

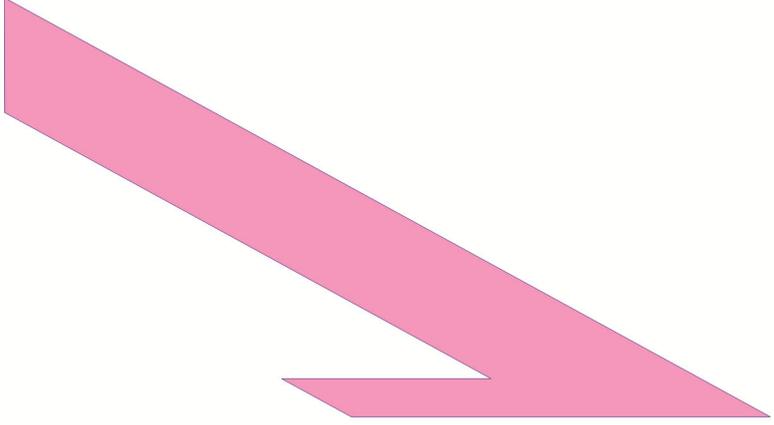


I T H E A



International Journal

**INFORMATION TECHNOLOGIES
&
KNOWLEDGE**



2012 Volume 6 Number 3



International Journal
INFORMATION TECHNOLOGIES & KNOWLEDGE

Volume 6 / 2012, Number 3

Editor in chief: **Krassimir Markov** (Bulgaria)

Victor Gladun (Ukraine)

Abdelmgeid Amin Ali	(Egypt)	Koen Vanhoof	(Belgium)
Adil Timofeev	(Russia)	Larissa Zaynutdinova	(Russia)
Aleksey Voloshin	(Ukraine)	Laura Ciocoiu	(Romania)
Alexander Kuzemin	(Ukraine)	Luis F. de Mingo	(Spain)
Alexander Lounev	(Russia)	Natalia Ivanova	(Russia)
Alexander Palagin	(Ukraine)	Nataliia Kussul	(Ukraine)
Alexey Petrovskiy	(Russia)	Nelly Maneva	(Bulgaria)
Alfredo Milani	(Italy)	Nikolay Lyutov	(Bulgaria)
Avram Eskenazi	(Bulgaria)	Orly Yadid-Pecht	(Israel)
Axel Lehmann	(Germany)	Radoslav Pavlov	(Bulgaria)
Darina Dicheva	(USA)	Rafael Yusupov	(Russia)
Ekaterina Solovyova	(Ukraine)	Rumyana Kirkova	(Bulgaria)
Eugene Nickolov	(Bulgaria)	Stefan Dodunekov	(Bulgaria)
George Totkov	(Bulgaria)	Stoyan Poryazov	(Bulgaria)
Hasmik Sahakyan	(Armenia)	Tatyana Gavrilova	(Russia)
Ilia Mitov	(Bulgaria)	Vadim Vagin	(Russia)
Irina Petrova	(Russia)	Vasil Sgurev	(Bulgaria)
Ivan Popchev	(Bulgaria)	Velina Slavova	(Bulgaria)
Jeanne Schreurs	(Belgium)	Vitaliy Lozovskiy	(Ukraine)
Juan Castellanos	(Spain)	Vladimir Ryazanov	(Russia)
Julita Vassileva	(Canada)	Martin P. Mintchev	(Canada)
Karola Witschurke	(Germany)	Zhili Sun	(UK)

International Journal "INFORMATION TECHNOLOGIES & KNOWLEDGE" (IJ ITK)
is official publisher of the scientific papers of the members of
the **ITHEA International Scientific Society**

IJ ITK rules for preparing the manuscripts are compulsory.

The **rules for the papers** for IJ ITK as well as the **subscription fees** are given on www.foibg.com.

Responsibility for papers published in IJ ITK belongs to authors.

General Sponsor of IJ ITK is the **Consortium FOI Bulgaria** (www.foibg.com).

International Journal "INFORMATION TECHNOLOGIES & KNOWLEDGE" Vol.6, Number 3, 2012
Edited by the **Institute of Information Theories and Applications FOI ITHEA**, Bulgaria, in collaboration with:
Institute of Mathematics and Informatics, BAS, Bulgaria,
V.M.Glushkov Institute of Cybernetics of NAS, Ukraine,
Universidad Politécnica de Madrid, Spain.

Printed in Bulgaria
Publisher ITHEA®

Sofia, 1000, P.O.B. 775, Bulgaria. www.ithea.org, www.foibg.com, e-mail: info@foibg.com
Technical editor: **Ina Markova**

Издател: ИТЕА (Регистрирана запазена марка на ЕТ ФОИ-КОМЕРС, София, България)
София 1000, ПК 775р България, www.ithea.org, e-mail: info@foibg.com

Copyright © 2012 All rights reserved for the publisher and all authors.

© 2007-2012 "Information Technologies and Knowledge" is a trademark of Krassimir Markov

ISSN 1313-0455 (printed)

ISSN 1313-048X (online)

ISSN 1313-0501 (CD/DVD)

ФОРМАЛИЗАЦИЯ ПРОБЛЕМЫ ИЗВЛЕЧЕНИЯ ЗНАНИЙ ИЗ ЕСТЕСТВЕННО ЯЗЫКОВЫХ ТЕКСТОВ

Александр Палагин, Сергей Крытый, Николай Петренко, Дмитрий Бибииков

Abstract: *Рассматривается формализация процесса анализа естественно-языковых текстов с целью извлечения знаний. Предлагается автоматизированный итеративный подход к реализации такого анализа.*

Keywords: *автоматизация обработки ЕЯТ, извлечение знаний, онтограф текстового документа.*

ACM Classification Keywords: *1.2 ARTIFICIAL INTELLIGENCE – 1.2.4 Knowledge Representation Formalisms and Methods.*

Введение

Проблема извлечения знаний из естественно-языковых текстов (ЕЯТ) является одной из главных проблем в исследованиях по искусственному интеллекту. Этой проблеме в последнее время уделяется большое внимание в основном из-за того, что потоки информации неуклонно возрастают и человек уже не в состоянии самостоятельно обрабатывать эту информацию. Это относится в первую очередь к информации, которая находится в текстах книг, разного рода электронных коллекциях, статей, газетах, Интернете и т. п. В связи с таким положением дел возникает необходимость в разработке средств автоматизации для анализа прежде всего ЕЯТ на предмет извлечения из них релевантной запросу пользователя информации или убедиться с их помощью, что такой информации в этих текстах нет. Создание таких средств наталкивается на сложность проблемы анализа, которая в свою очередь связана с семантической многозначностью значений предложений естественного языка. Кроме этой многозначности существует еще ряд вопросов, связанных с анализом эмоциональной окраски фраз, анализом фраз иронического и иносказательного характера, анализом при неполной информации (подразумевается нечто по умолчанию или вообще неизвестно) и т. п. Поскольку такого типа проблемы плохо формализуются, то отсюда и вытекают причины огромной сложности проведения такого анализа.

Краткая история проблемы

Проблема автоматизации процесса анализа ЕЯТ с целью извлечения знаний занимала многих исследователей. На первый план здесь выходит проблема формализации семантики естественного языка (ЕЯ) и такая попытка была предпринята еще в начале 30-х годов прошлого столетия в работах А. Тарского и его учеников. Однако, о такой необходимости говорили еще Аристотель, Лейбниц, Эйлер и др. В частности, Аристотель выделил четыре типа высказываний, которые были названы силлогизмами, а сам подход Аристотеля был назван силлогистикой [1, 7, 16]. Позже Эйлер изложил свое понимание силлогистики Аристотеля с помощью геометрической интерпретации его силлогизмов [16] (эту интерпретацию стали называть кругами Эйлера). Далее идеи Эйлера были развиты в работах

французского математика и астронома Ж.Д. Жергона, который ввел типы отношений и интерпретацию силлогистики Аристотеля в терминах этих отношений [19]. Он показал, что каждый тип силлогизма Аристотеля можно выразить в виде некоторых возможных вариантов таких отношений. Главная трудность в использовании жергоновых отношений состояла в том, что практически все их типы в сложном предложении требовали анализа большого числа вариантов.

Более существенные шаги на пути формализации ЕЯ были сделаны А. Тарским [2, 3], в результате которых появилось понятие выполнимости формул – понятия более общего, чем понятие истинности. Это понятие Тарский применил к открытым и замкнутым формулам (под замкнутой формулой в предложениях ЕЯ понимают фразу), что позволило сформулировать понятие истинности предложения ЕЯ и наложить ограничения на каждую открытую атомарную формулу, которая состоит из атомарного предиката. Поскольку таких формул имеется только конечное множество, то такой подход становится конструктивным.

Следующая попытка улучшения формализации А. Тарского, была предпринята Д. Дэвидсоном [4]. Он предложил добавить к понятию выполнимости и истины рекурсивное определение истины. В таком случае теория Т, которая включает рекурсивное определение истины, объясняет каким образом значения фраз зависят от значения (интерпретации) слов в этих фразах.

Монтегю [5] тоже верил в то, что методы формальной семантики можно применить к исследованию семантики ЕЯ. Но он, в отличие от Дэвидсона, отказался от применения логики предикатов первого порядка и предпочел категориальные грамматики. Эта грамматика включает в себя те категории, которые специалисты в области грамматик традиционно используют при определении ЕЯ, например, такие категории, как прилагательное или причастие. Это позволило Монтегю заменить понятие абсолютной истины понятием относительной истины в модели, потому что в разных моделях одно и то же предложение может иметь разные значения истинности. Такое расширение дало возможность определить понятие логической истинности и логического следствия для более широкого фрагмента ЕЯ. Таким образом, Монтегю выделил два элемента: интенцию (смысл) и экстенцию (денотат) и применил их к существительным, прилагательным и фразам. Существуют и другие подходы к анализу ЕЯ, которые базируются на понятиях семантической сети, фрейма и т. п. [18].

Недостатки и достоинства применения языка формальной логики к анализу предложений ЕЯ описаны в книгах [7] и [17], где исследуются рамки применения языков формальной логики к анализу предложений ЕЯ и указываются некоторые причины того, почему эти языки не всегда применимы к такого рода анализу. А результатом анализа предложений ЕЯ являются рассуждения (высказывания). Эти причины приводят к необходимости введения понятия достоверного и правдоподобного рассуждения (понятие правдоподобного рассуждения ввел Д. Пойа [6] и обосновал необходимость его использования в математике). Здесь под рассуждением понимается построение последовательности фактов, которые неизбежно приводят к принятию некоторого утверждения, являющегося целью рассуждения. Отметим, что понятие рассуждения существенно отличается от понятия логического вывода хотя бы тем, что оно может опираться на нелогические или металогические понятия как достоверного вывода, так и правдоподобного вывода.

Для последних десятилетий 20-го столетия в области исследования семантики ЕЯ характерными были попытки построения формальной теории семантики, которая была бы общей для естественных и

искусственных языков. Синтаксисом интересовались только в связи с семантикой, основной целью которой являлось объяснение понятий истины и логического следствия, а основной целью синтаксиса была характеристика синтаксических категорий, из которых строятся предложения.

Что касается данной работы, то она является продолжением исследований, начатых в работах [9-14] и связанных с проблемой анализа предложений ЕЯ с целью извлечения знаний.

Характеристика проблемы

Пусть заданы X – алфавит некоторого естественного языка, а $F(X)$ означает множество слов в алфавите X . Рассмотрим $L \subseteq F(X)$ – естественный язык в данном алфавите, предложения которого построены в соответствии с правилами грамматики P , где $P = \{p_i : i = 1, \dots, m\}$ – правила грамматики языка L . Правила грамматики определяют совокупность отношений

$$R_E = \{R_{p_i} : p_i \in P\},$$

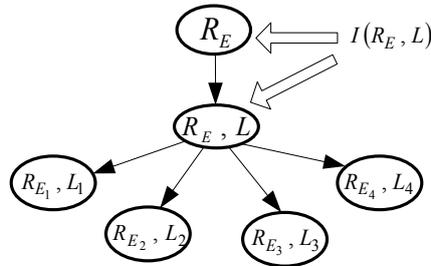
каждое из которых соответствует конкретному правилу грамматики. Пусть слова языка L разбиты в словаре этого языка на лексико-грамматические разряды с помощью лексико-грамматического отношения R [20–22]. Это означает, что в один класс попадают существительные, в другой класс – глаголы, в третий – прилагательные и т. д. Очевидно, что это отношение является отношением эквивалентности и в соответствии с этим отношением слова из L разбиваются на классы, элементы которых соответствуют лексико-грамматическим разрядам языка L , т. е. $L = L_1 \cup \dots \cup L_j \cup \dots \cup L_k$ – конечное множество лексико-грамматических разрядов в языке L .

