

СИСТЕМНЫЙ АНАЛИЗ НАПРАВЛЕНИЙ ВЫЧИСЛИТЕЛЬНОГО ИНТЕЛЛЕКТА

Юрий Зайченко, Михаил Згуровский

Аннотация: Исследование и анализ новой области искусственного интеллекта - вычислительный интеллект (ВИ) представлены в статье. Основные компоненты ВИ-технологий, методов и приложений, как и их взаимосвязи определяются и анализируются в этой работе. Обсуждены взаимоотношения ВИ и мягких вычислений.

Ключевые слова: вычислительный интеллект, мягкие вычисления, нейронные сети, нечеткая логика, эволюционное моделирование.

ACM Classification Keywords: I.2 Artificial Intelligence, I.2.11 Distributed Artificial Intelligence

Введение

Современный этап развития систем поддержки принятия решений характеризуется все большей их интеллектуализацией, причем процесс автоматизации творческих процессов принятия решений идет как вглубь, так и вширь, охватывая все новые и новые сферы, которые считались ранее целиком прерогативой человека. Стремление человечества переложить на компьютеры часть интеллектуальных задач, выполняемых ЛПР (лицом принимающим решения) появилось уже на заре компьютеризации. С появлением компьютеров стали возникать вопросы, может ли машина «мыслить»? Возможно ли поручить ЭВМ, часть функций мыслительных функций человека, появились работы по моделированию отдельных функций человеческого мозга, в частности по распознаванию образов, работы по моделированию мышления и психики.

В 1950 году А. Тьюринг сформулировал свой знаменитый тест проверки на интеллектуальность компьютера. Если поместить человека и компьютер в разные комнаты, а оператор с использованием клавиатуры будет задавать им вопросы и если ответы компьютера и человека совпадут, то оператор не сможет отличить человека от компьютера, и такой компьютер можно считать интеллектуальной машиной (т.е. системой искусственного интеллекта). А. Тьюринг полагал, что уже к 2000 году компьютер с объемом памяти в 10^9 бит и скоростью обработки (10^6 - 10^7) оп/с сможет пройти этот тест в течение 5 мин. с вероятностью 70%. Заметим, что в настоящее время компьютеры имеют память и быстродействие на несколько порядков выше, чем предполагал А. Тьюринг, однако пока ни один компьютер не в состоянии пройти тест Тьюринга в полном объеме, поскольку выполнение теста Тьюринга связано с рядом сложных проблем, а именно пониманием смысла текстов на естественном языке, автоматизацией поиска решений задач в различных предметных областях и рядом других, пока еще далеких от своего решения.

С момента появления первых работ в области искусственного интеллекта прошло более 50 лет. За прошедшие годы данная отрасль науки прошла сложный и поучительный путь развития. В ней сформировался ряд направлений, таких, как системы основанные на знаниях, логический вывод, поиск решений, системы распознавания образов, системы машинного перевода, обучение и самообучение,

планирование действий, агенты и мультиагентные системы, самоорганизация и самоорганизующиеся системы, нейронные сети, системы с нечеткой логикой и нечеткие нейронные сети, моделирование эмоций и психики, интеллектуальные игры, роботы и робототехнические системы.

В настоящее время существует достаточно много определений понятия искусственный интеллект. Многие из них приведены в монографиях [Рассел Стюарт, 2007], а также Ф. Люгера [Люгер Ф., 2006], в которых дается фундаментальное изложение основных направлений искусственного интеллекта. Нет смысла приводить их в данной работе. На наш взгляд, более важным является выделить **основные особенности и свойства, систем искусственного интеллекта**, отличающие их от обычных систем автоматизации. Эти свойства таковы [Згуровский М.З., 2013]:

- 1) *наличие цели или группы целей функционирования;*
- 2) *способность планирования своих действий и поиск решений задач;*
- 3) *способность к обучению и адаптации поведения в процессе работы,*
- 4) *способность работать в плохо формализованной среде, в условиях неопределенности; работать с нечеткими инструкциями;*
- 5) *способность к самоорганизации и саморазвитию,*
- 6) *способность понимать тексты на естественном языке*
- 7) *способность к обобщению и абстрагированию накопленной информации.*

Для того, чтобы создать машины, которые бы приближались по своим возможностям человеческому мозгу, необходимо прежде всего понять сущность интеллекта человека, раскрыть механизмы человеческого мышления. За истекшие десятилетия этой проблеме было посвящено много работ. Среди монографий, появившихся в последнее время, необходимо выделить монографию Джеффа Хокинса и Сандры Блейкли [Хокинс Джефф, 2007], в которой авторы, на наш взгляд, ближе всего подошли к пониманию основы человеческого интеллекта. В ней авторы, на стыке нейробиологии, психологии и кибернетики разработали пионерскую теорию, в которой построена модель мозга человека, главными функциями которой являются запоминание прошлого опыта и прогнозирование мозгом результатов восприятия окружающей действительности и своих действий. Авторы приводят множество убедительных примеров поведения человека в различной обстановке, подтверждающих эту идею. Дж. Хоукинс отмечает: «... Прогнозирование по моему мнению- это не просто одна из функций коры головного мозга, Это *первичная функция неокортекса и основа интеллекта*. Кора головного мозга является органом предвидения. Если мы хотим понять, что такое разум, что такое творчество, как работает наш мозг, и как научиться создавать разумные машины, нам нужно постичь природу прогнозов и понять, каким образом кора головного мозга их формирует. Даже поведение можно лучше всего представить как промежуточный продукт процесса прогнозирования» [Хокинс Джефф, 2007].

Целью настоящей работы является обзор и анализ современного направления в области ИИ, получившего название вычислительный интеллект.

Основные компоненты вычислительного интеллекта

В ходе развития работ в области ИИ в начале 90-х гг. путем интеграции ряда интеллектуальных технологий и методов сформировалось новое направление, получившее название *вычислительный интеллект* (computational intelligence).

Существует несколько определений термина вычислительный интеллект. Впервые термин «вычислительный интеллект» (ВИ - computational intelligence) был введен Бездеком [Bezdek J., 1994], который определил его так: «система является интеллектуальной вычислительно, если она: оперирует только с цифровыми данными; имеет компоненты распознавания образов; не использует знания в смысле искусственного интеллекта и вдобавок когда она проявляет:

- а) вычислительную адаптивность;
- б) вычислительную отказоустойчивость;
- в) уровень ошибок, аппроксимирующий характеристики человека.

В дальнейшем, это определение уточнялось и расширялось. Так, Маркс в определении ВИ делает акцент на составляющие технологии ВИ [Marks R., 1993]: «...нейронные сети, генетические алгоритмы, нечеткие системы, эволюционное программирование и искусственная жизнь являются строительными блоками ВИ».

