

РАЗВИТИЕ И ПРИМЕНЕНИЕ МНОГОЗНАЧНЫХ ЛОГИК И СЕТЕВЫХ ПОТОКОВ В ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМАХ¹

В.С.Сгурев, В.С.Йоцов, Л.В.Лютикова, А.В.Тимофеев
Институт информационных технологий Болгарской академии наук
P.O.Box 161, Sofia 1113, BULGARIA
<sgurev@bas.bg>, <jotsov@ieee.org>
Санкт-Петербургский институт информатики и автоматизации РАН, 199178,
Санкт-Петербург, 14-я линия, д.39.
<tav@iias.spb.su>

УДК 62-50

В.В.Сгурев, В.С.Йоцов, Л.В.Лютикова, А.В.Тимофеев. Развитие и применение многозначных логик и сетевых потоков в интеллектуальных системах // Труды СПИИРАН. Вып. 2, т. 1. — СПб.: СПИИРАН, 2004.

Аннотация. Рассмотрены методы решения задач искусственного интеллекта на базе многозначных и вероятностных логик. Предложен логический и сетевой подход для вывода по аналогии. Описаны и обобщены на случай многозначных логик метод поиска закономерностей в базах знаний, логико-аксиоматический и логико-вероятностный методы обучения понятиям и распознавания образов. — Библ. 15 назв.

UDC 62-50

V.V.Sgurev, V.S.Jotsov, L.V.Lutikova, A.V.Timofeev. Development and Application for Multi-Valued Logics and Network Flows in Intelligent Systems // SPIIRAS Proceedings. Issue 2, vol. 1. — SPb.: SPIIRAS, 2004.

Abstract. Methods for solution of artificial intelligence problems on the base of multi-valued and probabilistic logics have been discussed. Logic and network approach for analogous derivation has been suggested. Method for regularity search, logic-axiomatic and logic-probabilistic methods for learning of terms and pattern recognition in the case of multi-valued logic have been described and generalized. — Bibl. 15 items.

Введение

Классическая двузначная логика связана с формализацией строго корректных (формальных) рассуждений. Однако очень часто предметная область, на базе которой строятся основные понятия и выводы, обладает неполной, неточной, противоречивой и зачастую изменчивой информацией [1–7]. В связи с этим возникает необходимость в использовании и развитии новых неклассических методов формализации интеллектуальных процессов и информационных технологий.

На сегодняшний день арсенал разнообразных неклассических логик достаточно велик [2,3,7]. Однако плохо разработана методика использования этих логик в конкретных задачах. Кроме того, возможности этих логик (например, К-значных логик) не полностью отвечают потребностям, возникающим при разработке интеллектуальных систем и технологий.

До сих пор остаётся популярным статистический подход к анализу данных и принятию оптимальных решений. Однако он требует репрезентативности исходных данных. На практике обучающие выборки данных, из которых извлека-

¹ Работа выполнена при частичной поддержке грантов РФФИ № 03-01-00224а и гранта РГНФ № 03-06-12019в

ются знания и формулируются интеллектуальные решения, очень ограничены и поэтому не являются статистически репрезентативными.

В статье рассматриваются способы применения многозначных и вероятностных логик к решению интеллектуальных задач (в частности, к задачам обучения и поиска закономерностей на примере трёхзначных логик). Используется несколько подходов к построению вывода на основе модифицируемой информации: вывод по аналогии, логико-аксиоматический и логико-вероятностный методы и моделирование сетевых потоков. Предложенные методы позволяют кооперировать логические и вероятностные подходы и получать преимущества от каждого из них.

1. Определение классов и вывод по аналогии

Пусть универсум классов V составляют подмножества S_1, S_2, \dots и $S \in V$. Каждое из подмножеств типа S включает элементы $x_{S;1}, x_{S;2}, \dots$, которые формируют новую модель. Исходное множество S относится к одному из классов $S_i \in V$. Конечный результат анализа S идентифицируется с одним из классов S_i в U . На выходе формируется ответ типа $V_S = (T; F; ?)$ со тремя значениями: "истина", "ложь", "неопределенность". В случае ответа $V_S = ?$ или $V_S = F$ множество S может быть идентифицировано с больше, чем одним из известных классов $S_{i_1}, S_{i_2}, \dots (i_1 \neq i_2 \dots)$. Ответ $V_S = T$ получается тогда и только тогда, когда изучаемый класс S совпадает с S_i .