Пусть L_i – некоторый класс этого разбиения. Слова, которые входят в L_i структурируются в соответствии с лингвистическими и семантическими отношениями языка L , являющимися отношениями частичного порядка или квазипорядка (например, гипоним-гипероним, мероним-голоним или род-вид, целое-часть, класс-элемент, вышестоящий-нижестоящий, класс-подкласс). Отношения частичного порядка задаются в виде ориентированного или неориентированного графа $G = (V, E)$, где $V = L_i$, а E – множество пар слов (p, q) , $p, q \in L_i$ таких, что p доминирует над q по одному из указанных отношений. Такой граф называют онтографом (в математике такие графы называются диаграммами Хассе).

Заметим, что хотя в определении вершин онтографа использовались разряды $V = L_i$, на самом деле вершинами являются классы эквивалентных между собой слов из L_i относительно глобального отношения синонимии RS . Таким образом под $V = L_i$ следует понимать фактор множество $V = L_i / R_S$. В таком понимании онтограф $G = (V, E)$ представляет собой гиперграф, вершины которого соответствуют классам синонимичных слов языка. Далее под онтографом будем понимать именно такого типа гиперграф.

При конкретизация языка L и проблемы, которая исследуется, выполняется конкретизация отношений и построение соответствующего онтографа для данного естественного (украинского, русского, английского)

языка. Чаще всего в работах на эту тему рассматриваются следующие четыре разряда языка L : L_1 – существительные, L_2 – глаголы, L_3 – прилагательные и L_4 – наречия. В соответствии с такой семантической интерпретацией классов L_i как лексико-грамматических разрядов существительного, глагола, прилагательного и наречия ($\{L_1, L_2, L_3, L_4\}$) онтограф языково-онтологической картины мира (ЯОКМ) строится в соответствии со следующей схемой:



В данной работе нас будут интересовать отношения из $R_E = \{R_{p_i} : p_i \in P\}$ и их интерпретация $I(R_E, L)$. Отношения из R_E определяют синтаксические правила построения предложений языка L , но при анализе предложений такого языка на первый план выступают семантические отношения $I(R_E, L)$, как интерпретация отношений из R_E , поскольку синтаксически правильные предложения могут быть абсолютно бессмысленными с точки зрения здравого смысла.

Возникает вопрос: как определить семантические отношения $I(R_E, L)$ на синтаксически правильных предложениях языка L , т.е. предложениях, принадлежащих к отношениям из R_E ?

В связи с тем, что язык является естественным, то однозначного способа определения семантики, т.е. интерпретации предложений этого языка, дать невозможно. По видимому в этом и состоит главная трудность в решении проблемы анализа предложений естественного языка с целью извлечения знаний. Это обстоятельство дает определенную свободу в определении такой семантики, которая налагает определенные обязательства. Мы примем следующее определение семантики.

Определение 1. Синтаксически правильное предложение s языка L будем называть семантически правильным, если оно является фактически истинным, или достоверным, или правдоподобным с точки зрения здравого смысла группы индивидуумов или отдельно взятого индивидуума.

Понятно, что приведенное определение, не является строгим, но оно, в определенной степени, соответствует естественному положению дел. Приведем краткие комментарии понятий, фигурирующих в этом определении.

Под **фактически истинным** предложением понимается умозаключение, факты которого адекватно интерпретируются (т.е. имеют единственно возможное логическое значение) в языке формальной логики (логики высказываний, предикатов, модальной и т.п.) или получено из таких фактов по одному из формальных правил вывода в этой логике. Если удастся извлечь фактически истинные знания из предложений ЕЯ, то это позволяет полностью решить логические проблемы анализа.

Пример 1. Примерами фактически истинных предложений являются следующие высказывания: «Земля круглая и вращается вокруг Солнца», «Г.Ф. Вороной – украинский математик», «число 23 – простое».

А для иллюстрации анализа такого типа предложений, рассмотрим следующий текст: «*Жители сельской местности будут голосовать за кандидатуру президента (ГП) только в том случае, если он подпишет закон о праве сельских жителей на землю (ЗЗ). Рабочие заводов и олигархи (РО) не будут голосовать за кандидатуру президента, если он не наложит вето на этот закон (ВЗ). Очевидно, что президент либо подпишет закон или наложит на него вето. Следовательно, он либо потеряет голоса жителей сельских регионов, либо голоса рабочих и олигархов.*»

Анализ данного фрагмента текста допускает следующую интерпретацию в языке логики высказываний:

Дано: ГП \rightarrow ЗЗ, РО \rightarrow ВЗ, \neg ЗЗ \vee \neg ВЗ. Верно ли, что из этих фактов следует \neg ГП \vee \neg РО.

Методом резолюций справедливость этого следствия доказывается в два шага. Δ

Под **достоверным предложением** понимается такой факт, который подтверждается имеющимся опытом или знаниями либо группы людей либо отдельного взятого человека или следует из кажущихся им истинных предположений по какому-нибудь правилу умозаключения. При исследовании на непротиворечивость умозаключений, извлеченных из достоверных предложений, в широком смысле наиболее характерен дедуктивный вывод. Отсюда следует, что фактические (логические) следствия в языке формальной логики являются очень частным случаем достоверных умозаключений. Следует заметить, что дедукция «в высшей степени идеализированная и ограниченная форма рассуждений» [8], которая применима только в очень узких рамках к моделированию и анализу предложений ЕЯ, поскольку не полностью отражает такие понятия как здравый смысл, неопределенность, достоверность информации и т. п.

Далее, понимание разными людьми смысла одного и того предложения может быть разным, а отсюда вытекает, что точку зрения группы людей или отдельно взятого человека необходимо учитывать (кстати, разное понимание одного и того же предложения (высказывания, рассуждения) является источником возникновения *дискуссии*). В действительности мы очень редко пользуемся абсолютно достоверными фактами, поскольку многообразие мира, описываемое этими фактами, нельзя ограничить и втиснуть в какие-нибудь формальные рамки. Поэтому в процессе рассуждений в большинстве случаев мы оперируем, *опираясь на правдоподобные факты, а не на достоверные факты*. Это связано с тем, что часто при принятии решения о чем-нибудь, мы прибегаем к наблюдению и опыту. А это нам дает только правдоподобные факты, которые потом должны проверяться и доказываться, и только после этого приниматься или опровергаться.

Под **правдоподобным предложением** будем понимать такой факт, истинность которого опирается на кажущиеся правильными (истинными) умозаключения, с точки зрения группы людей или отдельно взятого человека. Для правдоподобных рассуждений характерным является индуктивный вывод и вывод по аналогии [6].

Отметим также, что многие исследователи этой проблемы также считают, что моделирование знаний, извлеченных из ЕЯТ, не может ограничиваться формализацией только лишь непогрешимого интеллекта. Естественным основанием для такого мнения является то, что важной чертой естественного интеллекта есть способность вырабатывать здравые рассуждения, которые могут оказаться и недостоверными. В случае неполной, неточной и изменчивой информации наши рассуждения часто становятся только

предположительными, а в следствии этого лишь правдоподобными и поэтому могут подлежать пересмотру и уточнению (модификации).

Другие исследователи считают, что проблема анализа ЕЯТ решена, если извлеченные знания представлены в базе знаний и все проблемы по их анализу решаются средствами баз знаний. Это мнение, с нашей точки зрения, не совсем соответствует реальному состоянию дел. Главной проблемой при обработке знаний в базах знаний является та же модифицируемость знаний. Модификация знаний необходима по многим причинам. В частности, в самой базе знаний объекты могут представляться интенционально (аксиоматически) или экстенционально (перечнем элементов), что вносит свои коррективы в процесс их обработки. А в общем случае различают два основных типа модифицируемых знаний внешнего характера: *предположительные* и *предполагаемо полные*.

Предположительные знания являются всего лишь правдоподобными. Это связано с тем, что они неточные, поскольку базируются на неполной, неточной и изменчивой информации, а также по причине их естественной неточности и модифицируемости. Примерами такого типа знаний являются рассуждения по умолчанию, рассуждения с прототипами и знания статистического характера. Неполнота знаний является естественной, поскольку в повседневной жизни мы часто общаемся путем молчаливо подразумеваемых фактов. Такого типа факты в силлогистике именуется *антимемами*. Антимемы неизбежны потому, что они существенно ускоряют процесс обмена мыслями между людьми и без них этот процесс сделался бы невыносимо скучным.

Пример 2. Рассмотрим такой текст: «Уважаемая госпожа Хадсон, если я за ужином выпью крепкий напиток вроде виски (ВВ), то я не смогу заснуть (З) всю ночь. Поэтому, с Вашего разрешения, я за ужином не буду пить виски.»

Это пример антимемы, которая выглядит таким образом: из $ВВ \rightarrow \neg З$ вытекает $\neg ВВ$. Очевидно, что пополненное пропущенным фактом рассуждение выглядит так: из $ВВ \rightarrow \neg З, З$ вытекает $\neg ВВ$. Δ

Предполагаемо полные знания – это знания, которые основаны на фактах, которые предполагаются информационно полными, но которые таковыми не являются или перестают быть таковыми. В действительности часто встречается ситуация, когда выдвигаются модифицируемые (а иногда и неявные) соглашения, для наращивания наших знаний в условиях неполной или неизвестной информации. Основываясь на таких знаниях наши выводы могут быть логически корректными по отношению к этим добавленным знаниям. Однако, эти рассуждения оказываются модифицируемыми, так как они основываются на изменчивом состоянии знаний. Следует отметить, что некоторые корректные формы вывода, которые формализует классическая математическая логика, также могут оказаться модифицируемыми. Это объясняется тем, что они применяются к базе знаний, которая зачастую пополняется всего лишь правдоподобными знаниями. Например, знания, занесенные одним исследователем в базу знаний, могут быть неправильно или неточно поняты другими и поэтому будут подвергаться модификации другими исследователями.

Формальная постановка задачи

В связи с проблемой, которая нас интересует, необходимо определить понятия «знание» и «процесс извлечения знаний» из предложений ЕЯ¹. Строгого определения понятия «знание» не существует, однако это понятие вызывало большой интерес ученых, начиная с древних греков. Его изучали Платон и Аристотель, которые ввели еще целый ряд понятий, характеризующих знание: «рассудок», «мнение», «математическое мышление» и др. В 20-м столетии в связи с развитием такой области как программирование появились понятия «процедурное» и «декларативное» знание. Процедурное знание содержит в себе информацию о том, как нужно действовать, чтобы получить нужный результат, а декларативное знание содержит в себе информацию о том, над чем надо выполнять эти действия. В частности, в области искусственного интеллекта закрепилось и употребляется следующее определение понятия «знание». **Знание – это обоснованное истинное убеждение (вера)**. Это определение для решения нашей задачи мало что дает, поскольку не очерчивает материальный объект.

В целях более точной формулировки понятий «знание» и «извлечение знаний», которыми будем пользоваться в этой работе, рассмотрим следующие определения, пользуясь нотацией констрейнтного программирования [15].

Пусть дано некоторое множество D , на котором определена конечная совокупность $R = \{R_1, \dots, R_k\}$ отношений $R_i \subseteq D^n$, $i = 1, 2, \dots, k$, конечной арности. Языком ограничений L на D называется непустое множество $L \subseteq R$. Проблема выполнимости ограничений из L формулируется следующим образом.

Определение 2. Для произвольного множества D и языка ограничений L на D проблемой выполнимости ограничений $CSP(L)$ является решение такой комбинаторной задачи:

дана тройка $P = (V, D, C)$, где

- $V = \{v_1, \dots, v_m\}$ - конечное множество переменных;
- $C = \{c_1, \dots, c_q\}$ - конечное множество ограничений, где ограничение c_i из C - пара (s_i, R_i) , где $s_i = (v_{i_1}, \dots, v_{i_j})$ - кортеж, состоящий из переменных, $R_i \in L - n_j$ -арное отношение на D ;

найти функцию $\varphi: V \rightarrow D$ такую, что $\forall (s_i, R_i) \in C$ кортеж $(\varphi(v_{i_1}), \dots, \varphi(v_{i_j})) \in R_i$ либо убедится в том, что ее не существует, $i = 1, 2, \dots, q$. Множество D в этом случае называется областью проблемы, а функция φ называется интерпретацией $CSP(L)$.