Другая попытка определения ВИ была сделана Фогелем [Fogel D., 1995]: «эти технологии нейронных сетей, нечетких и эволюционных систем были интегрированы под вывеской «вычислительный интеллект»-сравнительно новый термин, предложенный для обобщенного описания методов вычислений, которые могут быть использованы, чтобы адаптировать решения к новым проблемам и не базироваться на явных человеческих знаниях.

За прошедшие годы было выполнено большое число работ, посвященных различным направлениям в области ВИ, регулярно проводятся международные конференции и конгрессы по вычислительному интеллекту, в Международном институте IEEE издается специализированный журнал, посвященный проблематике вычислительного интеллекта IEEE Transactions on Computational Intelligence.

Анализ этих работ позволяет дать следующее определение ВИ [Згуровский М.З., 2013].

Под вычислительным интеллектом (ВИ, computational intelligence) будем понимать совокупность технологий, моделей, методов и программных средств, предназначенных для решения неформальных, творческих задач в различных сферах человеческой деятельности с использованием аппарата и логики, в определенной степени отождествляющих мыслительную деятельность человека (нечеткость рассуждений, качественный и интуитивный подходы, креативность, логический вывод, самообучение и т.д.) в частности принятии решений, классификации, распознавании образов и т.д.

Следует отметить взаимосвязь между искусственным интеллектом (ИИ) и вычислительным интеллектом. ВИ – это составная часть направлений (разделов) современного ИИ, использующих специальные модели, методы и технологии и ориентированных на решение определенных классов задач.

Структура направлений и методов ВИ приведена на рис.1.

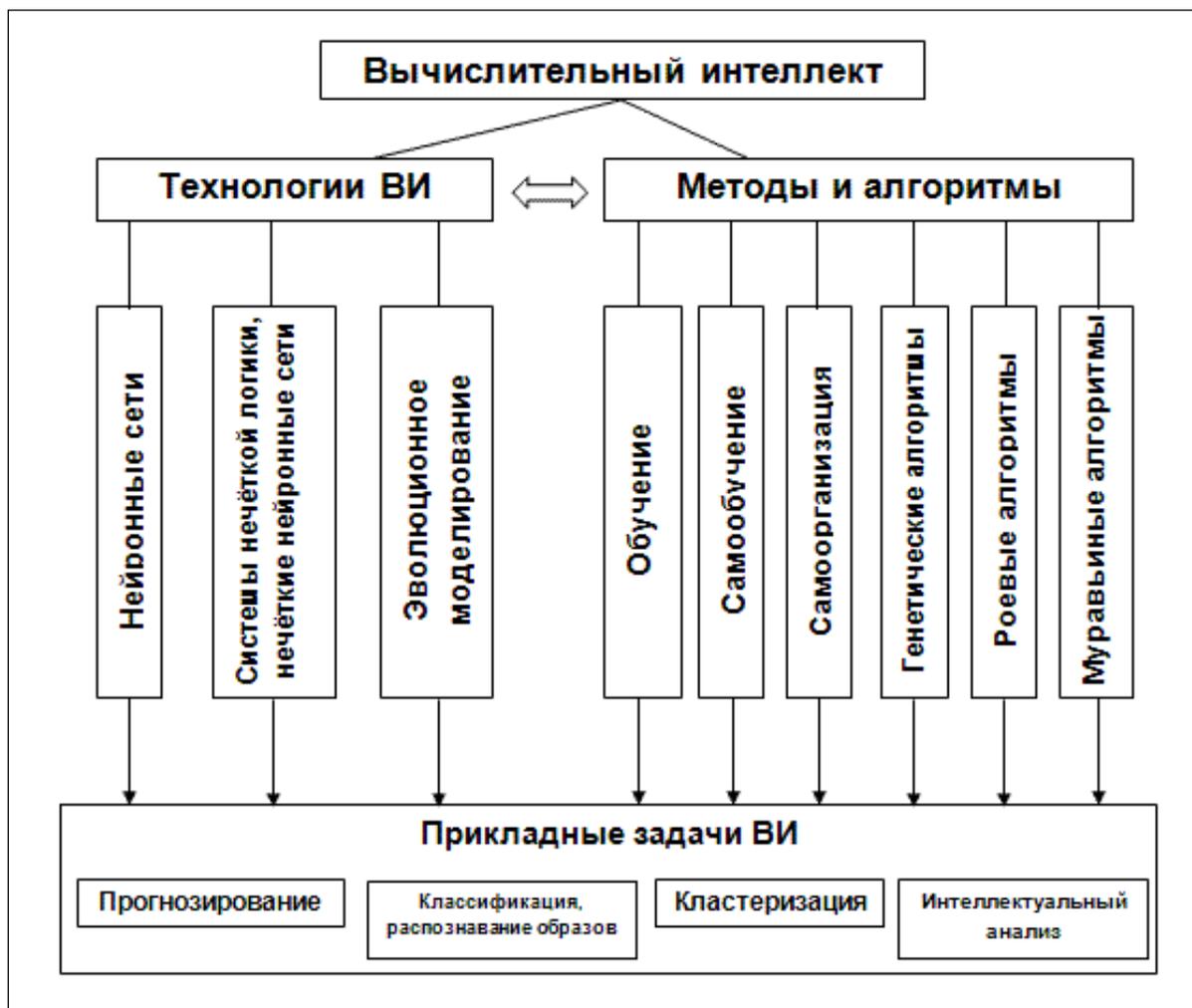


Рис. 1 Структура вычислительного интеллекта

В структуре ВИ можно выделить следующие компоненты [Згуровский М.З., 2013]:

- технологии;
- модели, методы и алгоритмы;
- прикладные задачи.

Технологии ВИ включают:

- нейронные сети (НС);
- системы нечеткой логики (СНЛ) и нечеткие нейросети (ННС);
- эволюционное моделирование (ЭМ).

К методам и алгоритмам ВИ можно отнести:

- методы обучения;
- методы самообучения;
- методы самоорганизации;

-
- генетические алгоритмы (ГА);
 - роевые алгоритмы;
 - муравьиные алгоритмы.

Технологии и методы ВИ используются при решении соответствующих задач ИИ. Логично выделить следующие основные классы задач ВИ, характерных для мыслительной деятельности человека:

- прогнозирование и предвидение;
- классификация и распознавание образов;
- кластеризация, самопроизвольное разбиение множества объектов на классы однородных объектов;
- интеллектуальный анализ данных;
- логический вывод
- принятие решений.

Взаимосвязь вычислительного интеллекта и «мягких вычислений»

Термин ВИ близок по своему значению, широко используемому в зарубежной литературе термину «soft computing» [Zadeh L.A., 1984] (т.е. «мягкие» вычисления), под которым понимается совокупность моделей, методов и алгоритмов, базирующихся на применении нечеткой математики (нечетких множеств и нечеткой логики).