Среди классов S_i существует зависимость типа "предшественник - наследник", (например, S_i - предшественник S_{i_1}). Таким образом, можно формировать семантические сети. Необходимо заметить, что элементы $x_{Si_1;1}, x_{Si_1;2}, \dots$, дают различия между каждым классом S_{i_1} и другими наследниками общего предшественника S_i . Все различия, которые выступают в процессе сравнения S_{i_1} с другими классами, не являющимися прямыми наследниками S_i , определяются после использования механизма наследственности.

Вывод (ответ) $V_S = T$ формируется, когда для всех предшественников

$$A'_1 \wedge A'_2 \wedge \dots \wedge A'_n,$$

S_i и для S_{i_1} , соответствующие конъюнктивные члены имеют следующий вид:

где A_k есть x_k или $\neg x_k$, A'_k может совпадать с A_k или включать A_k и аналогичные члены для других переменных.

Пусть импликативные правила типа Хорна описывают некоторую предметную область:

$$B \leftarrow \bigwedge_{i \in I} A_i. \tag{1}$$

При использовании двузначной логики в указанных правилах, если по крайней мере одна переменная A_i - не "истина", то истинность B не определена, т.е. B может иметь значения "истина" или "ложь". В случае, если соответствующее исключение из конъюнкции (1) правила, основываются на подключении члена с любым $A_k (k \in I)$, процедура для вывода изменяется. В случае, если исключение $E(C, A_k)$ и C является истинным и A_k ложным, правая часть правила B может быть истинна (как исключение).

Расширенные модели вывода с исключением были предложены, обобщены в формализованных в [9,10] в следующем виде.

$$B \leftarrow \frac{\bigwedge_{i=1}^z A_i C, E(C, A_k), \neg A_k \leftarrow C}{B \leftarrow A_1 \wedge A_2 \wedge \dots \wedge A_{p-1} \wedge \neg A_k \wedge \dots \wedge A_z} \quad (2)$$

$$B \leftarrow A_1 \wedge A_2 \wedge \dots \wedge A_{p-1} \wedge \neg A_k \wedge \dots \wedge A_z \quad (3)$$

$$C, B \leftarrow \frac{\bigwedge_{i=1}^z A_i, E(C, A_k)}{B \leftarrow A_1 \wedge \dots \wedge A_{k-1} \wedge A_{k+1} \wedge \dots \wedge A_n} \quad (4)$$

$$B \leftarrow A_1 \wedge \dots \wedge A_{k-1} \wedge A_{k+1} \wedge \dots \wedge A_n \quad (4)$$

$$C, B \leftarrow \frac{\bigwedge_{i=1}^z A_i, E(C, A_k)}{B \leftarrow A_1 \wedge \dots \wedge A_{k-1} \wedge A_{k+1} \wedge (A_k \vee C) \wedge A_{k+2} \wedge \dots \wedge A_n} \quad (4)$$

$$B \leftarrow A_1 \wedge \dots \wedge A_{k-1} \wedge A_{k+1} \wedge (A_k \vee C) \wedge A_{k+2} \wedge \dots \wedge A_n \quad (4)$$

Из формул (2)-(3) видно, что исключения — это своего рода подключение специальных правил и области их действия. Интерпретация формулы (2) основана на следующем: если существует исключение $E(C, A_k)$, связанное одним из правил с заключением B и в результате его воздействия получен A_k , то конъюнкт A_k должен быть заменен на $\neg A_k$. В случае, если C — не “истина”, то соответствующая замена не выполнима. При обращении к правилу Modus Ponens, связь между B и $\neg A_k$ ведёт к формальному логическому противоречию.

Следовательно, формирование исключений типа $E(C, A_k)$ может приводить к противоречивому результату, вызванному неполнотой в описании предметной области. В случае, если C истинно, исключение $E(C, A_k)$ подключает это значение к конъюнкту A_k , чтобы обнулить значение последних выводов. В результате A_k заменяется на C , так как проверка его значения не влияет на вывод. В случае, когда C истинно, соответствующий конъюнкт A_k прямо заменяется на C .

Для увеличения эффективности поиска предлагается использовать вывод по аналогии. Этот метод подробно изложен в [8–10].