Применительно к анализу предложений ЕЯ с целью извлечения знаний множество D интерпретируется как множество объектов (сущностей), извлеченных из предложений входного текста T , удовлетворяющих отношениям из $R_E = \{R_{p_i} : p_i \in P\}$, которое факторизовано по некоторому отношению эквивалентности R_S^s (это отношение представлено в онтографе вершинами синонимичных объектов, которые факторизуются по предметно-ориентированному отношению синонимии). Переменные из

¹ Один из вариантов таких определений мы привели в [14].

множества $V = \{v_1, v_2, \dots, v_m\}$ принимают свои значения в этом факторизованном множестве объектов D , фигурирующих в тексте T (это могут быть более широкие лексико-грамматические разряды, такие, как конкретные личности, даты, конкретные предметы и т. п.). А в качестве $\varphi: V \rightarrow D$ выступает интерпретация $I(R_E, L)$, в результате которой появляются отношения (предикаты) $\{\phi_1, \phi_2, \dots, \phi_m\}$.

Отношения $\{\phi_1, \phi_2, \dots, \phi_m\}$ из $I(R_E, L)$, извлеченные из текста T ЕЯ, будем называть **знаниями**.

Это определение, по нашему мнению, уточняет данное выше определение знания в том смысле, что оно материализует объект поиска и механизм этого поиска.

Пример 3. Пусть имеем такой текст: «*Гуляя набережной Черного моря писатель заметил, что облака (О) подобны пляшущим сатирам (С). Но присмотревшись, он понял, что эти сатиры плывут, а не пляшут*».

В этом тексте множество $D = \{\text{набережная Черного моря (Н(м,ч)), писатель (П), облака (О), сатиры (С)}\}$ $R = \{\text{гулять, замечать, понимать, плыть, плясать, присматриваться, подобные}\}$. В данном случае Н(м,ч) означает объект **набережная**, а **(м,ч)** – атрибуты (ограничения) этого объекта.

Из первого предложения получаем такие отношения:

гулять(П, Н(м,ч)), замечать(П,О), подобные(О,С), плясать(С).

Из второго предложения получаем такие отношения:

присматриваться(П,С), понимать(П), плыть(С), \neg плясать(С).

Отношение *подобные(О,С)* факторизует множество D на классы по объектно-ориентированному отношению синонимии. Заметим, что глобальное отношение синонимии не внесет объекты «сатиры» и «облака» в один класс синонимичных объектов, в то время как предметно-ориентированное отношение внесет объекты «облака» и «сатиры» в один класс синонимии (по отношению подобия). Таким образом, факторизованное множество D принимает вид: $D = \{\text{Н(м,ч), (П), \{О,С\}}\}$. Эта факторизация позволяет все высказывания относительно С считать аналогичными высказываниями относительно О и уточнить полученные отношения следующим образом:

$\text{гулять(П, Н(м,ч))} \wedge \text{замечать(П,О), плясать(С)} \leftrightarrow \text{плясать(О)}$

для первого предложения и

$(\text{присматриваться(П,С)} \wedge \text{понимать(П)}) \rightarrow (\text{плыть(С)} \wedge \neg \text{плясать(С)})$

для второго предложения.

Соединяя эти факты вместе и выполняя эквивалентные преобразования, получаем

$\text{гулять(П, Н(м,ч))} \wedge \text{замечать(П,О),} \quad (\text{присматриваться(П,С)} \wedge \text{понимать(П)}) \rightarrow \text{плыть(С)},$

$(\text{присматриваться(П,С)} \wedge \text{понимать(П)}) \rightarrow \neg \text{плясать(С)}, \quad \text{плыть(С)} \leftrightarrow \text{плыть(О)}.$

В процессе логического анализа этих фактов (вследствие эквивалентных преобразований) возникает необходимость модификации извлеченных знаний.

Заметим, что этот текст описывает некоторое поэтическое (метафорическое) восприятие наблюдаемых фактов, которые в действительности не имеют места. Δ

Автоматизированный итеративный метод анализа предложений ЕЯ

Исходя из выше сказанного, можно предложить такой итеративный способ автоматизированной обработки ЕЯТ [13].

Шаг 1. Морфологический анализ заданного текста T с целью построения словаря для текста T и разбиения на классы $\{L_1, L_2, L_3, L_4\}$ (или более мелкого разбиения, включающего и другие части речи). Кроме того, на этом шаге вычисляется парадигма всех словоформ изменяемых частей речи и исходная лексема, выделение отглагольных существительных и др.

Шаг 2. Построение множества объектов D , исходя из результатов синтаксического анализа текста T и результатов шага 1. Кроме того, на этом шаге вычисляются многословные термины, анафорические связи, антимемы и т. п.

Шаг 3. Построение онтографа, исходя из множества объектов D (построение отношения R_S) на классах $\{L_1, L_2, L_3, L_4\}$. Онтограф текста строится на основе онтографов предложений применением правил конъюнкции и упрощения, алгоритмы применения которых описаны в [14.]

Шаг 4. Построение интерпретации $I(R_E, L)$ на множестве объектов D , исходя из онтографа и предметно-ориентированного отношения синонимии на D .

Шаг 5. Внесение полученных на шаге 4 отношений $\{\phi_1, \phi_2, \dots, \phi_m\}$ в базу знаний.

Шаг 6. Выполнить анализ множества отношений $\{\phi_1, \phi_2, \dots, \phi_m\}$ средствами базы знаний.

Шаг 7. Если результаты анализа удовлетворяют пользователя, то закончить процесс иначе выполнить уточнение множества D и интерпретации $I(R_E, L)$ и перейти на шаг 3.

В приведенной последовательности шагов многие из них требуют комментариев. Заметим, что первые три шага детально изучались многими исследователями и для их реализации имеются соответствующие средства, работающие в автоматическом или автоматизированном режиме [10, 17, 20–22].

Наиболее проблемными являются шаги 4 и 7, что является следствием неформального определения семантически правильного предложения. На шаге 4 предполагается такая обобщенная схема взаимосвязей структурных компонент текста T , которая следует из семантической интерпретации соответствующих частей речи:

- объекты – это существительные;
- отношения (предикаты) – это глаголы;
- атрибуты объектов – это прилагательные (ограничения на объекты);
- атрибуты отношений (предикатов) – это наречия (ограничения на предикаты).

Такая интерпретация согласовывается с определением 2, со структурой предложений текста T и с другими известными концепциями (в частности, с концепцией системы WordNet).

Этот шаг, по-видимому, необходимо выполнять в автоматизированном (например, в диалоговом) режиме, консультируясь с пользователем, являющимся (или с пользователями являющимися) авторами текста T или экспертами в той предметной области, к которой относится данный текст T . На этом шаге сначала определяются имена отношений и их арности, которые связываются прежде всего с глаголами (как в вышеприведенном примере 3 имена отношений ПЛЯШУТ и ПЛЫВУТ и их арность 1). Затем, полученные

таким образом отношения, уточняются в процессе взаимодействия с пользователем. Если такое уточнение выполнено, то осуществляется переход на шаг 5.

Шаги 5 и 6 детально комментировать нет необходимости, поскольку представляется понятным, что на этих шагах должно выполняться. Извлеченные из текста знания заносятся в базу знаний таким образом, чтобы можно было эффективным способом проводить их анализ. Выбор способа представления знаний в такой базе зависит от того, какие алгоритмы вывода будут использованы. Что касается такой обработки извлеченных из текста знаний в базе знаний, то о методах их представления, обработки и анализа свойств мы отсылаем читателя к монографии [8], где описаны основные методы и средства различных типов вывода.

Процесс выполнения шага 7 заключается в том, что если результаты анализа в базе знаний удовлетворяют пользователя, или подтверждают факты, полученные опытным путем, или соответствуют ожидаемым результатам, то дальнейший анализ можно не проводить и закончить работу алгоритма. В противном случае, если результаты анализа приводят к противоречиям, или являются абсурдными с точки зрения здравого смысла, или носят недостоверный характер, или не правдоподобны, то необходимо сделать повторный анализ семантических отношений, присутствующих в тексте, сделать необходимые уточнения или другие предположения и выполнить соответствующую модификацию, после чего повторить шаги 3 – 7. В процессе повторного анализа необходимо убедиться в правильности построенных семантических отношений, правильности дополнительных предположений, правильности трактовки некоторых понятий, объектов и отношений между этими объектами с точки зрения здравого смысла, если не удается эти отношения проинтерпретировать в языке математической логики.

Пример работы алгоритма

Рассмотрим пример работы первых трех шагов вышеприведенного алгоритма, поскольку продемонстрировать работу всех шагов нет возможности. С этой целью проанализируем такой текст (анализируемый текст написан на украинском языке, в связи с тем, что в данной версии в системе функционирует только толковый словарь украинского языка):

«В сучасному розумінні, комп'ютер – це універсальний електронний пристрій, призначений для автоматизації накопичення, збереження, опрацювання, передачі та відтворення даних.

Структурно комп'ютер складається з чотирьох основних пристроїв відповідно до тих завдань, які вони вирішують при опрацюванні даних.

Пристрої вводу призначені для вводу (накопичення) інформації та управління роботою комп'ютера користувачем. До цих пристроїв відносяться: клавіатура, маніпулятор миша, сканер, джойстик.

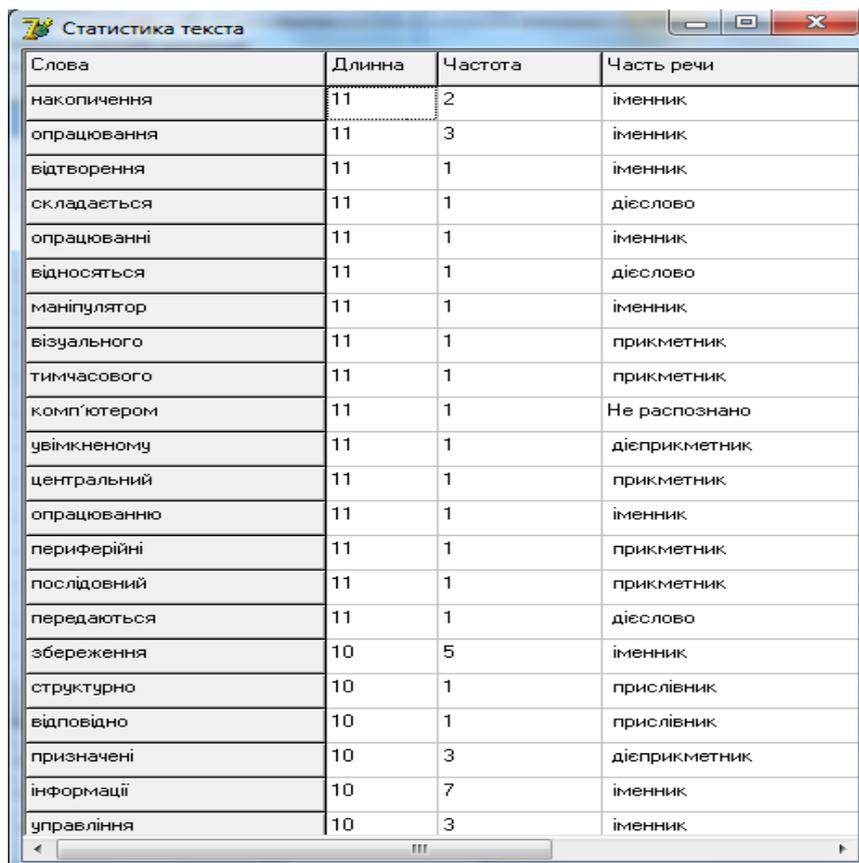
Пристрої виводу призначені для виводу інформації з метою візуального спостереження за роботою комп'ютера та створення твердих копій документів. До них належать монітор, принтер (друкуючий пристрій), плоттер.