Понятие мягких вычислений (soft computing) впервые было упомянуто в работе Л. Заде (Lotfi A. Zadeh) по анализу мягких (soft) данных в 1981 году. Мягкие вычисления (МВ) - это сложная компьютерная методология, основанная на нечеткой логике (НЛ), генетических вычислениях, нейрокомпьютинге и вероятностных вычислениях [Zadeh L.A., 1984]. Составные части не конкурируют, но создают эффект синергизма. Ведущий принцип МВ - это учет неточности, неопределенности, частичной истины и аппроксимации для достижения робастности, низкой цены решения, большего соответствия с реальностью.

Четыре составные части мягких вычислений включают в себя:

- нечеткую логику - приближенные вычисления, грануляция информации, вычисление на словах;
- нейрокомпьютинг - обучение, адаптация, классификация, системное моделирование и идентификация;
- генетические вычисления - синтез, настройка и оптимизация с помощью систематизированного случайного поиска и эволюции;
- вероятностные вычисления - управление неопределенностью, сети доверия, хаотические системы, предсказание.

Традиционные компьютерные вычисления (hard computing) слишком точные для реального мира. Имеется два класса задач для мягких вычислений: во-первых, существуют проблемы, для решения которых полная и точная информация не может быть получена и, во-вторых, проблемы, определение которых является не достаточно полным.

Каждая из отдельных компонент МВ обладает рядом внутренних проблем, взаимный учет которых и создает синергетический эффект. Например, нейронные сети решают задачи аппроксимации или

классификации, причем настройка весов происходит за счет обучения. Но знания распределены по многим связям и результаты работы нейронной сети кажутся пользователю необъяснимыми, то есть необходима экспликация знаний. В тоже время, системы нечеткого вывода, наоборот, обладают явными знаниями в виде продукций и позволяют легко построить протокол объяснений. Однако такая ясность достигается долгим и творческим по своей сути предшествующим процессом извлечения знаний и последующей отладкой совокупности правил.

Отсутствие возможности извлечения знаний из данных, настройки параметров функций принадлежности в ходе обучения, автоматизированной редукции правил – это характерные проблемы систем логического вывода. Применение генетических алгоритмов (ГА) требует предварительного исследования проблемы с целью выбора вероятностной модели, формулировки целевой функции.

Выбор отдельной компоненты МВ в качестве базовой обычно определяет общую архитектуру гибридной системы (под гибридной системой в данном случае понимается “мягкая” система).

Таким образом, между вычислительным интеллектом (ВИ) и мягкими вычислениями много общего: общие парадигмы, принципы, методы, технологии и прикладные задачи. Отличия их, на наш взгляд, состоят в подходах, МВ ориентированы на методологические, философские и математические проблемы, а ВИ ориентирован в основном на вычислительные алгоритмы и технологии и практическую реализации соответствующих моделей и методов.

Общая характеристика технологий и методов вычислительного интеллекта

Между технологиями и методами ВИ с позиций системного анализа можно выделить следующие взаимосвязи:

- 1) методы обучения широко используются для настройки весов связей нейронных сетей и нечетких нейронных сетей. В качестве критерия обучения здесь используются следующие критерии:
 - СКО (MSE):

$$\overline{q^2} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i(w))^2 \rightarrow \min_w \quad (1)$$

- средняя модульная процентная ошибка (MAPE):

$$\mathcal{E}_{cano}^2 = \frac{1}{N} \sum_{i=1}^N \frac{|y_i - \hat{y}_i(w)|}{|y_i|} \rightarrow \min_w \quad (2)$$

где y_i - желаемый выход НС для i -го входного сигнала; \hat{y}_i - фактический выход НС для i -го входа; W - матрица весов, $W = |w_{kj}|_{k=1, \dots, n, j=1, \dots, m}$.

- 2) при обучении, как правило, задается обучающая выборка $L = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$, где x_i - входной вектор ($x \in R^{(n)}$), y_i - выходной сигнал (желаемый выход – цель обучения).

- 3) для обучения используются методы или алгоритмы оптимизации, в частности: градиентный метод и его модификации [16, 20, 44], метод сопряженных градиентов (СГ), методы второго порядка (метод Ньютона), а также широко используются ГА [15].

В том случае, когда учитель отсутствует (т.е. цель y_i неизвестна или не задана) используются методы и алгоритмы самообучения. В этом случае задача состоит в автоматическом разбиении выборки L на ряд подмножеств, называемых «кластерами». При этом задаются соответствующие критерии и метрики разбиения.

Одной из первых технологий ВИ явились искусственные нейронные сети (ИНС). Они являются моделями нейронной системы (мозга) человека и высокоорганизованных животных. Первой известной ИНС явился разработанный Фр. Розенблаттом в 1962-65 гг. перцептрон (от слова perceptio – восприятие) [Розенблатт Ф., 1965].

Первый вариант перцептрона – трехслойный перцептрон (элементарный) состоял из 3 типов нейронов: S - нейронов – сенсорных элементов, A - нейронов – ассоциативных, и R - нейронов – реагирующих, представляющих выходную реакцию нейросети на входной сигнал.

Перцептрон является моделью зрительного механизма мозга человека.

При проектировании изображения на сенсорное поле из S - нейронов, представляющих чувствительные фотоэлементы, S - нейроны (i, j) , на которые попадают элементы изображения, срабатывают и выдают сигнал $S_{ij}(t) = 1$. Сенсорные элементы соединены случайными связями с A - элементами, аналогично тому как в зрительном механизме человека. A - связи (i, j) бывают двух типов: возбуждающие, их вес $w_{ij} = 1$, и тормозящие $w_{ij} = -1$. Число входных возбуждающих связей x , а тормозящих y - одинаково для всех A -нейронов.

Суммарный сигнал g_j на входе ассоциативного нейрона A_j определяется так:

$$g_j = \sum_{i \in P_j} w_{ij} s_i = \sum_{i \in P_j^a} s_i - \sum_{i \in P_j^o} s_i \quad (3)$$

где P_j - множество входных связей (сигналов нейрона A_j), P_j^a - множество возбуждающих связей у нейрона A_j , P_j^o - множество тормозящих связей. Очевидно, $|P_j^a| = x$, $|P_j^o| = y$.

Выходной сигнал ассоциативного нейрона A_j определяется

$$a_j = f_1(g_j)$$

где $f_1(\cdot)$ - функция активации A -нейронов.

Здесь чаще всего используется 2 вида функций:

- а) релейная $a_j = \begin{cases} 1, & \text{если } g_j \geq \theta, \text{ где } \theta - \text{величина порога;} \\ 0, & \text{в противном случае.} \end{cases}$

б) сигмоидальная $f_1(x) = \frac{1}{1 + e^{-x}}$.