2. Сетевые потоки и вывод по аналогии

В интеллектуальных системах важную роль играют графовые модели и сетевые потоки. Пусть задан граф $G(N, U)$ с множеством ребер U и множеством вершин N . В [10] показано, что вывод по аналогии может быть представлен как сетевой поток на графе. Геометрическая интерпретация этого представления изображена на рис. 1

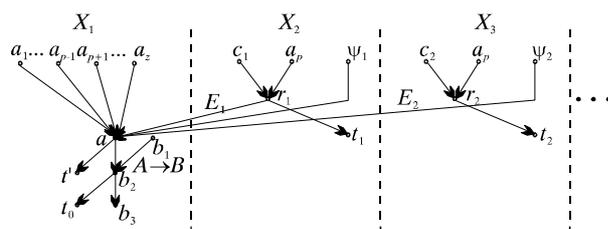


Рис. 1. Графическая интерпретация

Рис.1 разделён пунктирами на области. Каждая область содержит данные, соответствующие приблизительно одному объекту X_j . Множество S содержит все элементы A_j , исключения E и ψ_j . Множество заключений T содержит все t_v , t_0 , и t' . Тогда для всех $X, y \in N$ справедливы следующие соотношения:

$$f(y, X) - f(X, y) = \begin{cases} v(a_j), & \text{если } y \in S, \\ 0, & \text{если } y \notin S, T, \\ v(t_k), & \text{если } y \in T. \end{cases} \quad (5)$$

Здесь A_j соответствует потоковой функции $f(a_j, a)$, а C_v, A_p и $E(C_v, A_p)$ - функциям, $f(a_p, r_j), f(r_j, a)$ соответственно. Функция $f(\psi_j, a)$ первоначально принимает значение 1, если $\psi(X_j, X_1) < T$ или значение 0 - в противном случае. Некоторые соответствующие исключения $E(C_v, A_p), i \geq j$ могут быть включены в знания об объекте X_j . Пары входящих ребер дизъюнктивно соединены в вершинах r_i , и конъюнктивно соединены в вершинах a и b_2 . Функциональная зависимость v выглядит следующим образом:

$$v(a_j) = f(a_j, a), v(t_j) = f(a_j, t_j), \quad (6)$$

Конъюнкция всех A_j , и ψ_j обозначается через A и соответствуют ребру (a, b_2) . Импликация $A_j A \rightarrow B$ есть множество ребер (b_1, b_2) , а результат дедуктивного вывода $f(b_2, b_3)$ имеет значение истинности A . Тогда вывод по аналогии можно представить следующей системой равенств и неравенств [6–10]:

$$f(a, b_2) - f(a_i, a) \leq 0; i = 1, \dots, z, \quad (7)$$

$$f(a, b_2) - f(r_j, a) \leq 0; i = 1, \dots, n, \quad (8)$$

$$f(a, b_2) - f(r_j, a) \leq 0; i = 1, \dots, n, \quad (9)$$

$$f(r_j, a) - f(c_j, r_j) \geq 0; i = 1, \dots, n, \quad (10)$$

$$2f(r_j, a) - f(c_j, r_j) - f(a_p, r_{1j}) = 0; i = 1, \dots, n, \quad (11)$$

$$(2n + z - 1)f(a, b_2) - \sum_{i=1}^{z-1} f(a_i, a) - \sum_{j=1}^n f(r_j, a) - \sum_{j=1}^n f(\psi_j, a) = 0 \quad (12)$$

$$2f(b_2, b_3) - f(a, b_2) - f(b_1, b_2) = 0, \quad (13)$$

$$f(r_j, a) = 0, \text{ или } 1, \quad (14)$$

$$f(x, y) \geq 0; (x, y) \in U, \quad (15)$$

$$f(r_j, t_j) \leq 1, \quad (16)$$

$$f(a, t) \leq 2n + z - 2 f(b_2, t_0) \leq 1. \quad (17)$$

Таким образом, проблема доказательства по аналогии сведена к задаче линейного программирования с целевой функцией вида:

$$\sum_{(x,y) \in D} f(x, y) \rightarrow \max \quad (18)$$

при ограничениях (7)–(17). Более подробно эта задача описана в [10].