Запам'ятовуючі пристрої призначені для збереження інформації, як тривалого, так і тимчасового, на час їх опрацювання комп'ютером. Пам'ять поділяють на внутрішню та зовнішню. Внутрішня призначена для збереження інформації під час роботи комп'ютера. Сюди відносять оперативну, постійну, кеш та CMOS-пам'ять. Вміст цієї пам'яті, як правило, зберігається лише при увімкненому живленні. Зовнішня пам'ять

призначена для тривалого збереження інформації незалежно від того, чи є живлення. Для зовнішньої пам'яті розрізняють пристрої пам'яті (накопичувачі) та носії даних (дискети, диски, магнітні стрічки, лазерні диски, тощо). Пристрій опрацювання інформації здійснює її переробку та загальне управління роботою всіх інших пристроїв. Цим пристроєм є процесор.

Головним пристроєм комп'ютера є центральний процесор. Він і виконує основні операції по опрацюванню даних та управління роботою інших пристроїв. По способу розташування пристроїв відносно процесора їх поділяють на внутрішні та зовнішні (периферійні). До внутрішніх відносять деякі види пам'яті. Зовнішніми є пристрої вводу-виводу інформації, пристрої для її тривалого збереження. Узгодженість роботи окремих пристроїв здійснюють апаратні інтерфейси. Їх поділяють на послідовні та паралельні. Через послідовний інтерфейс дані передаються по одному біту. Вони прості за будовою, не вимагають синхронізації роботи передавача та приймача даних. Однак пропускна здатність їх менша, а коефіцієнт корисної дії нижчий. Їх використовують для підключення "повільних" пристроїв – наприклад, різноманітних датчиків».

Результат роботи системи аналізу цього тексту на шаге 1 показан (частично) на рис. 1 (снимок екрана комп'ютера), где приведені елементи класов существительных (іменники), глаголов (дієслова), прилагательных (прикметники) и причастий (прислівники).



Слова	Длина	Частота	Часть речи
накопичення	11	2	іменник
опрацювання	11	3	іменник
відтворення	11	1	іменник
складається	11	1	дієслово
опрацюванні	11	1	іменник
відносяться	11	1	дієслово
маніпулятор	11	1	іменник
візуального	11	1	прикметник
тимчасового	11	1	прикметник
комп'ютером	11	1	Не распознано
увімкненому	11	1	дієприкметник
центральний	11	1	прикметник
опрацюванню	11	1	іменник
периферійні	11	1	прикметник
послідовний	11	1	прикметник
передаються	11	1	дієслово
збереження	10	5	іменник
структурно	10	1	прислівник
відповідно	10	1	прислівник
призначені	10	3	дієприкметник
інформації	10	7	іменник
управління	10	3	іменник

Рис. 1. Результат работы системы анализа текста на шаге 1

Далее, система также выдает номер предложения и номер вхождения данного слова в тексте. Эта информация используется на шаге 4 при построении и интерпретации предложений и извлечения

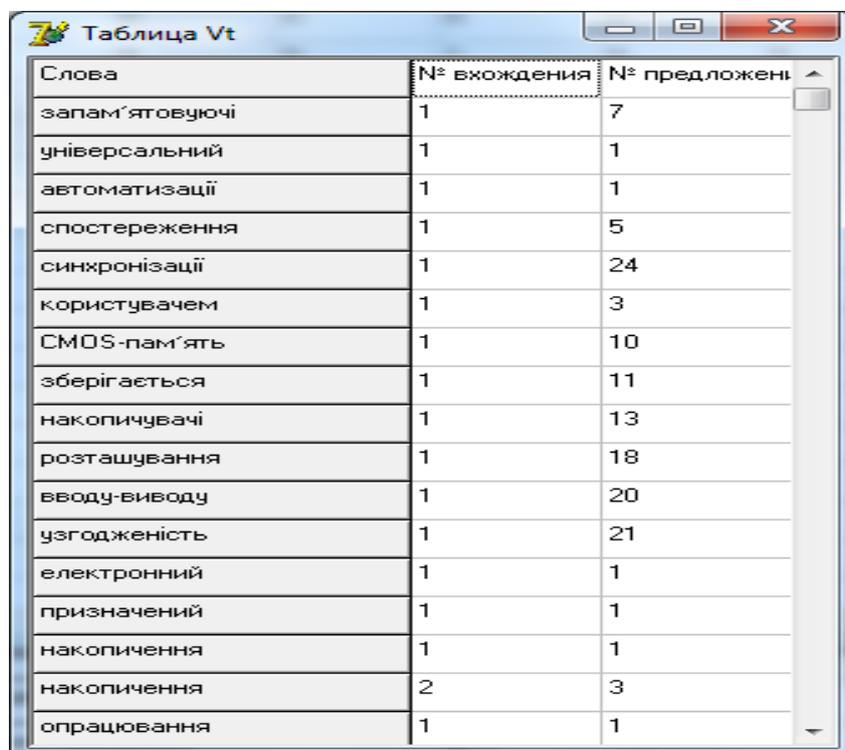
первичных знаний из текста и, в частности, при определении арностей извлекаемых предикатов (отношений).

Из приведенных рисунков видно, как очерчивается область интерпретации D и классы разбиения слов текста на классы $\{L_1, L_2, L_3, L_4\}$. В данном случае $L_1 = \{\dots\text{накопичення, опрацювання, маніпулятор, збереження, інформації, управління,}\dots\}$, $L_2 = \{\dots\text{складається, відносяться, передаються,}\dots\}$.

Система также выдает частоту появления слов в тексте и длину этих слов (рис. 2), что позволяет выполнить частотный анализ исходного текста с целью идентификации плагиата или чего-либо другого.

Кроме того, результат построения классов $L_1 - L_6$ для первых трех предложений анализируемого текста представлен в таблице ниже. А на рис. 3 – рис. 5 представлены онтографы предложений 1–4.

Для рассмотренного примера множество R_s состоит из элементов $\{R_1$ (эквивалентность), R_2 (мероним-гипероним), R_3 (назначение), R_4 (тема), R_5 (обрабатывать), R_6 (принимать решение)}.



Слова	N° вхождения	N° предложени
запам'ятовуючі	1	7
універсальний	1	1
автоматизації	1	1
спостереження	1	5
синхронізації	1	24
користувачем	1	3
СМОС-пам'ять	1	10
зберігається	1	11
накопичувачі	1	13
розташування	1	18
вводу-виводу	1	20
узгодженість	1	21
електронний	1	1
призначений	1	1
накопичення	1	1
накопичення	2	3
опрацювання	1	1

Рис. 2. Результат работы системы анализа текста на шаге 2

Заключение

Описанные в данной работе понятия и подходы к автоматизации обработки ЕЯТ составляют основу как теоретического, так практического анализа извлечения знаний из ЕЯТ. Данный подход используется в Институте кибернетики им. В.М. Глушкова НАН Украины и Киевском национальном университете им. Т. Шевченко в экспериментальной системе автоматизации анализа ЯЕТ с целью извлечения знаний в рамках онто-логического подхода к представлению и обработки информации. Используя эту основу и, прежде всего ее реализацию, предполагается наращивание ее мощности за счет построения новых метаотношений над построенными отношениями, являющимися отдельными частями знаний в исследуемом тексте, а также анализа текстов, написанных на русском языке.

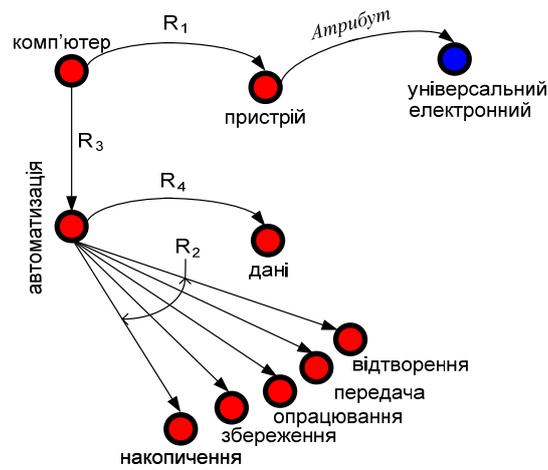


Рис. 3. Онтограф лексики предложения 1

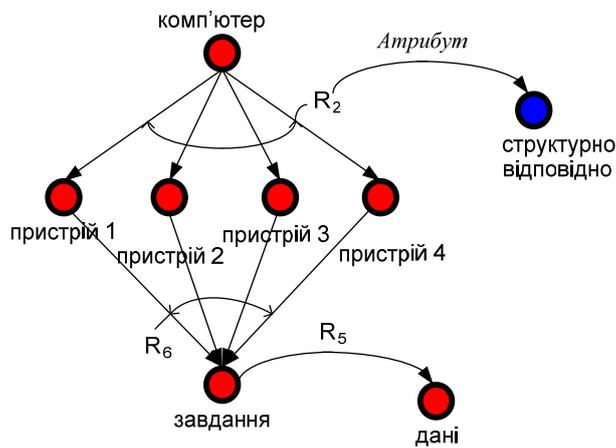


Рис. 4. Онтограф лексики предложения 2

Речення 3-4

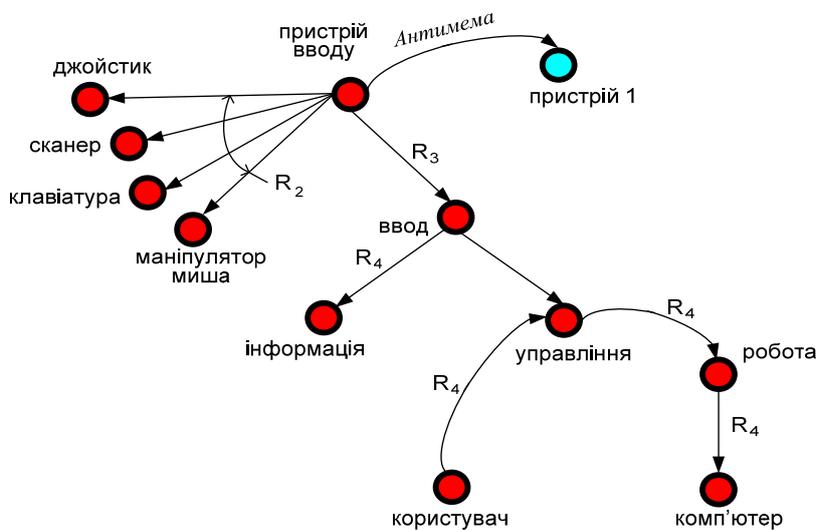


Рис. 5. Онтограф лексики предложений 3 и 4

Литература

1. Колмогоров А.Н., Драгалин А. Г. Введение в математическую логику. – Издательство Московского университета. – 1982. – 118 с.
2. Tarski A. The semantic conception of truth. *Philosophy and phenomenological Research*. – v.4. – 1944. – P. 241–375.
3. Tarski A. *Logique, Semantique and Metamathematique (1923-1944)*. Colin. – Paris. – 1972.
4. Davidson D. *Proceedings of Philosophical Logic*. – Reidel. – Dordrecht. – 1969.
5. Montague R. *Universal grammars. Theoria. Formal Phylisophy: Selected Papers of R. Montague*. – Yale University Press.– 1974. – vol. 36. – P. 222–246.
6. Пойа Д. Математика и правдоподобные рассуждения. – М.: Наука.–1975.– 462 с.
7. Клини С. Математическая логика. – М.: Мир. – 1973. – 480 с.
8. Вагин В. А., Головина Е. Ю., Загорянская А. А., Фомина М. В. Достоверный и правдоподобный вывод. – М.: Физматлит. – 2004. – 703 с.
9. Палагин А. В., Крывый С. Л., Петренко Н. Г., Знание-ориентированные информационные системы с обработкой естественно-языковых объектов: основы методологии и архитектурно-структурная организация. – ж. УСиМ. – 2009. – №3. – С. 42–55.
10. Палагин А. В., Крывый С. Л., Бибииков Д. С. Обработка предложений естественного языка с использованием словарей и частоты появления слов. – *Natural and Artificial Intelligence Intern. Book Series. – Intelligent Processing. – ITNEA. – Sofia. – N 9. – 2010. – P. 44–52.*
11. Палагін О. В., Кривий С. Л., Петренко М. Г., Бібіков Д. С. Алгебро-логічний підхід до аналізу та обробки текстової інформації. – ж. «Проблемы программирования». – 2010. № 2–3. – С. 318–329.
12. Палагін О. В., Кривий С. Л., Бібіков Д. С., Величко В. Ю., К. Марков, К. Иванова, І. Мітов Формально-логічний підхід до побудови системи аналізу знань в різних предметних областях. – ж. «Проблемы программирования». – 2010. – № 2–3. – С. 382–389.
13. Крывый С. Л. Бибииков Д. С. Итеративный подход к анализу естественно-языковых текстов: логический аспект. – ж. «Проблемы программирования». – 2012. – № 2–3. – С. 318–329.
14. Палагин А. В. Онтологические методы и средства обработки предметных знаний / А. В. Палагин, С. Л. Крывый, Н. Г. Петренко. – [монография] – Луганск: изд-во ВНУ им. В. Даля, 2012. – 324 с.
15. Cohen D., Jeavons P. The Complexity of Constraint Languages. In "Handbook of Constraint Programming. – Edited by F. Rossi, P. van Beek and T. Walsh. – 2006. – P. 245 – 280.
16. Кулик Б. А. Логика естественных рассуждений. – С.-Петербург: Невский диалект. – 2001. – 127 с.
17. Рубашкин В.Ш. Представление и анализ смысла в информационных системах. – М.: Наука. – 1989. – 188 с.
18. Тейз А., Грибомон П., Луи Ж. и др. Логический подход к искусственному интеллекту. От классической логики к логическому программированию. – М.: Мир. – 1990. – 429 с.
19. Тейз А., Грибомон П., Юлен Г и др. Логический подход к искусственному интеллекту. От модальной логики к логике баз данных. – М.: Мир. – 1998. – 494 с.