A -нейроны связаны с R -нейронами «каждый с каждым». Обозначим через w_{jk}^o - вес связи нейрона A_j с нейроном R_k . Суммарный сигнал S_k на входе R_k определяется из условия:

$$S_k = \sum_{j=1}^J a_j w_{jk}^o \quad (4)$$

Срабатывает тот нейрон R_k , на входе которого сигнал S_k максимален. Сработавший нейрон определяет соответствующий класс объектов (образ). Если обозначить через r_k выходную реакцию нейрона R_k , то

$$r_k = \begin{cases} 1, & \text{если } S_k = \max_l S_l; \\ 0, & \text{в противном случае.} \end{cases} \quad (5)$$

Веса связей $\|w_{jk}^o\|$ настраиваются так, чтоб обеспечить требуемую классификацию изображений. Для этого используются соответствующие алгоритмы обучения. Фр. Розенблатт разработал несколько алгоритмов обучения элементарного перцептрона, α -алгоритм обучения и γ -алгоритм обучения. Это так называемые алгоритмы «с поощрением – наказанием» [Розенблатт Ф., 1965].

В случае правильной реакции перцептрона веса активных входных связей срабатывающего нейрона R_k увеличиваются, поощряются так:

$$w_{jk}(t+1) = w_{jk}(t) + \Delta w. \quad (6)$$

В случае неправильной реакции, наоборот, наказываются:

$$w_{jk}(t+1) = w_{jk}(t) - \Delta w.$$

Это так называемый « α -алгоритм обучения».

γ -алгоритм обучения отличается тем, что в нем веса связей изменяются так, что общая сумма весов у каждого нейрона остается прежней.

Фр. Розенблатт провел множество экспериментальных исследований 3-х слойного перцептрона по распознаванию простейших изображений: цифр, букв и геометрических фигур. Эти эксперименты показали способность перцептрона обучаться правильной классификации образов. Тем самым были убедительно продемонстрированы интеллектуальные возможности перцептрона как модели зрительного механизма мозга.

Опыты Розенблатта стимулировали интерес ученых к нейронным сетям. В конце 60-х – начале 70-х годов возник «бум» работ по исследованию НС. Появились новые архитектуры нейронных сетей, в частности работы Б. Уидроу предложившего нейрон Adaline и НС Madaline, обучающие матрицы Штейнбуха, распознающая система Альфа А.Г. Ивахненко.

Усилия многих ученых были направлены на выяснение потенциальных возможностей перцептрона. Математик П. Новиков доказал ряд теорем о сходимости процесса обучения перцептрона к искомой реакции. Однако ожидания ученых относительно чрезвычайно широких возможностей перцептрона как модели механизма мозга не оправдались. В процессе дальнейших исследований было выяснено, что элементарный трехслойный перцептрон не обладает такими неотъемлемыми свойствами интеллекта человека, как способность к экстраполяции (обобщению), т.е. перцептрон обученный распознавать некий объект в одной области сенсорного поля не способен правильно распознать его в другой области (без дополнительного обучения). Кроме того, перцептрон не обладает инвариантностью к изменению масштаба и повороту изображения.

Наконец, в 1971 г. профессор из МТИ М. Минский и С. Пайперт опубликовали книгу [Минский М., 1971], в которой показали, что трехслойный перцептрон не способен решать даже простую логическую задачу – реализовать логическую функцию XOR – исключительное «Или», в которой бинарные объекты $X_1 = [0; 0]$ и $X_2 = [1; 1]$ должны быть отнесены к одному классу ($k=1$), а $X_3 = [1; 0]$ и $X_4 = [0; 1]$ - к другому ($k=2$) (см. рис.3).

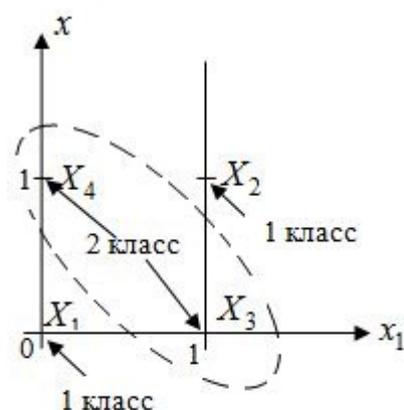


Рис. 3. Задача XOR для размерности $n=2$

Правда, Фр. Розенблатт в ответ на эту критику предложил усложнить структуру перцептрона, перейдя к 4-х слойному перцептрону с двумя настраиваемыми слоями связей. Такой перцептрон способен легко решить задачу XOR. Однако в нем возникает новая проблема – непонятно, как настраивать 2 слоя весов, чтобы обеспечить сходимость обучения к требуемой реакции. В связи с этой критикой математиков работы по НС были незаслуженно приостановлены на целых 10 лет, и был нанесен урон целому направлению в области ВИ.

Лишь спустя 10 лет, в результате работ Андерсона, Хинтона, Уильямса была предложена НС Back Propagation (обратного распространения) [Хайкин Саймон, 2006], представляющая по структуре многослойный перцептрон (MLP). Для этой сети был предложен алгоритм обратного распространения (ошибки), который позволяет настраивать сеть с произвольным числом слоев и обеспечивать сходимость к требуемой реакции (классификации или прогнозу).

Была доказана теорема об универсальной аппроксимации с помощью НС, в которой утверждается, что «существует НС ВР, позволяющая реализовать любое непрерывное отображение: $y = F(x)$ при некоторых достаточно обобщенных предположениях».

Эта теорема обосновывает широчайшие потенциальные возможности НС ВР, и стимулировала резкий рост исследований в области НС.

Кроме классических ИНС с последовательными связями между слоями нейронов, в последние годы появились нейронные сети с обратными связями (рекуррентные НС [Хайкин Саймон, 2006]). Например, НС Хопфилда и Хемминга. Такая сеть за счет использования положительной ОС обладает удивительным свойством: при определенных условиях поведение системы в процессе ее работы стремится к устойчивому состоянию (неподвижной точке), которое называется аттрактором. Аттракторы представляют собой память системы; а нейронная сеть Хопфилда – модель ассоциативной памяти человека. Когда на вход НС Хопфилда подается зашумленное или искаженное изображение объекта, она как бы по ассоциации «узнает» его и выдает на выходе неискаженное изображение – эталон, который является одним из аттракторов.

В настоящее время НС широко применяются в разнообразных сферах, в частности, прогнозировании, классификации и распознавании образов, аппроксимации функций, управлении технологическими процессами.

Достоинством НС является способность к обучению и настройке весов связей между нейронами с целью получения требуемой реакции. Вторым достоинством НС является параллельная обработка входной информации, что обеспечивает высокую производительность НС в целом, несмотря на медленную скорость передачи информации по нейронным связям и низкую скорость работы самих нейронов.

Важным свойством интеллекта является *способность к самоорганизации, самообучению*, когда «учитель» в системе отсутствует, и «правильная» классификация объекта неизвестна (если речь идет о распознавании образов). В таком случае система сама, рассматривая различные объекты (изображения) в пространстве признаков, разбивает их на подмножества, называемые кластерами, по степени «похожести», близости объектов в пространстве признаков. При этом могут использоваться самые различные метрики близости.