3. Логический вывод в задачах поиска закономерностей в базах данных и распознавания образов

Предположим, что информация о некоторой предметной области задана в виде базы данных, интерпретируемой как обучающая выборка для поиска (извлечения) логических закономерностей, связывающих эти данные. Пусть множество $Z = \{X_j, Y_j\}_{j=1}^m$ является некоторой базой данных (обучающей выборкой), связанных неизвестной зависимостью вида

$$Y = f(X), \quad (19)$$

где X и Y являются многозначными предикатами. Требуется определить зависимость (закономерность) (19) по обучающей базе данных Z мощности m . Сначала рассмотрим случай кодирования обучающей выборки двузначными предикатами. В этом случае исходная предметная область может быть описана правилами продукций вида:

$$\bigwedge_{j=1}^n x_{ij} \rightarrow y_j, j = 1, \dots, m. \quad (20)$$

Поскольку любое правило продукции является импликацией, оно может быть представлено в виде совершенной дизъюнктивной нормальной формы (СДНФ). В случае двузначной логики правило перевода импликации в ДНФ осуществляется по формулам:

$$A \rightarrow B = \neg A \vee B. \quad (21)$$

Поэтому в случае кодирования знаний двузначными предикатами любое предложение может быть представлено правилом продукции и переведено в СДНФ вида:

$$\bigvee_{j=1}^n x_j^{\sigma^{ij}} \vee y_j, \quad (22)$$

где σ^{ij} принимает значения 0 или 1.

Далее требуется объединить все формулы обучающей выборки в одну логическую функцию или систему функций, дающих однозначную интерпретацию исходной предметной области.

Таким образом, неизвестная зависимость $Y = f(X)$ может быть восстановлена непосредственно по обучающей выборке Z .

Любая логическая функция, записанная в виде СДНФ, может быть подвергнута сокращению. Поэтому и система логических знаний, как правило, тоже может быть подвергнута сокращению. Тогда сокращение СДНФ, соответствующей логической функции, может интерпретироваться как минимизация сложности исходной базы знаний.

Для сокращения СДНФ с учётом специфики предметной области предлагается следующий алгоритм:

1. Если в ДНФ имеются однолитерные дизъюнкты x и $\neg x$, то ДНФ общезначима;
2. Если некоторая переменная входит в ДНФ с одним знаком, то удаляем все дизъюнкты, содержащие эту переменную (данная переменная неинформативна);
3. Если в ДНФ имеется какой-то однолитерный дизъюнкт x , то выполняем следующие действия:
 - а) удаляем все дизъюнкты вида $x \wedge \dots$ (правило поглощения);
 - б) заменяем дизъюнкты вида $\neg x \wedge s \dots$ на дизъюнкты вида $s \wedge p \dots$.

В результате такого сокращения получаем самые “сильные” логические правила, описывающие исходную предметную область.

Описанный метод можно использовать для обучения понятием (классом) в задачах распознавания образов. Синтезированные понятия можно интерпретировать как аксиомы классов (образов) $A_k(\omega)$ в предметной области, заданной обучающей базой данных. Тогда задача распознавания образов сводится к поиску логического вывода с использованием метода резолюций Робинсона или обратного метода Маслова [11].

Задача идентификации изображения $\hat{\omega}$ k -го класса (образа) на сложном изображении ω с логическим описанием $D(\omega)$ сводится к выводу формулы:

$$D(\omega) \rightarrow \exists \hat{\omega} A_k(\hat{\omega}), \hat{\omega} \in \omega \quad (23)$$

Смысл этой формулы заключается в следующем: сложное изображение ω , имеющее логическое описание $D(\omega)$, содержит изображение $\hat{\omega}$ k -го класса, на котором истинна аксиома $A_k(\hat{\omega})$. Это позволяет автоматически идентифицировать и локализовать (выделить) изображение k -го класса (образ) на сложном изображении, содержащем изображения (образы) из M разных классов S_1, S_2, \dots, S_M .

Множественное применение логико-аксиоматического метода для каждого $k=1, 2, \dots, M$ позволяет распознать (классифицировать) все изображения всех классов, содержащиеся на сложном изображении [11].

4. Использование многозначных и вероятностных логик в задачах обучения и поиска закономерностей

Описанный метод поиска логических закономерностей может быть обобщён на случай многозначного кодирования обучающей выборки и поиска многозначных закономерностей. Использование многозначных логик осложняется неоднозначностью интерпретации функций отрицания, импликации и т.д. Поэтому рассмотрим наиболее общий вариант в случае использования трёхзначной логики.