20. Леонтьева Н.Н., Семенова С.Ю. Семантический словарь РУСПАН как инструментарий компьютерного понимания. – М.: МГГИИ. – 2003. – С. 41–46.

21. Леонтьева Н.Н. К теории автоматического понимания естественных текстов. Часть 1. Моделирование системы "мягкого понимания" текста: информационно-лингвистическая модель. – М., МГУ, 2000. – 43 с.

22. Леонтьева Н.Н. К теории автоматического понимания естественных текстов. Часть 2. Семантические словари: состав, структура, методика создания. – М., МГУ, 2001. – 41 с.

Информация об авторах



Александр Палагин – Ин-т кибернетики им. В.М. Глушкова НАН Украины, Киев-187 ГСП, 03680, просп. акад. Глушкова, 40; e-mail: palagin_a@ukr.net

Область исследований: Общая теория знание-ориентированных информационных систем



Сергей Крывый – Киевский национальный университет им. Т. Шевченко, Украина Киев-187, ГСП, 03680, просп. акад. Глушкова, 40; email: krivoi@i.com.ua

Область исследований: Дискретная математика, теория автоматов, прикладная математическая логика, верификация программного обеспечения, программирование с ограничениями, онтологии.



Николай Петренко – Ин-т кибернетики им. В.М. Глушкова НАН Украины, Киев-187 ГСП, 03680, просп. акад. Глушкова, 40; e-mail: petrng@ukr.net

Область исследований: Методология и инструментальные средства автоматизированного проектирования онтологий предметных областей, системная интеграция междисциплинарных научных знаний



Дмитрий Бибиков – Ин-т кибернетики им. В.М. Глушкова НАН Украины, Киев-187 ГСП, 03680, просп. акад. Глушкова, 40; e-mail: gbbcoff@gmail.com

Область исследований: Искусственный интеллект, автоматизация поиска доказательств в формальных логических языках

КОМПРОМИССНОЕ РЕШЕНИЕ ЗАДАЧ УСЛОВНОЙ ОПТИМИЗАЦИИ

Альберт Воронин

Аннотация: Рассмотрена возможность получения компромиссного решения в задачах условной оптимизации. Проблема состоит в том, чтобы полученное решение отражало компромисс между противоречивыми требованиями экстремизации целевой функции и выполнения ограничений. Для решения рассматриваемой проблемы предпринимается подход многокритериальной оптимизации с применением нелинейной схемы компромиссов. Приведен модельный пример.

Ключевые слова: нелинейное программирование, целевая функция, многокритериальная оптимизация, ограничения, нелинейная схема компромиссов.

ACM Classification Keywords: H.1 Models and Principles – H.1.1 – Systems and Information Theory; H.4.2 – Types of Systems.

Содержание проблемы

Задачи оптимизации в различных предметных областях обычно сводятся к поиску экстремума целевой функции, ограниченной условиями, наложенными на аргументы оптимизации. Для решения таких задач привлекается аппарат математического программирования. Изложим кратко некоторые основные понятия из этой области.

Математическое программирование – это математическая дисциплина, в которой разрабатываются методы отыскания экстремальных значений целевой функции $f(x)$ среди множества ее возможных значений, определяемых ограничениями $\psi(x)$.

Достаточно общая задача математического программирования формулируется так:

– найти оптимальное сочетание аргументов оптимизации, доставляющее экстремум целевой функции при заданных ограничениях:

$$x^* = \arg \operatorname{extr}_{x \in X} f(x),$$

при $X = \{x \mid x_i \geq 0, \forall i \in [1, n], \psi_j(x) \leq 0, \forall j \in [1, m]\}$. Допускаются $n > m$, $n = m$, $n < m$. Функции $f(x)$ и $\psi_j(x)$ – произвольные.

В зависимости от свойств целевой функции и функций ограничений все задачи математического программирования делятся на два основных класса:

- задачи линейного программирования,
- задачи нелинейного программирования.

Если целевая функция и функции ограничений – линейные функции, то соответствующая задача поиска экстремума является задачей линейного программирования. Если хотя бы одна из указанных функций нелинейная, то соответствующая задача поиска экстремума является задачей нелинейного

программирования. В настоящей работе ограничимся рассмотрением задач нелинейного программирования.

Для конструктивного решения задачи делаются дополнительные частные предположения. В наиболее простом случае рассматривается задача *без ограничений* (оптимизация в открытой области, безусловная оптимизация). Обычно используется необходимое условие экстремума функции как следствие из теоремы Ферма:

$$\partial f(x) / \partial x_i = 0, i \in [1, n],$$

имея в виду, что достаточное условие (минимум или максимум) вытекает из физического смысла задачи. При необходимости достаточное условие определяется по знаку второй производной в точке экстремума (знак плюс означает минимум функции, а минус – максимум).

Решая эту систему уравнений, получаем требуемый набор $x^*=x^0$ из n аргументов оптимизации.

Наличие ограничений делает задачи математического программирования принципиально отличными от обычных задач математического анализа по отысканию экстремальных значений функции. В этом случае имеет место *условная оптимизация*.

В классической постановке рассматривается случай строгого равенства ограничений некоторым константам: $\psi_j(x)=b_j, \forall j \in [1, m]$. Это так называемая изопериметрическая задача («задача Дидоны»).

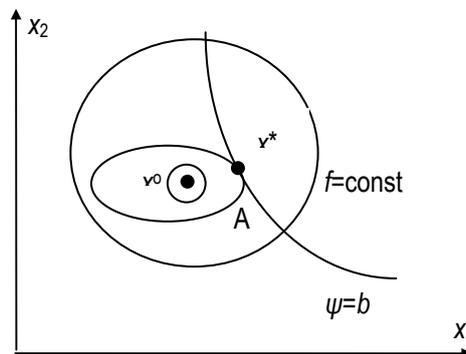


Рис.1

На Рис.1 показана двумерная по аргументам и с одним ограничением иллюстрация изопериметрической задачи. Поверхность функции $f(x)$ представляется линиями равного уровня $f(x) = \text{const}$ (как на топографических картах), а ограничение представляется проекцией пространственной кривой $\psi(x)=b$ на координатную плоскость.

Для решения изопериметрических задач пользуются методом множителей Лагранжа. Составляется функция Лагранжа

$$\Phi(x, \lambda) = f(x) + \sum_{j=1}^m \lambda_j [b_j - \psi_j(x)],$$

где $\lambda_j, j \in [1, m]$ – неопределенные множители Лагранжа, которая исследуется на *безусловный* экстремум. Для этого используется необходимое условие экстремума функции и решается система уравнений

$$\begin{aligned}\partial\Phi / \partial x_i &= 0, i \in [1, n]; \\ \partial\Phi / \partial \lambda_j &= \psi_j - b_j = 0, j \in [1, m].\end{aligned}$$

Получается n значений вектора x^* (точка А на Рис.1) и, как побочный результат, m значений вектора λ^* множителей Лагранжа, при котором уравнения связей выполняются строго.

Анализ физического смысла множителей Лагранжа говорит о том, что по величине l -го множителя λ_l , $l \in [1, m]$ можно судить о мере несоответствия l -го условия связи и требования экстремизации целевой функции. Действительно, пусть кривая $\psi(x)=b$ на Рис.1 такова, что она проходит через точку $x^{(0)}$ экстремума функции $f(x)$. В этом случае процедура множителей Лагранжа дает результат $x^*=x^{(0)}$ и $\lambda=0$, то есть здесь требования выполнения ограничения и экстремизации целевой функции не противоречат друг другу. И, наоборот, чем дальше точка x^* от $x^{(0)}$, тем больше значение коэффициента λ . Следовательно, множители Лагранжа λ в данной задаче являются векторной мерой того, насколько строгие уравнения ограничений «мешают» достичь экстремума функции $f(x)$.

Если ограничения заданы как неравенства $\psi(x) \leq 0$, а функции $f(x)$ и $\psi(x)$ выпуклые, то для решения задачи привлекается теорема Куна-Таккера, которая обобщает теорему Лагранжа для неклассических задач. Краткая формулировка теоремы Куна-Таккера: для того, чтобы целевая функция $f(x)$ достигала в точке x^* глобального условного экстремума, необходимо и достаточно, чтобы x^* , λ^* являлась глобальной седловой точкой функции Лагранжа. Это значит, что

$$\Phi(x, \lambda^*) \geq \Phi(x^*, \lambda^*) \geq \Phi(x^*, \lambda)$$

для всех x, λ из области определения задачи.

В соответствии с этой теоремой, при заданных предположениях поверхность функции Лагранжа в пространстве (x, λ) имеет оптимальную точку x^*, λ^* , которая является седловой и определяется как

$$x^*, \lambda^* = \arg \min_x \max_{\lambda} \Phi(x, \lambda).$$

Алгоритм решения задач поиска экстремумов в случае функций с седловыми точками обычно сводится к решению системы уравнений

$$x^*_{*i} \left(\frac{\partial \Phi}{\partial x_i} \right)_{x^*_{*i}} = 0, i \in [1, n]; \lambda^*_{*j} \left(\frac{\partial \Phi}{\partial \lambda_j} \right)_{\lambda^*_{*j}} = 0, j \in [1, m].$$

Анализ задач условной оптимизации приводит к выводу, что требования, предъявляемые к экстремизации целевой функции и выполнения ограничений *противоречивы*. При этом в изопериметрической задаче приоритет полностью отдается выполнению ограничений, а целевая функция экстремизируется при условии их строгого выполнения. В задаче Куна-Таккера эти противоречивые требования удовлетворяются строго в одинаковой мере.

Между тем, само понятие противоречивости требований предполагает возможность *компромисса* при их удовлетворении. Результатом этого вывода может быть методика решения, при которой условия связи выполняются не точно, а компромиссно. Если принять это допущение, то решения, полученные с помощью множителей Лагранжа, представляются всего лишь точками в континууме возможных компромиссных решений.

Проблема состоит в том, чтобы полученное при этом допущении решение отражало компромисс между противоречивыми требованиями экстремизации целевой функции и выполнения ограничений. Если указанные противоречивые требования облечь в форму частных критериев, то такая ситуация дает основание для привлечения к решению задачи условной оптимизации подхода многокритериальной (в данном случае двухкритериальной) оптимизации.

Субъективность в методике решения задачи

По своей сути компромисс является прерогативой человека, лица, принимающего решение (ЛПР). Получая компромиссное решение, мы намеренно вводим в методику решения задачи субъективный фактор в виде предпочтений ЛПР. Остановимся на этом подробнее.

В настоящее время в науке все чаще проявляются черты смены парадигмы, перехода от стиля классицизма к концепциям постклассической науки. Характерной чертой стиля современной науки является включение в методику решения задач человека (эксперт, ЛПР), в отличие от классической науки, возводившей формализацию в абсолют и ставившей целью исключать при описании действительности все сколько-нибудь субъективные суждения.