Нейронные сети с самоорганизацией (в частности, НС Кохонена), используя соревновательное обучение, проявляют способность к автоматической классификации объектов самой различной природы [Хайкин Саймон, 2006].

В задачах обучения и самообучения нейронных сетей настраиваются веса связей $w = w'$ с учителем или без него. При этом структура сети не изменяется.

Однако зачастую встречаются задачи, когда в процессе настройки необходимо не только настраивать веса, но и менять структуру сети. В таких случаях необходимо использовать модели и методы самоорганизации.

К числу методов самоорганизации относится, в частности, метод индуктивного моделирования, так называемый метод группового учета аргументов (МГУА). Метод используется для восстановления неизвестной функции прогноза или классификации $y = f(x_1, x_2, \dots, x_n)$ по экспериментальным

данным. Метод был предложен в конце 60-х годов 20-го века академиком А.Г. Ивахненко [Ивахненко А.Г., 1976].

Он использует основные идеи эволюции скрещивания (кроссинг-овер) родительских пар и генерацию потомков, селекцию (отбор лучших потомков), проверку условий останова.

Для восстановления неизвестной функции n переменных в классе полиномов (т.н. полиномов Колмогорова-Габора) вида

$$y = f(x_1, x_2, \dots, x_n) = a_0 + \sum_{i=1}^n a_i x_i + \sum_{j \geq i} \sum_{i=1}^n a_{ij} x_i x_j + \sum_{k \geq j} \sum_{j \geq i} \sum_{i=1}^n a_{ijk} x_i x_j x_k + \dots \quad (7)$$

используются элементарные функции от каждой пары переменных, называемые частичными описаниями (11) следующего вида:

$$y = \varphi(x_i, x_j) = \begin{cases} a_0 + a_i x_i + a_j x_j - \text{линейные;} \\ a_0 + a_i x_i + a_j x_j + a_{ii} x_i^2 + a_{jj} x_j^2 + a_{ij} x_i x_j - \text{квадратичные.} \end{cases} \quad (8)$$

Для каждого такого частичного описания (8) по обучающей выборке $L_{обуч}$ методом МНК, находятся оценки неизвестных коэффициентов $\{a_i\}$, $\{a_{ij}\}$, а по проверочной $L_{пров}$ находим наилучшую модель.

Для этого используются внешние критерии селекции:

а) регулярности

$$\bar{\varepsilon}_{np}^2 = \frac{1}{N_{np}} \sum_{i=1}^{N_{np}} (y_i - \hat{y}_i(x))^2, \quad (9)$$

где y_i - реальный выход для i -ой точки, $\hat{y}_i(x)$ - выход модели, N_{np} - объем проверочной выборки;

б) несмещенности

$$N_{см} = \frac{1}{N_{np}} \sum_{i=1}^{N_{np}} (y_i^* - y_i^{**})^2 \quad (10)$$

Далее происходит процедура селекции – отбор F лучших частичных описаний по \min критерия $\bar{\varepsilon}_{np_i}^2$.

Число F называется «свободой выбора». Отобранные лучшие частичные описания являются входами следующего ряда синтеза. На этом 1 итерация метода заканчивается.

Процесс синтеза продолжается до тех пор, пока не будет достигнут минимум критерия регулярности, т.е.

$\min \bar{\varepsilon}_{np}^2(k)$, где $\bar{\varepsilon}_{np}^2(k) = \min_i \bar{\varepsilon}_{np_i}^2(k)$, $\bar{\varepsilon}_{np_i}^2(k)$ - значения критерия y i -ой модели на k -ой

итерации.

Таким образом, находится оптимальная модель минимальной сложности. Найдя искомую (оптимальную) модель, далее двигаемся в обратном направлении, по её связям с предыдущим рядом, и делая замену переменных, приходим в конце к искомой модели в исходных переменных x_1, x_2, \dots, x_n . Метод в

отличии от других методов идентификации моделей сложных систем, позволяет автоматически восстановить структуру искомой модели.

Это обеспечивается за счет использования *принципа самоорганизации* [1, 3]: с ростом сложности модели S , значения критерия регулярности $\bar{\mathcal{E}}_{np}^2$ сначала падает, достигает минимума, а затем остается постоянным (при отсутствии шумов) или начинает расти. Здесь S – число членов полинома Колмогорова-Габора.

Одним из существенных свойств интеллекта человека является его *способность работать с неполной нечеткой информацией*, в условиях неопределенности. Возникла потребность в создании аппарата формализованного описания нечеткой и качественной информации для принятия решений в условиях неопределенности. Такая проблема была решена Л. Заде, который ввел понятие нечеткого множества (НМ) (Fuzzy set) и разработал аппарат операций над НМ [Zadeh L.A., 1965]. Основным атрибутом НМ является его *функция принадлежности* $\mu_a(x)$, такая что $\mu_a(x) \in [0; 1]$.

Величина $\mu_a(x_i)$ – есть степень принадлежности элемента x_i к НМ A , лежащая между 0 и 1. Здесь принципиальное отличие от обычных множеств, для которых степень принадлежности элементов равна либо 1 (элемент x принадлежит), либо 0 (элемент x не принадлежит множеству A).

Следует подчеркнуть, что ФП не является плотностью распределения случайной величины A , а степень принадлежности $\mu_a(x_i)$ – это не вероятность появления значения x_i , а шансы появления такого значения по оценке ЛПР, и отображает его субъективную оценку.

Дальнейшим развитием теории НМ явилось введение Л.Заде понятия лингвистической переменной её использование для принятия решений в условиях качественной и нечеткой информации.

По определению [Zadeh L.A., 1975], лингвистическая переменная (ЛП) – это переменная (количественная или качественная), задаваемая пятеркой кортежей:

$$\text{ЛП} = \langle N, X, T, \mu, G \rangle, \quad (11)$$

где N – имя переменной;

$T = \{T_j\}$ – множество базовых термов или значений ЛП;

$X = [x_{\min}, x_{\max}]$ – универсальная шкала;

μ – оператор, ставящий в соответствие каждому значению T_j его ФП $\mu_{T_j}(x)$;

G – генератор новых термов из базовых с использованием предикатов: «НЕ», «ОЧЕНЬ», «БОЛЕЕ-МЕНЕЕ».

