае, когда анализируется все 1000 знаков) возможно с использованием любой из 10 характеристических последовательностей, т.е. каждая из последовательностей $H_0, H_1, H_2, \ldots, H_{999}$ имеет уникальное распределение всех десяти цифр. Таким образом, объем необходимой диагностической информации может быть сокращен (для каждой неисправности вместо сохранения всей исходной последовательности достаточно хранить любую из 10 характеристических последовательностей).

Кроме того, отмечены следующие свойства, позволяющие более существенно сократить объем диагностической информации:

1. при использовании всего 10 первых точек характеристических последовательностей распознается 85% последовательностей;
2. при использовании 20 первых точек характеристических последовательностей распознается более 97% последовательностей;
3. при использовании 25 первых точек характеристических последовательностей распознается более 99% последовательностей;
4. для распознавания всех 1000 последовательностей достаточно 60 первых точек характеристических последовательностей, выделяющих расположение цифры 8 в последовательностях $H_0, H_1, H_2, \ldots, H_{999}$;
5. использование любого из 10 сечений геометрического образа для распознавания всех 1000 последовательностей возможно при использовании 90 первых точек характеристических последовательностей.

Проведенный вычислительный эксперимент показывает, что использование метода диагностирования с использованием декомпозиции геометрических образов автоматов (на примере класса (π, m, d(π)) – автоматов, построенного по первым 1000000 знакам числа π) позволяет существенно сократить объем диагностической информации.

Эффективность предложенного метода сокращения диагностической информации исследована еще на 9 классах автоматов, построенных по 9 последовательностям длины 1 млн.знаков , задающим приближения следующих математических величин: e, , $\varphi = \frac{1+\sqrt{5}}{2}$ (т.н. золотое сечение), $\sqrt{2}$, $\sqrt[3]{2}$, ln(2), ln(10), $\zeta(3) = \sum_{x=1}^{\infty} \frac{1}{x^3}$, константы Каталана $C = \sum_{n=0}^{\infty} \frac{(-1^n)}{(2n+1)^2}$, константы Эйлера $\gamma = \lim_{n\to\infty}(1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}-\ln(n))$. Для каждого класса определены специфические свойства. Например, наибольшее число распознаваемых последовательностей (88%) при использовании всего 10 первых точек характеристических последовательностей, отмечено в классе автоматов, построенном по первому миллиону знаков числа $\sqrt[3]{2}$. В классе (e, m, d(e)) – автоматов для распознавания всех 1000 последовательностей достаточно знания первых 54 точек в характеристической последовательности, выделяющей расположение цифры 1.

Заключение

Используемый геометрический подход позволяет исследовать свойства законов функционирования дискретных детерминированных динамических систем большой размерности на основе анализа свойств геометрических кривых и числовых последовательностей. В работе построены и проанализированы классы дискретных детерминированных автоматов, определенные на основе математических свойств геометрических образов, задающих законы функционирования автоматов. Изложенные результаты показывают возможность практического использования аппарата геометрических образов для задания и исследования свойств законов функционирования дискретных детерминированных динамических систем. Предложен метод сокращения диагностической информации на основе декомпозиции геометрических образов. Приведена иллюстрация метода на примере класса $(\pi, \mathrm{m}, \mathrm{d}(\pi))$ – автоматов, построенного по первому миллиону знаков последовательности, задающей приближение числа π.

Литература

1. *Мур Э.* Умозрительные эксперименты с последовательными машинами / Автоматы. Сб. статей под ред. К. Шеннона и Д. Маккарти, ИИЛ, М., 1956. – С. 179-213.
2. *Гилл А.* Введение в теорию конечных автоматов. – М.: Наука, 1966. – 272с.
3. *Резчиков А.Ф., Твердохлебов В.А.* Техническое диагностирование мехатронных систем // Мехатроника, автоматизация, управление. – 2003. – № 2. – С. 2-6.
4. *Твердохлебов В.А.* Геометрические образы законов функционирования автоматов. – Саратов: Изд-во «Научная книга»,2008. – 183с. ISBN 978-5-9758-0924-7
5. *Твердохлебов В.А.* Дискретные пространства в задачах управления и диагностирования // Доклады Академии военных наук, №9, Саратов, 2003. – С. 102-108.
6. http://oeis.org/Seis.html (дата обращения 23.05.2022).
7. *Хопкрофт Д., Мотвани Р., Ульман Д.* Введение в теорию автоматов, языков и вычислений. Перев. с англ. Васылык О.И., Саит-Аметова М., Ставровского А.Б. Изд.дом "Вильямс", М.-СПб-Киев, 2002. –528 с.