Считалось, что познание следует освобождать от случайностей, пристрастий и слабостей индивидуального человеческого исполнения. Наука строится по строгим правилам логики, и поэтому часто кажется, что личностное, человеческое – наши идеи, пристрастия, общее представление о мире – тут никакой роли не играет. Но любая логика начинает развиваться от постулатов, от аксиом, которые сами по себе никак логически не обоснованы. Бесспорно, естественные науки изучают объективный мир природы, находящейся вне нас и от нас не зависимый. Но исследователь рассматривает этот мир глазами человека определенной эпохи, определенной культуры, наконец, определенного психического и интеллектуального склада, – и все это накладывает отпечаток на его деятельность и результаты этой деятельности.

Уже создание и выбор моделей, конкретизация целевого функционала и выбор метода исследования является чисто человеческим делом и несет печать личности исследователя, а в области изучения сложных систем незаменим его личный опыт и интуиция [1]. Наука включает в свой предмет человека, допуская элементы субъективности в объективно истинном знании. Наблюдатель осознает себя частью исследуемого мира, активно взаимодействующей с наблюдаемым объектом. Познание в концепциях современной науки – диалогично.

Возвращаясь к понятию компромисса, заметим, что вся теория многокритериальной оптимизации построена на идеях компромисса и все ее проблемы сводятся к способам формализации субъективных предпочтений ЛПР.

Постановка задачи

Логика компромисса требует, чтобы при решении задачи на условный экстремум *назначались* границы допустимых, с точки зрения ЛПР, отступлений от строгих равенств по ограничениям, а также определялась предельно допустимая для него величина ухудшения целевой функции. Эти назначенные величины образуют пространство компромисса, в котором ищется приемлемое для данного ЛПР компромиссно-оптимальное решение.

Пусть выражения $\psi_j(x)=0, \forall j \in [1, m]$ таковы, что по физическим соображениям могут быть допущены отступления от строгих равенств в некоторых назначенных пределах:

$$\psi_j(x) \leq A_j, j \in [1, m].$$

Положим для определенности, что целевая функция $f(x)$ является минимизируемой. Определена некоторая предельная величина, выше которой минимизируемая целевая функция, с точки зрения ЛПР, не может (или не должна) подниматься:

$$f(x) \leq B.$$

Ставится задача: определить компромиссно-оптимальное решение x^* .

Метод решения

В данной задаче выражения для ограничений $\psi_j(x)$ и целевая функция $f(x)$ могут рассматриваться как своего рода минимизируемые, неотрицательные и ограниченные *критерии*. Возникают все предпосылки для применения к задаче условной оптимизации подхода многокритериальной оптимизации.

Суть понятия «компромисс» заключается в ответе на вопрос: сколькими единицами выигрыша по одному критерию можно, по мнению ЛПР, компенсировать неизбежный проигрыш единицы по другому критерию (другим) в заданной ситуации? На основании этой информации формулируется конкретная схема компромиссов для данной многокритериальной задачи и в итоге находится искомое решение.

Таким образом, определение многокритериального решения по своей природе компромиссно и основано на использовании субъективной информации. Имея эту информацию и выбрав схему компромиссов, можно перейти от общего векторного выражения к скалярной свертке частных критериев, что является основой для построения конструктивного аппарата решения многокритериальных задач. Возможность решения проблемы основана на гипотезе существования некоторой *функции полезности*, возникающей в сознании ЛПР при решении конкретной многокритериальной задачи. Можно утверждать, что практически все подходы к определению скалярной свертки критериев сводятся к построению той или иной математической модели функции полезности ЛПР.

Предполагается, что существуют некоторые инварианты, правила, обычно являющиеся общими для всех ЛПР независимо от их индивидуальных склонностей, которых они одинаково придерживаются в той или иной ситуации. Согласно [2] субъективность ЛПР имеет свои границы. В деловых решениях человек обязан быть рациональным, чтобы иметь возможность аргументировать мотивы своего выбора, логику своей субъективной модели. Поэтому любые предпочтения ЛПР должны находиться в рамках определенной рациональной системы. Это и делает возможной формализацию.

В данном случае предметом исследования является такая тонкая субстанция, как воображаемая функция полезности, возникающая в сознании ЛПР при решении конкретной многокритериальной задачи. У каждого ЛПР своя функция полезности. Тем не менее, можно получить сведения для задания вида содержательной модели критериальной функции, если выявить и проанализировать некоторые общие закономерности, наблюдаемые в процессе принятия многокритериальных решений различными ЛПР в разных ситуациях. При этом напряженность ситуации характеризуется близостью критериев к своим предельно допустимым значениям. Такой анализ [2,3] позволил сформулировать универсальную

скалярную свертку частных критериев, которая выражает схему компромиссов, адаптирующуюся к ситуации принятия многокритериальных решений (нелинейная схема компромиссов).

Тогда может быть предложена следующая процедура условной оптимизации, основанная на концепции нелинейной схемы компромиссов:

$$x^* = \arg \min_{x \in X} \left\{ \sum_{j=1}^m A_j [A_j - \psi_j(x)]^{-1} + B[B - f(x)]^{-1} \right\}.$$

Из этого выражения видно, что полученное решение является результатом компромисса между стремлением удовлетворить строгие ограничения $\psi_j(x)=0, \forall j \in [1, m]$ и тенденцией минимизировать целевую функцию $f(x)$. При этом решение осуществляется по нелинейной схеме компромиссов, основанной на принципе «подальше от предельно допустимых значений $A_j, j \in [1, m]$ и B ».

Скалярная свертка уравнений связи и целевой функции по нелинейной схеме в данной формуле приведена в унифицированной форме. При необходимости дополнительные индивидуальные предпочтения ЛПР могут быть учтены введением весовых коэффициентов:

$$x^* = \arg \min_{x \in X} \left\{ \alpha \sum_{j=1}^m \alpha_j A_j [A_j - \psi_j(x)]^{-1} + (1 - \alpha) B [B - f(x)]^{-1} \right\}, \alpha_j \geq 0, \sum_{j=1}^m \alpha_j = 1$$

где $\alpha_j, j \in [1, m]$ – весовые коэффициенты, отражающие относительную важность выполнения соответствующих ограничений; α – весовой коэффициент, устанавливающий относительную важность для ЛПР противоречивых тенденций: выполнения ограничений и минимизации целевой функции.

ПРИМЕР. Задана целевая функция в виде эллиптического параболоида (Рис.2):

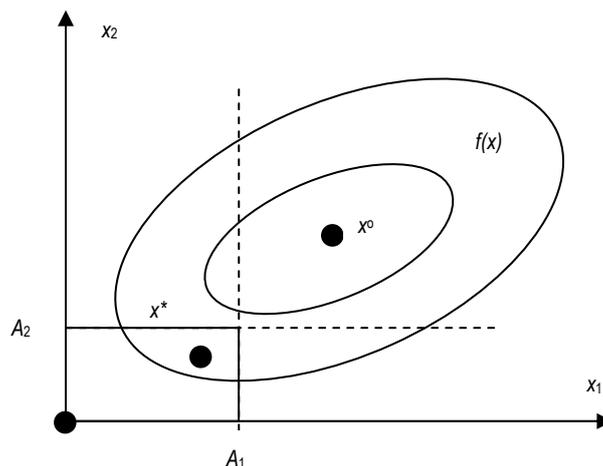


Рис. 2

$$f(x) = \frac{(x_1 - 6)^2}{2} + \frac{(x_2 - 5)^2}{4}$$

с минимумом в точке $x^0=(6; 5)$. В пространстве (x_1, x_2) заданы уравнения связи (их пересечение дает начало координат) в виде ограничений $\psi_1(x) = x_1 = 0$ и $\psi_2(x) = x_2 = 0$, которые желательно

выполнять строго, но ЛПР допускает отклонения в пределах $\psi_1 = x_1 \leq A_1 = 4$ и $\psi_2 = x_2 \leq A_2 = 3$.

Максимальная величина минимизируемой целевой функции достигается при строгом выполнении ограничений (в нашем примере в начале координат $x_1, x_2=0$) и составляет $f(x) \leq B = 24,25 \approx 25$. ЛПР считает эту величину предельно допустимой.

Ставится задача: найти такие значения аргументов оптимизации x_1^* и x_2^* , при которых требования минимизации целевой функции и выполнения ограничений удовлетворяются компромиссно.

Подставляем численные значения в выражение для компромиссной условной оптимизации в унифицированной форме (без весовых коэффициентов):

$$x^* = \arg \min_{x \in X} \left\{ \frac{4}{4-x_1} + \frac{3}{3-x_2} + \frac{25}{25 - \left[\frac{(x_1-6)^2}{2} + \frac{(x_2-5)^2}{4} \right]} \right\}.$$

Решить эту задачу можно и аналитически, выполнив безусловную минимизацию скалярной свертки. Однако в практических случаях аналитический расчет может оказаться громоздким. Обычно пользуются численными методами экстремизации функций, но гораздо удобнее применить компьютерную программу.

Для решения широкого спектра оптимизационных задач разработана и изложена в работе [2] программа векторной оптимизации TURBO-OPTIM. Программа выполнена на языке Borland C++3.1 с использованием библиотеки Turbo Vision, что обеспечивает эффективное использование ресурсов ЭВМ, стандартизованную и удобную среду для пользователя, простоту модификации и отладки.

Для работы с программой необходимо выполнить следующие этапы:

- выделить набор частных критериев так, чтобы все они принимали неотрицательные значения и требовали *минимизации*;
- определить допустимое предельное значение для каждого критерия;
- выделить набор параметров (независимых переменных), от которых зависят частные критерии;
- определить диапазон изменения каждого параметра (минимальное, стартовое и максимальное значения);
- задать ограничения для параметров, имеющие вид неравенств $g_j(r) \leq 0, j \in [1, k]$, где k – количество ограничений;
- определить вид зависимости частных критериев от параметров.

Программа позволяет решать задачи оптимизации для следующих случаев связи частных критериев с аргументами оптимизации (параметрами):

- критерии выражаются через параметры явно, известны аналитические зависимости;
- критерии являются некоторыми функционалами и для расчета их значений требуется решение системы дифференциальных уравнений;
- зависимости критериев от параметров не известны и для определения значений параметров необходимо проведение экспериментов;
- значения критериев можно получить, выполнив написанную пользователем программу;
- есть таблица зависимости частных критериев от параметров.

В каждом из перечисленных случаев программа представляет пользователю средства нахождения минимума обобщенного критерия, построенного по нелинейной схеме компромиссов, одним из методов

оптимизации: 1) метод симплекс-планирования в модификации Нелдера-Мида и 2) нелокальный метод нелинейного программирования (дуальный метод оптимизации).

В соответствии с этапами работы с программой, устанавливаем: режим «Аналитика», метод оптимизации «Симплекс-планирование» (по умолчанию) и далее вводим числовые данные

$$x_{1\min} = 0; x_{1\text{start}} = 2; x_{1\max} = A_1 = 4$$

$$x_{2\min} = 0; x_{2\text{start}} = 2; x_{2\max} = A_2 = 3$$

$$y_{1\max} = \psi_{1\max} = A_1 = 4; y_{2\max} = \psi_{2\max} = A_2 = 3; y_{3\max} = f_{\max} = B = 25$$

$$\hat{A} \tilde{n}$$

После этого даем команду «Выполнить» и программа определяет искомые значения аргументов оптимизации: $x_1^* = 1,82$, $x_2^* = 0,45$ и компромиссно-оптимальное значение целевой функции $f(x)^* = 13,89$.

Заключение

Предложенный компромиссный метод позволяет обойтись без вспомогательной категории неопределенных множителей Лагранжа, используемой как при решении изопериметрических задач, так и в процедуре седловых точек Куна-Таккера. Обычно множители Лагранжа играют лишь вспомогательную, техническую роль и представляют собой побочный результат решения задачи условной оптимизации. Иногда даже незначительный отход от строгого выполнения ограничений (или от седловой точки) позволяет существенно улучшить целевую функцию. Возможность назначения допустимых пределов для изменения ограничений и целевой функции по физическим соображениям дает ЛПР дополнительные степени свободы при решении практических оптимизационных задач.

Благодарности

Статья частично финансирована из проекта **ITHEA XX1** Института Информационных теорий и приложений FOI ITHEA и Консорциума FOI Bulgaria (www.ithea.org, www.foibg.com).