При этом ФП новых термов определяются по соотношениям:

$$\mu_{\text{не}T_j}(x) = 1 - \mu_{T_j}(x); \mu_{\text{очень}T_j}(x) = (\mu_{T_j}(x))^2; \mu_{\text{более-менее}T_j}(x) = \sqrt{\mu_{T_j}(x)}. \quad (12)$$

Приведем примеры:

- 1) Пусть количественная переменная «Доход фирмы» рассматривается как ЛП, возможный доход фирмы изменяется в диапазоне $[1; 10^6]$ грн. Введем базовые термы: T_1 - низкий, T_2 - низкий - средний, T_3 - средний, T_4 - средний - высокий, T_5 - высокий. Тогда ЛП Доход фирмы задается так:

$$\text{ЛП Доход фирмы} = \langle N, X, \{T_j\}, \{\mu_{T_j}(x)\}, G \rangle$$

где N - Доход фирмы;

$X = [1; 10^6]$; $T = \{T_j\}$ - набор базовых термов; а их ФП $\mu_{T_j}(x)$ приводятся на рис.4

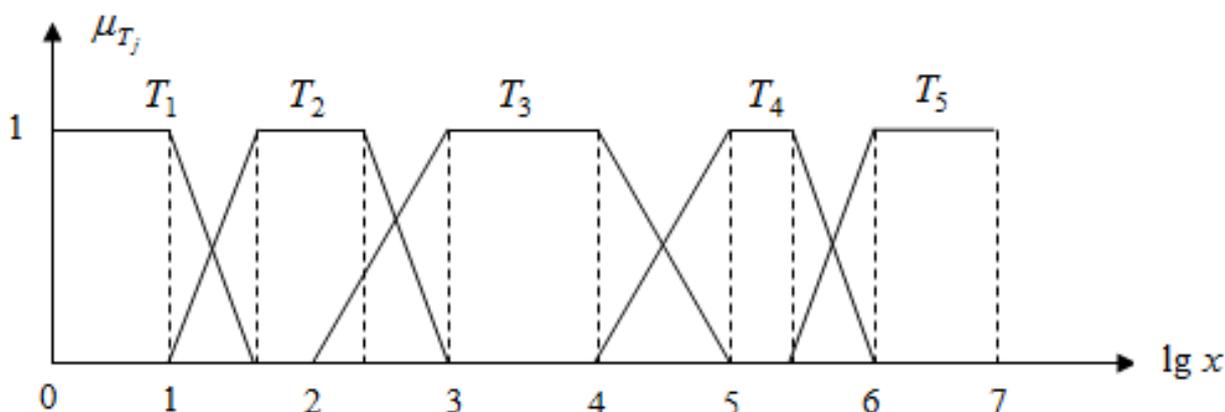


Рис. 4. ФП $\mu_{T_j}(x)$

- 2) Рассмотрим теперь качественную переменную. Например, внешность. Такую переменную задаем на универсальной шкале $X = [0; 1]$. Введем набор базовых термов для ЛП «внешность»: T_1 - безобразная, T_2 - несимпатичная, T_3 - приятной наружности (симпатичная), T_4 - красивая, T_5 - идеал красоты (например, Сикстинская Мадонна Рафаэля). ФП базовых термов представим на универсальной шкале (см. рис. 5).

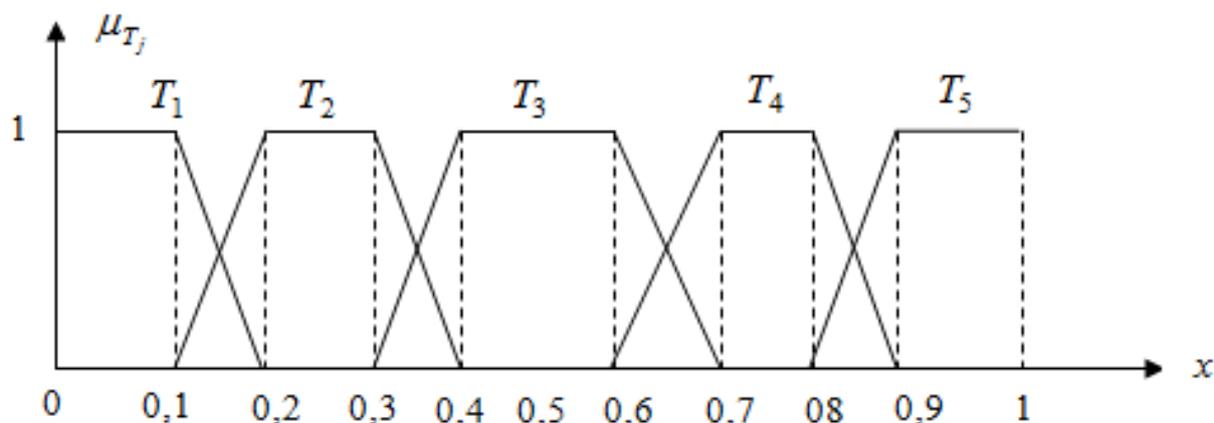


Рис. 5. ФП $\mu_{T_j}(x)$

Эти ФП отражают информацию эксперта о качественной переменной «красота», его субъективную оценку соответствия этой переменной степени наличия данного признака у конкретного лица.

Лингвистические переменные используются для принятия решений в условиях качественной и недостоверной информации в системах нечеткой логики (СНЛ). Главным элементом таких систем является нечеткая база правил (НБП), отображающая нечеткие знания эксперта о соответствующей предметной области. Нечеткие правила вывода R_k имеют вид:

$$R_1 : \text{если } x_1 \text{ есть } A_{11}, x_2 \text{ есть } A_{21}, \dots, x_n \text{ есть } A_{n1}, \text{ то } z \text{ есть } C_1 ;$$

.....

$$R_k : \text{если } x_1 \text{ есть } A_{1k}, x_2 \text{ есть } A_{2k}, \dots, x_n \text{ есть } A_{nk}, \text{ то } z \text{ есть } C_k ;$$

где x_1, x_2, \dots, x_n - входные переменные; z - выходная переменная; $\{A_{ij}\}$ и $\{C_k\}$ - значения лингвистических переменных (термов).

Рассмотрим в качестве примера нечеткие правила в задаче анализа риска банкротства корпораций в условиях неопределенности [Згуровский М.З., 2013].

Пусть для оценки финансового состояния корпораций используются следующие финансовые показатели:

x_1 - коэффициент автономности;

x_2 - коэффициент обеспеченности активов собственными средствами;

x_3 - коэффициент промежуточной ликвидности;

x_4 - коэффициент абсолютной ликвидности;

x_5 - коэффициент оборачиваемости активов;

x_6 - рентабельность всего капитала.

Все эти показатели рассматриваются как лингвистические переменные со значениями: очень низкий (ОН), низкий (Н), средний (Ср), высокий (В) и очень высокий (ОВ).

Пусть выходная переменная «риск банкротства» принимает аналогичные значения: очень высокий (ОВ), высокий (В), средний (Ср), низкий (Н) и очень низкий (ОН).

Тогда можно записать следующие правила:

R_1 : если x_1 - «ОН», x_2 - «ОН», x_3 - «ОН», x_4 - «ОН», x_5 - «ОН», x_6 - «ОН», то риск банкротства «В»;

R_k : если x_1 - «Ср», x_2 - «Ср», x_3 - «Ср», x_4 - «Н», x_5 - «ОН», x_6 - «ОН», то риск банкротства «В»;

R_M : если x_1 - «ОВ», x_2 - «ОВ», x_3 - «ОВ», x_4 - «ОВ», x_5 - «ОВ», x_6 - «ОВ», то риск банкротства «ОН».