Пусть множество значений истинности имеет вид $\{0, 1, 2\}$ со следующей интерпретацией:

$x = 0$ — ложь, $x = 1$ — бессмыслица (неопределённость), $x = 2$ — истина.

Тогда введем понятие инверсии как $\neg x = 1 \vee 0$, т. е. отрицание истины может быть либо ложью, либо бессмыслицей. Это понятие определено таблицей 1. Такое задание инверсии обеспечивает включение всех возможных интерпретаций инверсии в разных логиках.

Таблица 1. Значения истинности.

X	$\neg X$
0	$1 \vee 2$
1	$0 \vee 2$
2	$0 \vee 1$

Введем еще несколько функций трехзначной логики. Наиболее важными из них являются характеристические функции, определяемые следующим образом:

$$I_i(x) = \begin{cases} k-1, & \text{если } x = i, \\ 0, & \text{если } x \neq i, \end{cases} \quad (24)$$

$$J_i(x) = \begin{cases} 1, & \text{при } x = i, \\ 0, & \text{при } x \neq i, \end{cases} \quad (25)$$

Основные правила операций с этими функциями имеют вид:

$$I_\sigma(x)I_\tau(x) = \begin{cases} I_\sigma(x), & \text{если } \sigma = \tau, \\ 0, & \text{если } \sigma \neq \tau, \end{cases} \quad (26)$$

$$\sigma \wedge \tau = \min(\sigma, \tau), \quad \sigma \vee \tau = \max(\sigma, \tau), \quad (27)$$

Воспользуемся также двузначным аналогом импликации в данной трехзначной логике, т.е.

$$A \rightarrow B = \neg A \vee B = I_0(A) \vee I_1(A) \vee B. \quad (28)$$

Форма (28) так же, как инверсия, является расширением, включающим в себя ряд возможных импликаций трехзначной логики. Такое широкое задание основных функций логики удобно при моделировании интеллектуальных систем в тех случаях, когда не представляется возможным описать интеллектуальные процессы какой-либо одной конкретной многозначной логикой.

Теперь вернемся к решению исходной задачи в терминах трехзначной логики. Каждую строку обучающей выборки опишем правилами продукции:

$$A_{j=1}^n X_{ij} \rightarrow Y_i, \quad i = 1, \dots, m. \quad (29)$$

Тогда аналогом СДНФ будет следующая функция трехзначной логики:

Поскольку каждую закономерность (знание), согласующееся с обучающей выборкой, можно записать в виде предложенной функции трехзначной логики, хотелось бы иметь возможность представления всех закономерностей, образующих базу знаний, функцией или системой функций трехзначной логики.

Однозначное соответствие легко получить, если, например, логически перемножить правила продукции. Это соответствует рассуждениям следующего типа: знаем частные (локальные) правила и таким образом знаем все вместе локальные правила (закономерности), определяющие глобальную базу знаний, построенную по обучающей выборке.

В результате получим трёхзначную функцию, определяющую искомую закономерность. К этой функции можно применить алгоритм сокращения в адаптированном для многозначных логик варианте, а именно:

1. Если некоторая переменная входит в ДНФ с одним знаком $I_i(x)$, $i = const$. во всех дизъюнктах, то удаляем все дизъюнкты, содержащие эту переменную (данная переменная неинформативна);

2. Если в ДНФ имеется какой-то однолитерный дизъюнкт $I_i(x)$, то выполняем следующие действия:

а) удаляем все дизъюнкты вида $I_i(x) \wedge \dots$ (правило поглощения);

б) заменяем дизъюнкты вида $I_i(x) \wedge \dots (i \neq j)$ дизъюнктами вида $s \wedge p \dots$

Результатом использования алгоритма является многозначная функция, построенная по исходной обучающей выборке, однозначно её характеризующая и дающая множество наиболее существенных правил (закономерностей), определяющих исходную область знаний.

При добавлении нового правила продукции (нового знания) проверяем, выводимо ли данное правило из уже существующих или нет. Если это правило выводимо, то функция остается той же. В противном случае к базе знаний присоединяем новое правило (закономерность) путем логического многозначного перемножения имеющейся функции и нового правила продукции, записанного в виде многозначной СДНФ.