Библиография

1. Воронин А.Н. Экспертные модели многокритериальной оптимизации // International Journal "Information Technologies & Knowledge" Vol. 5, Number 3, 2011. – P. 217-223.
2. Воронин А.Н., Зиатдинов Ю.К., Козлов А.И. Векторная оптимизация динамических систем. – К.: Техніка, 1999. – 284 с.
3. Воронин А.Н. Нелинейная схема компромиссов в многокритериальных задачах оценивания и оптимизации // Кибернетика и системный анализ. – 2009. – № 4. – С. 106-114.

Сведения об авторе



Воронин Альберт Николаевич – профессор, доктор технических наук, профессор кафедры компьютерных информационных технологий Национального авиационного университета, проспект Комарова, 1, Киев-58, 03058 Украина; e-mail: alnv@voliacable.com

МОДЕЛИРОВАНИЕ ВРЕМЕННЫХ ЗАВИСИМОСТЕЙ В ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМАХ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ НА ОСНОВЕ ПРЕЦЕДЕНТОВ

Александр Еремеев, Павел Варшавский, Иван Куриленко

Аннотация: В работе рассматриваются вопросы, связанные с моделированием временных зависимостей в интеллектуальных системах поддержки принятия решений (ИСППР) на основе прецедентов. Описываются особенности применения систем на основе прецедентов (СВР-систем) для поддержки принятия решений в открытых и динамических предметных областях при наличии достаточно жестких временных ограничений. Рассмотрен пример использования СВР-системы с темпоральными прецедентами для контроля нештатных ситуаций в работе ИСППР для управления крупными парковочными комплексами.

Ключевые слова: интеллектуальные системы поддержки принятия решений, моделирование временных зависимостей, темпоральные прецеденты, СВР-системы.

ACM Classification Keywords: H.4.2 [Information systems applications]: Types of systems – Decision support; I.2.3 [Artificial intelligence]: Deduction and Theorem Proving – Uncertainty, "fuzzy," and probabilistic reasoning; I.2.4 [Artificial intelligence]: Knowledge Representation Formalisms and Methods – Temporal logic.

Введение

Актуальной задачей в области искусственного интеллекта (ИИ) и создания перспективных интеллектуальных систем типа ИСППР является задача моделирования правдоподобных рассуждений [Вагин и др., 2008]. Наличие подобных механизмов рассуждений в ИСППР, предназначенных для мониторинга и управления сложными объектами и процессами различной природы, позволяет осуществлять оперативное диагностирование проблемной ситуации и помогает лицам, принимающим решения (ЛПР), находить адекватные и эффективные управляющие воздействия с целью нормализации проблемной ситуации.

Наряду с широко используемыми в области ИИ методами правдоподобных рассуждений на основе индукции, абдукции, аргументации и аналогии активно применяются методы на основе прецедентов, в том числе, на основе активно разрабатываемых в настоящее время темпоральных прецедентах [Вагин и др., 2008], [Варшавский и др., 2009]. Данные методы могут быть эффективны в ИСППР для мониторинга и поиска решения в реальном масштабе времени и в условиях различного рода неопределенности в исходных данных и знаниях, которые содержатся в базе знаний (БЗ) системы, а также для обучения и тренировки оперативно-диспетчерского персонала.

Методы на основе прецедентов в настоящее время активно применяются в таких областях, как медицинская диагностика, юриспруденция, мониторинг и диагностика технических систем, поиск решения

в проблемных ситуациях и т.д. Данный подход составляет основу машинного обучения и предоставляет широкие возможности для формирования корпоративной памяти.

Несмотря на имеющиеся разработки в данной области [Watson et al., 1994], на данный момент ощущается нехватка в программных средствах для моделирования рассуждений на основе прецедентов в интеллектуальных системах типа ИСППР реального времени (ИСППР РВ), систем экспертной диагностики и SCADA-систем.

Наиболее актуальными задачами в данной области являются задачи организации представления, хранения и обработки данных и знаний (прецедентов) в CBR-системах с возможностью учета временных (темпоральных) зависимостей. Решению указанных задач уделено основное внимание в данной работе.

Рассуждения на основе прецедентов

В большинстве энциклопедических источников **прецедент** (от латинского *praecedentis* – предшествующий) определяется как случай, имевший место ранее и служащий примером или оправданием для последующих случаев подобного рода.

Вывод на основе прецедентов (Case-Based Reasoning, CBR) является подходом, позволяющим решить новую, неизвестную задачу, используя или адаптируя решение уже известной задачи, т.е. используя уже накопленный опыт решения подобных задач.

Подход на основе прецедентов возник в процессе развития исследований в области создания экспертных систем (систем, основанных на знаниях). Большинство практических задач в данной области являются плохо формализованными, причем неопределенность может иметь не вероятностный характер. При поиске решения таких задач, ориентированных на открытые и динамические предметные области, необходимо применение методов правдоподобных рассуждений, позволяющих найти некоторое приемлемое (которое может не быть оптимальным) решение. Один из подходов базируется на том факте, что человеку (эксперту, ЛПП) свойственно на первом этапе поиска решения новой (неизвестной) задачи пытаться использовать решения, которые принимались ранее в подобных случаях и при необходимости адаптировать их к возникшей проблеме (текущей проблемной ситуации). Данный подход с использованием накопленного опыта лег в основу методов моделирования рассуждений на основе прецедентов. Основой для разработки отмеченного подхода и CBR-систем (в частности, первой CBR-системы CYRUS) послужила работа [Schank, Abelson, 1977], затрагивающая проблемы организации памяти и представления знаний.

В настоящее время интерес к CBR-технологии и CBR-системам значительно возрос, регулярно проводятся международные конференции и семинары ICCBR, ECCBR, UKCBR (<http://www.iccbr.org/>).

Как правило, CBR-методы включают в себя четыре основных этапа, образующие так называемый цикл рассуждения на основе прецедентов или CBR-цикл (рис. 1) [Варшавский и др., 2006]:

- **Retrieve** – извлечение наиболее подобного прецедента (или прецедентов) для сложившейся ситуации из библиотеки (базы) прецедентов (БП).
- **Reuse** – повторное использование извлеченного прецедента (прецедентов) с целью предложения решения для текущей проблемы.
- **Revise** – пересмотр полученного решения и адаптация его, в случае необходимости, в соответствии с текущей проблемой.
- **Retain** – сохранение вновь принятого решения как части нового прецедента.

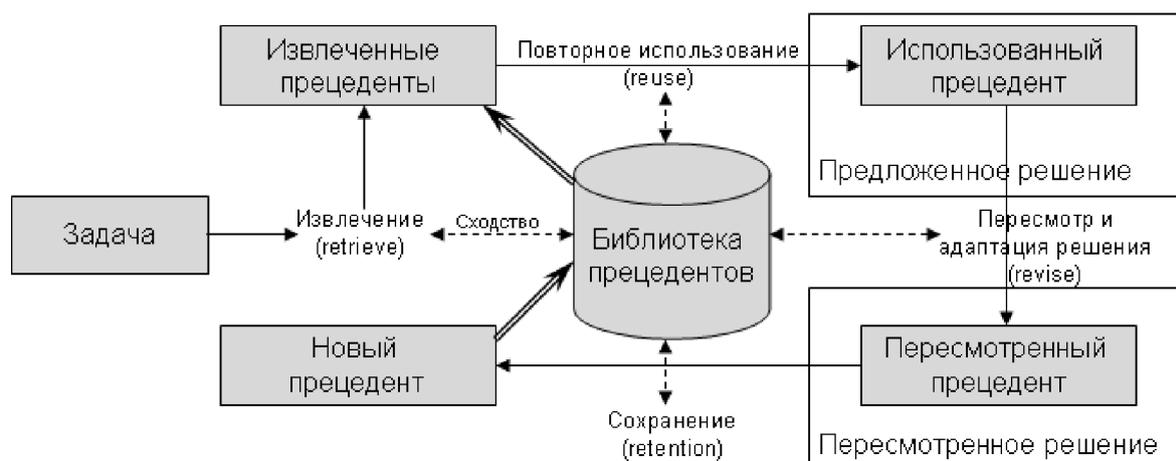


Рис. 1. CBR-цикл

Основная цель использования аппарата прецедентов в рамках ИСППР заключается в выдаче готового решения оператору (ЛПР) для текущей ситуации на основе прецедентов, которые уже имели место в прошлом при управлении данным или подобным объектом (системой). Часть этих прецедентов изначально добавляется в БП системы экспертом, а остальные прецеденты пополняют БП системы за счет выполнения последнего этапа CBR-цикла.

Способы представления прецедентов

Существуют различные способы представления и хранения прецедентов [Варшавский и др., 2009], [Варшавский и др., 2006] – от простых (линейных) до сложных иерархических. Следует отметить, что простые способы, базирующиеся, как правило, на технологии реляционных баз данных, требуют значительно меньших затрат на реализацию, а также на поддержание и сопровождение БП, чем сложные.

Прецедент в общем случае может включать следующие компоненты:

- **описание задачи** (проблемной ситуации);
- **решение задачи** (диагностирование проблемной ситуации и рекомендации ЛПР);
- **результат** (или прогноз) применения решения.

Результат может включать список выполненных действий, дополнительные комментарии и ссылки на другие прецеденты. Прецедент может иметь как положительный, так и отрицательный исход применения решения, также в некоторых случаях может приводиться обоснование выбора предложенного решения и возможные альтернативы.

В большинстве случаев для представления прецедентов достаточно простого параметрического представления, т.е. представления прецедента в виде набора параметров с конкретными значениями и решением (диагнозом и рекомендациями ЛПР): $CASE=(x_1, x_2, \dots, x_n, R)$, где x_1, \dots, x_n – параметры ситуации, описывающей данный прецедент; $x_1 \in X_1, x_2 \in X_2, \dots, x_n \in X_n$, где n – количество параметров прецедента, а X_1, \dots, X_n – области допустимых значений соответствующих параметров, R – диагноз и

рекомендации ЛПР. Дополнительно может присутствовать описание результата применения найденного решения и дополнительные комментарии.

Существует целый ряд методов извлечения прецедентов и их модификаций [Варшавский и др., 2006], [Егемеев et al., 2009], такие как метод ближайшего соседа, метод извлечения прецедентов на основе деревьев решений, метод извлечения прецедентов с учетом их применимости (адаптивности), метод извлечения прецедентов с использованием аппарата искусственных нейронных сетей и другие.

Как правило, классические CBR-методы позволяют извлекать прецеденты на основе значений параметров контролируемой системы в текущий момент времени, но без учета динамики процесса (т.е. ситуации и прецеденты содержат «мгновенные снимки» контрольных параметров наблюдаемого объекта или системы и нет учета истории их изменения). Так, например, с помощью метода ближайшего соседа текущие значения параметров могут сравниваться с прецедентами, и на их основе может быть выбрана рекомендация (решение). Однако, большинство физических процессов развивается в соответствии с некоторым временным законом и, учитывая историю изменения состояний наблюдаемого объекта или процесса, можно находить более качественные решения и рекомендации, чем на основании анализа только текущего состояния. Необходим новый способ представления прецедентов, позволяющий учитывать историю изменения параметров – метод на основе **темпоральных прецедентов**, а также адаптированные к учету фактора времени алгоритмы извлечения прецедентов.

В рамках данной работы предлагается расширение CBR-методов, позволяющее учитывать поведение контролируемого процесса или объекта во времени. При учете фактора времени появляется возможность рассмотреть проблемную ситуацию в динамике, то есть текущая ситуация сравнивается не с какими-либо фиксированными значениями прецедентов, а отслеживается процесс изменения значений, что позволяет строить предположения, используя не только критерий сходства, но и учитывать временные зависимости.

Построение темпорального расширения CBR-методов

Рассмотрим расширение методов извлечения прецедентов с учетом фактора времени. Наиболее простым способом адаптации существующих алгоритмов вывода на основе прецедентов к учету фактора времени является подмена «мгновенного» снимка ключевых параметров историей их изменения (рис. 2).