Логический вывод осуществляется за следующие этапы [Згуровский М.З., 2013]:

- 1) Фаззификация (введение нечеткости).

- 2) *Логический вывод*, состоящий из двух подэтапов – определение степени выполнения условий правил и нахождение выхода каждого правила.
- 3) *Композиция* выходов правил.
- 4) *Дефаззификация*. Приведение к четкости (т.е. нахождение четкого выхода).

Имеется несколько алгоритмов нечеткого логического вывода, основными из которых являются: алгоритм Мамдани; Цукамото; Сугено; Ларсена.

Отметим основные достоинства систем с нечеткой логикой СНЛ [Згуровский М.З., 2013].

- 1) Они позволяют учитывать качественную и нечеткую информацию, работать в условиях неполноты и неопределенности исходной информации;
- 2) Они позволяют учитывать нечеткие знания экспертов в виде базы нечетких правил.

Вместе с тем СНЛ имеют следующие недостатки.

Для работы СНЛ необходимо задать ФП лингвистических переменных в нечетких правилах.

- 1) Эту работу выполняет человек-эксперт. Для упрощения проблемы предполагается, что все ЛП имеют один класс ФП (например, гауссовские, треугольные и трапецеидальные, и ФП разных термов отличаются лишь параметрами, которые должен указать эксперт). Однако эксперт может не знать их;
- 2) База нечетких правил, формулируемая экспертом, может оказаться неполной или противоречивой.

Для устранения указанных недостатков необходимо сделать СНЛ аддитивными и обучать по входной выборке.

Чтобы использовать арсенал накопленных методов и алгоритмов обучения, СНЛ реализуют структурно в виде нечетких нейронных сетей, в которых имеется слой входных нейронов X_1, X_2, \dots, X_n , слой нейронов правил R_1, R_2, \dots, R_k и слой выходных нейронов C_1, C_2, \dots, C_m .

При этом связи между слоями w_{ij} являются нечеткими, задаваемыми в виде некоторых ФП с неопределенными параметрами, настраиваемыми в процессе обучения. Для этого используются алгоритмы обучения, разработанные для обычных НС.

Таким образом, интеграция двух технологий: нейронных сетей и систем с нечеткой логикой позволила создать новую *гибридную технологию ВИ – нечеткие нейронные сети*, позволяющие использовать достоинства обоих исходных технологий. Подробно системы нечеткой логики и ННС рассмотрены в главах 3, 4, 5 монографии.

Системы нечеткой логики и ННС находят широкие применения во многих задачах ВИ. Их широкое распространение базируется на теореме FAT – Fuzzy Approximation Theorem об универсальной аппроксимации систем нечеткой логики.

Ванг доказал, что СНЛ с Гауссовскими ФП при увеличении числа правил $n \rightarrow \infty$ является универсальным аппроксиматором [Wang F., 1992]. Несколько позднее Коско доказал аналогичную теорему, если ФП нечетких множеств (термов) являются треугольными [Kosko B., 1994].

Заметим, что данная теорема является полным аналогом соответствующей теоремы об универсальной аппроксимации для обычных нейронных сетей. Таким образом, между свойствами обычных НС и нечетких НС имеются глубокие взаимосвязи.

Одним из свойств живых организмов биосистем в том числе человеческого мозга является способность к саморазвитию, приспособлению к применяющимся условиям внешней среды, самосовершенствованию в процессе эволюции. Как известно, биологическая эволюция базируется на следующих механизмах:

- а) скрещивание родительских пар и генерация потомков, берущих гены у обоих родителей;
- б) мутации, воздействие случайных факторов;
- в) селекция – отбор лучших особей, переходящих в следующее поколение.

Этот процесс продолжается многократно, в результате идет постепенное целенаправленное улучшение особей.

Именно эти механизмы были использованы в эволюционном моделировании (ЭМ) – важной технологии ВИ. Основная идея ЭМ – это заменить процесс построения сложной системы (распознавание образов, технической диагностики) процессом его эволюции [Fogel L. J., 1966]. Этот процесс проводится естественно в ускоренном масштабе времени, требуются тысячи или десятки тысяч поколений, чтобы достичь желаемого результата, желаемого качества функционирования.

ЭМ широко используется в задачах структурного синтеза сложных систем РО и классификации, медицинской диагностики, прогнозирования, систем управления сложными объектами и процессами.

Следует отметить, что крайне важным в ЭМ является использование механизма мутаций, который обеспечивает необходимое разнообразие особей и исключает возможность вырождения популяций.

Вышеперечисленные технологии ВИ базируются и широко используют соответствующие методы и алгоритмы. В частности, НС и ННС используют обучения и адаптации, такие как градиентный, сопряженных градиентов, стохастическая аппроксимация, РМНК, современные алгоритмы обучения Качмаржа, Марквардта, Гульвина-Рамаджи-Кейнса и др.

Кроме того, они также используют для своей работы методы и алгоритмы самообучения и самоорганизации: Кохонена, методы кластерного анализа, к-средних (четкий), дисперсионный алгоритм, метод Уорда, «ближайшего соседа», нечеткие методы кластерного анализа: К-средних, Густавссона Кесселя, на основе нечеткого отношения квазитолерантности, адаптивные нечеткие методы кластерного анализа [Згуровский М.З., 2013].

Наконец, технология ЭМ базируется на использовании генетических алгоритмов (ГА), использующих те же механизмы эволюции. ГА отличаются способами выбора родительских пар, кроссовера, механизмов мутаций способом селекций [Gen M, 1996].

Технологии и методы ВИ широко используются для решения прикладных задач ВИ. Так, нейронные сети и нечеткие НС, метод МГУА используются для прогнозирования нестационарных временных процессов, в частности в экономике и финансовой сфере.

НС и МГУА широко используются в задачах классификации и распознавания образов, в задачах диагностики, в том числе технической и медицинской [Згуровский М.З., 2013].

Нейронные сети с самоорганизацией, методы кластерного анализа (четкие и нечеткие) применяются в задачах автоматической классификации объектов по их признакам сходства – различия.

Эволюционное моделирование, генетические алгоритмы используются для структурного синтеза сложных систем распознавания образов, классификации, а также оптимизации структуры информационно-компьютерных сетей. Кроме того, ГА находят широкое применение в задачах комбинаторной оптимизации на графах и сетях.

Системы с нечеткой логикой и ННС эффективно используются в задачах анализа финансового состояния корпорации, прогнозирования риска банкротства корпораций и банков, оценке кредитоспособности заемщиков банковских капиталов в условиях неопределенности [Згуровский М.З., 2013].

Таким образом, подводя итог следует сказать, что современные технологии и методы ВИ тесно взаимодействуют друг с другом, имеется глубокое взаимопроникновение методов и алгоритмов вычислительного интеллекта в соответствующие технологии, и наоборот. А в целом технологии и методы ВИ являются отображением и технической реализацией свойств и способностей интеллекта человека в различных областях практической деятельности человека.