Другой метод обучения понятиям и поиска многозначных закономерностей по заданным базам данных основан на локально-оптимальных логико-вероятностных алгоритмах [12,13]. Этот метод обеспечивает автоматический синтез, оптимизацию (по точности) и минимизацию сложности баз знаний в терминах многозначных предикатов с произвольной значностью по обучающим базам данных. Он допускает интерпретацию и реализацию синтезированных знаний (закономерностей) в виде трёхслойных или многослойных нейронных сетей полиномиального типа с самоорганизующейся архитектурой [14,15].

Заключение

Предложен подход для вывода по аналогии на базе трёхзначной логики и сетевых потоков, ориентированный для применения в системах искусственного интеллекта и поддержки принятия решений. Обсуждены особенности и общие характеристики различных типов доказательства по аналогии.

Разработан метод, в котором отменяемое доказательство выполнено в форме сетевого потока. При этом подходе проблемы логического программирования сводятся к соответствующим проблемам линейного программирования.

Предложен многозначный логический подход к решению задач обучения и поиска закономерностей в базах данных, позволяющий однозначно описывать предметную область, структурно анализировать исходную информацию, сокращать её и изменять по мере формирования новых знаний, не выводимых из исходных данных.

Логико-аксиоматический и логико-вероятностный методы обучения понятиям и распознавания образов обобщены на случай многозначных логик. Показано, что синтезируемые понятия и распознающие правила могут быть реализованы в виде многозначных нейронных сетей полиномиального типа и использованы в системах интеллектуального и нейронного управления [13–14].

Литература

- [1] *Dung P., Moncarella P.* Production systems with negation as failure, IEEE, Trans/ Knowledge and Data Engineering, April 2002. Vol. 14, no.2. P. 336–352,.
- [2] *Rescher N.* Many Valued Logics. Basil Blackwell, Oxford, 1979.
- [3] *Яблонский С.В.* Введение в дискретную математику. М.: Высшая Школа, 2001.
- [4] *Gladun V.* Process of new knowledge formation. Pedagog 6, Sofia, 1994.
- [5] *Markov Z.* Inductive Machine Learning Methods, TEMPUS JEN 1497, Softex, Sofia, 1996.
- [6] *Jotsov V.* Knowledge discovery and data mining in number theory: some models and proofs // Proc. Methods and Algorithms for Distributed Information Systems Design. / Institute for Information Transmission Problems of RAS, 1997. P.197–218.
- [7] *Kowalski K.* Logic for Problem Solving. North-Holland Inc. Amsterdam, 1979.
- [8] *Halpern, J., Rabin M.* A logic to reason about likelihood // Artificial Intelligence. 1987. No. 32. P. 379–405.

- [9] *Sgurev V., Jotsov V.* Some characteristic features of the application of three-valued logical schemes in intelligent systems // Proc. First International IEEE Symp. 'Intelligent Systems' / T. Samad and V. Sgurev, Eds. Varna, Bulgaria, September 10–12, 2002. Vol. I. P. 171–176.
- [10] *Sgurev V., Jotsov V.* Some defeasible inference methods by using network flows // J. Problems of Engineering Cybernetics and Robotics. 2001. No. 51. P. 110–116.
- [11] *Timofeev A., Kossovskaya T.* Logic-Axiomatic Method for Knowledge Representation and Recognition in Intelligent Manufacturing Systems., Amsterdam: Elsevier & New York: Oxford, 1990. P. 3–6.
- [12] *Timofeev A., Shibzuchov Z.* The Synthesis and Optimization of Knowledge Bases Based on the Local-Optimization Logic-Probabilistic Algorithms // Int. Journal of Information Theories and Applications. 1995. Vol. 3, no. 2.
- [13] *Timofeev A., Yusupov R.* Evolution of Intelligent Control in Adaptive Systems // Int. Journal of Adaptive Control and Signal Processing. 1992. Vol. 6. P.193–200.
- [14] *Тимофеев А.В., Шеожев А.В., Шибзухов З.М.* Синтез нейросетевых архитектур по многозначному дереву решений // Нейрокомпьютеры: разработка и применение. 2002. № 5–6. С. 44–49.
- [15] *Timofeev A.V.* Polynomial Neural Networks with Self-Organizing Architecture // Int. Journal on Optical Memory and Neural Networks. 2004. Vol.1.