Заключение

В работе дано определение термина вычислительный интеллект и определены его основные компоненты. Указаны взаимосвязи между вычислительным интеллектом и мягкими вычислениями.

Дан обзор современных направлений вычислительного интеллекта, проанализированы их основные свойства, взаимосвязи и возможности применения.

Отмечены взаимосвязи между направлениями и методами вычислительного интеллекта. Дальнейшее развитие вычислительного интеллекта по-видимому будет идти в нескольких направлениях.

Во-первых – это расширение сфер применения задач ВИ, новые приложения в различных задачах и предметных областях, например в экономике, финансовой сфере, телекоммуникационных системах, управлении технологическими процессами и т.д.

Во-вторых – это развитие и совершенствование самих методов ВИ, в частности, генетических (ГА) и эволюционных алгоритмов (ЭА), роевых алгоритмов оптимизации, иммунных алгоритмов, муравьиных алгоритмов. Одним из перспективных направлений здесь является адаптация и самообучение параметров ГА и ЭА с целью ускорения сходимости и повышения точности в задачах оптимизации. Актуальным является также развитие и совершенствование параллельных генетических алгоритмов.

В – третьих, это дальнейшая интеграция различных технологий ВИ, например, интеграция нечеткой логики и генетических и эволюционных алгоритмов в задачах принятия решений и распознавания образов в условиях неполноты информации, роевых алгоритмов и алгоритмов их обучения и самообучения.

Acknowledgement

The paper is published with financial support by the project ITHEA XXI of the Institute of Information Theories and Applications FOI ITHEA (www.ithea.org) and the Association of Developers and Users of Intelligent Systems ADUIS Ukraine (www.aduis.com.ua).

Bibliography

- [Fogel D., 1995] D. Fogel, (1995) Review of "Computational intelligence: imitating life," IEEE Trans. Neural Networks, 6, 1995.-pp. 1562–1565
- [Fogel L.J., 1966] L.J. Fogel, A. Owens, and M. Walsh. Artificial Intelligence through Simulated Evolution. John Wiley & Sons, 1966.
- [Bezdek J., 1994] J.C. Bezdek, "What is computational intelligence?" in Computational Intelligence Imitating Life, Zurada, J. M., R. J. M. II, and C. J. Robinson, Eds., IEEE Press, New York, 1–12, 1994.
- [Gen M, 1996] Gen M. and Cheng R.: Genetic Algorithms and Engineering Design, John Wiley & Sons, Inc., 1996, 410 p.
- [Holland J.H., 1990] J.H. Holland. ECHO: Explorations of Evolution in a Miniature World. In J.D. Farmer and J. Doyne, editors, Proceedings of the Second Conference on Artificial Life, 1990.
- [Kosko B., 1994] Kosko B.: Fuzzy Systems as Universal Approximators, // IEEE Transaction on Computers, 1994, N.11, pp. 1329-1333
- [Marks R., 1993] Marks R., (1993) "Intelligence: computational versus artificial," IEEE Trans. Neural Networks, 1993. 4, 737–739.
- [Wang F., 1992] Wang F.: "Neural Networks Genetic Algorithms, and Fuzzy Logic for Forecasting," Proceedings, International Conference on Advanced Trading Technologies, New York, July 1992, pp. 504-532.
- [Zadeh L.A., 1965] Zadeh L.A.: Fuzzy sets // Information and Control, 1965, V.8, № 3, p. 338-353.
- [Zadeh L.A., 1975] Zadeh L.A.: The Concept of a Linguistic variable and its application to approximate reasoning, // Part 1 and 2, Information Sciences, 1975, v.8-p. 199-249, 301-357.
- [Zadeh L.A., 1984] Zadeh L.A. Theory of commonsense knowledge: aspects of vagueness / Dordrecht: D.Reidel, 1984, p. 257-296.
- [Згуровский М.З., 2011] Згуровский М.З. Зайченко Ю.П. Модели и методы принятия решений в нечетких условиях.-К.: Изд «Наукова думка, 2011.-275 с.
- [Згуровский М.З., 2013] Згуровский М.З. Зайченко Ю.П. Основы вычислительного интеллекта.-К.: Изд. «Наукова думка», 2013.- 406 с.
- [Ивахненко А.Г., 1976] Ивахненко А.Г., Зайченко Ю.П. Димитров В.Д.: Принятие решений на основе самоорганизации. – Москва, Сов. Радио. – 1976. – 363стр.
- [Минский М., 1971] Минский М., Пайперт С. Перцептроны. Пер. с англ. Под ред. В.А. Ковалевского. Киев. «Наукова Думка», 1971.
- [Рассел Стюарт, 2007] Рассел Стюарт и Норвиг Питер. Искусственный интеллект: современный подход, Второе издание.: Пер. с англ. М. : Изд. Дом «Вильямс» , 2007- 1408 с.
- [Розенблатт Ф., 1965] Розенблатт Ф. Принципы нейродинамики. Перцептроны и теория механизмов мозга.- М.: Мир, 1965.
- [Ф. Люгер, 2006] Люгер Ф. Искусственный интеллект. Пер. с англ. М.: Изд. Дом « Вильямс», 2006.
- [Хайкин Саймон, 2006] Хайкин Саймон. Нейронные сети: полный курс, 2-е изд. ,испр. Пер. сангл. М.: Изд. Дом Вильямс, 2006.- 1104 с.
- [Хокинс Джефф, 2007] Хокинс Джефф, Блейкли Сандра. Об интеллекте. Пер. с англ. –М.: ООО «ИД. Вильямс», 2007, 240с.

Authors' Information



Згуровский Михаил – ректор НТУУ «Киевский политехнический институт», директор Института прикладного системного анализа», академик НАН Украины, д.т.н., профессор. 03056, Киев-56, проспект Победы, 37, Украина. phone: 38044 -2366913 e-mail: mzgurovsky@gmail.com

Области научных исследований: системный анализ, модели и методы устойчивого развития, принятие решений



Зайченко Юрий - д.т.н., профессор ННК «Институт прикладного системного анализа», 03056, Киев-56, проспект Победы, 37, Украина phone: 38044 -4068393; e-mail: baskervil@voliacable.com,

Области научных исследований: теория принятия решений в условиях неопределенности, модели и методы вычислительного интеллекта в задачах прогнозирования и анализа в экономике и финансовой сфере, моделирование и оптимизация компьютерных сетей

System Analysis of Computational Intelligence Main Trends

Michael Zgurovsky, Yuriy Zaychenko

Abstract: The survey and analysis of new field in AI -computational intelligence (CI) is performed. The main components of CI- technologies, methods and applications are outlined, their interconnections are determined and analyzed. The interrelations between CI and soft-computing are established and discussed.

Keywords: computational intelligence, soft computing, neural networks, fuzzy logic, evolutionary modeling.