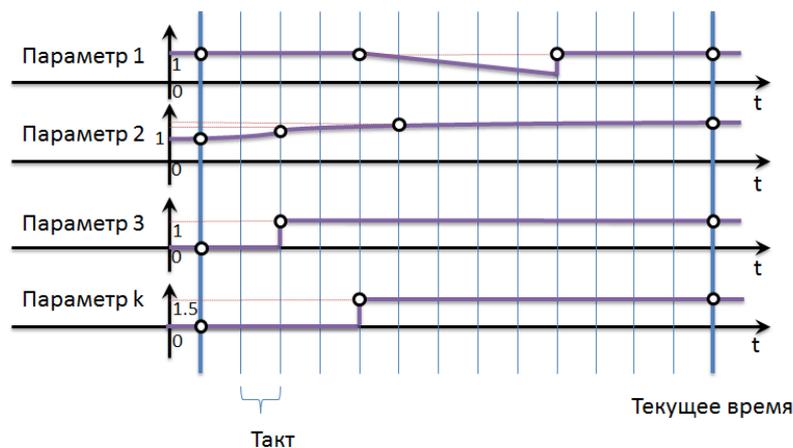


Рис. 2. Представление истории изменения параметров в виде графика

Эта история может быть представлена и в виде матрицы значений, составленной на определенный период наблюдения с определенным интервалом (тактом) (рис. 3).

Параметр	1	2	3	4	5	6	7	...	N
1	1	1	1	1	1	0.85	0.75	...	1
2	1	1.1	1.2	1.24	1.27	1.28	1.29	...	1
...
K	0	0	0	0	1	1	1	...	1

Рис. 3. Представление истории изменения параметров в виде матрицы

Для обработки темпоральных прецедентов могут использоваться упомянутые выше методы. При этом в прецедент включается не только значение параметров в текущий момент времени, но и их значения на некоторый период времени до этого (их история). Эксперту посредством интерфейса может представляться более удобная для интерпретации и заполнения матричная форма, а ее трансформацию в массив для применения того или иного метода CBR-система может выполнять автоматически.

В первую очередь следует определить отрезок времени для анализа проблемной ситуации (глубину анализа). Рассматриваемый промежуток времени разбивается на N равных отрезков с некоторым шагом (тактом). Далее в сформированной БП, где каждый прецедент представляется в параметризованном виде, происходит сравнение значений параметров ситуации в каждый момент времени t_i , где $i = 1, \dots, N$. Значения параметров прецедентов сравниваются между собой по методу ближайшего соседа с учетом выбранной метрики (например, Евклидовой) и соответствующего порогового значения [Варшавский и др., 2009], [Варшавский и др., 2006].

Также для адаптации существующих алгоритмов поиска решения к темпоральным прецедентам может использоваться подход на основе учета и анализа решений, полученных ранее. Проиллюстрируем данный подход на примере расширения метода ближайшего соседа.

Таким образом, для всех точек i получаются варианты прецедентов с определенными интегральными оценками, которые могут быть выбраны как соответствующие ситуации в этой точке для текущего порогового значения сходства.

На следующем этапе по интегральным оценкам в точках i выбирается наиболее вероятный прецедент или группа прецедентов, удовлетворяющих поисковому условию. Каждому результирующему прецеденту в соответствие ставится мера сходства (расстояние Хэмминга), которая определяется по правилу: «При совпадении всех параметров в описании прецедента и текущей ситуации степень сходства будет равна 1, а каждый совпавший параметр дает вклад равный $1/K$, где K – число параметров в описании прецедента и текущей ситуации». Используя значения параметров в точках i , можно с помощью интерполяции построить прогноз развития проблемной ситуации (на уровне параметров).

Следует отметить, что на результат поиска прецедентов, удовлетворяющих проблемной ситуации, оказывает влияние несколько факторов. Во-первых, это выбор метрики. В каждом конкретном случае этот выбор производится по-разному, в зависимости от целей ЛПР, физической и статистической природы используемой информации при управлении сложным объектом и других ограничений и

факторов, влияющих на процесс поиска решения. Во-вторых, влияние оказывает выбор шага, с которым разбивается на промежутки рассматриваемый отрезок времени. Вариация размеров шага дает различное количество дискретных точек, по которым происходит сравнение значений параметров прецедентов. Это оказывает влияние не только на точность выбора прецедента, но и на скорость выполнения поискового алгоритма, а так же на величину погрешности, с которой может быть определено соответствие. И, в-третьих, важным оказывается выбор порогового значения, определяющего степень сходства.

Проиллюстрируем вышесказанное на несложном примере [Еремеев и др., 2011]. Предположим, что существует некая проблемная ситуация, описываемая набором из трех параметров (рис. 4), и имеется заданная БП, включающая прецеденты X, Y, Z (рис. 5-7). Произведя сравнительный анализ значений в каждой точке i из отрезка времени $[0, N]$, можно получить диаграмму интегральных оценок соответствия каждого прецедента текущей проблемной ситуации во всех точках выбранного отрезка времени (рис. 8).

Из рассмотренного примера видно, что прецедент Y будет выбран как прецедент, наиболее удовлетворяющий условиям поиска и имеющий большую степень сходства с текущей проблемной ситуацией по сравнению с другими прецедентами X и Z, так как интегральные оценки прецедента Y в каждой точке выше, чем соответствующие оценки других прецедентов в этих точках.

Предложенные методы могут быть использованы как по отдельности, так и в комбинации друг с другом. В практических приложениях важно правильно определить величину N и размер такта. Если следует учитывать вероятность скачкообразного изменения поведения параметров контролируемого объекта или процесса, то предпочтительнее использовать первый из рассмотренных выше методов – метод подмены «мгновенного» снимка ключевых параметров историей их изменения.

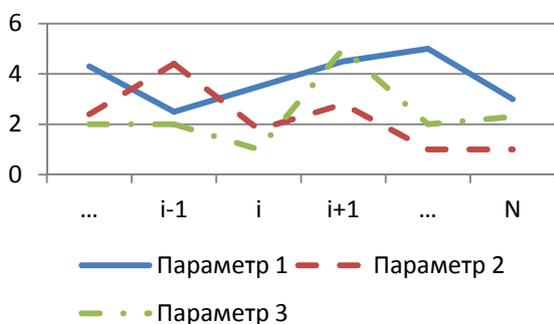


Рис. 4. Значения параметров текущей проблемной ситуации

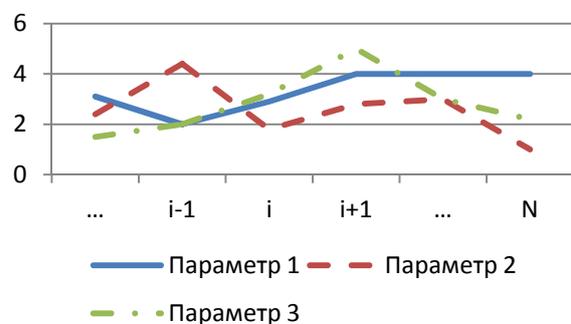


Рис. 5. Значения параметров прецедента X

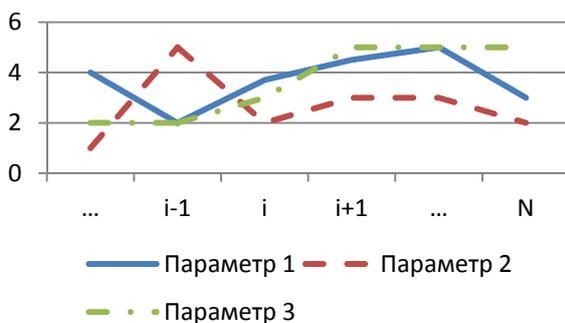


Рис. 6. Значения параметров прецедента Y

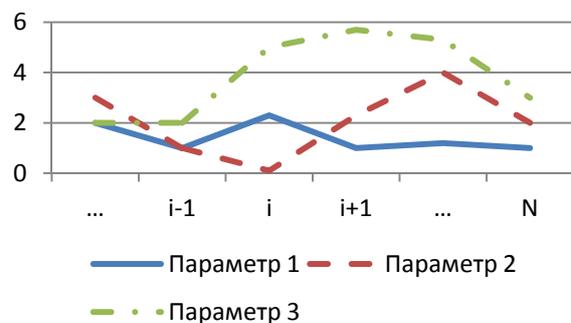


Рис. 7. Значения параметров прецедента Z

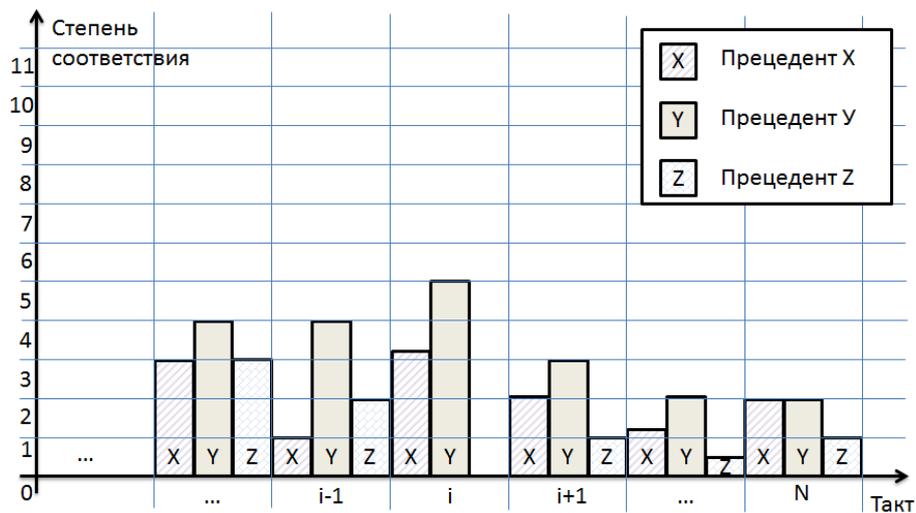


Рис. 8. Варианты precedентов с интегральными оценками

Использование темпоральной логики

Для реализации механизма вывода на основе темпоральных precedентов могут быть применены качественная точечная временная логика и метрическая временная логика [Еремеев и др., 2007]. Данный выбор обусловлен тем, что для этих логик существуют алгоритмы вывода с полиномиальной оценкой сложности [Еремеев и др., 2010].

Метрическая точечная задача согласования временных ограничений (МЗСВО) определяется как $Z=(V,D,C_1,C_2)$, где $V = \{V_1, V_2, \dots, V_m\}$ - конечное множество временных переменных, соответствующих моментам времени; D - область значений временных переменных (является либо множеством вещественных чисел, либо множеством целых чисел); C_1 - конечное число бинарных временных ограничений вида $C_{ij}=\{[a_1, b_1], \dots, [a_k, b_k]\}$, где интервалы попарно не пересекаются; C_2 - конечное число унарных временных ограничений вида $C_i = \{[a_1, b_1], \dots, [a_i, b_i]\}$, где интервалы попарно не пересекаются. Бинарные ограничения позволяют определить разрешенную дистанцию между моментами времени и интерпретируются как $(a_1 \leq V_j - V_i \leq b_1) \cup \dots \cup (a_k \leq V_j - V_i \leq b_k)$. Унарные - интерпретируются как $(a_1 \leq V_i \leq b_1) \cup \dots \cup (a_i \leq V_i \leq b_i)$ и могут быть представлены с помощью бинарных (через введение нулевого момента времени V_0 и преобразовании унарных ограничений C_i в C_{0i}). Для решения задачи требуется скорректировать все ограничения так, чтобы они не противоречили друг другу. Если подобная корректировка возможна МЗСВО считается согласованной (иначе - несогласованной).

Рассмотрим основные операции над метрическими точечными ограничениями. **Отрицание** ограничения $C_{ij}=\{[a_1, b_1], \dots, [a_k, b_k]\}$ определяется как $\sim C_{ij}=\{[-b_1, -a_1], \dots, [-b_k, -a_k]\}$. **Пересечение** $T \cap S$ ограничений T и S содержит значения, которые одновременно присутствуют в обоих ограничениях. **Композиция** $T \cdot S$ ограничений T и S содержит значения r , для которых существуют значения $t \in T$ и $s \in S$, такие, что $r = t + s$.

МЗСВО может быть применена для описания наблюдаемых ситуаций и precedентов. При этом фиксируется как сам факт наступления тех или иных событий, их порядок, так и время их появления (метрика). Далее будем предполагать, что эти МЗСВО согласованы и преобразованы к минимальной

форме. Для вычисления соответствия наблюдаемой ситуации и прецедента могут быть предложены несколько методов – метод с «жесткими» ограничениями и метод с «мягкими» ограничениями.

Метод с «жесткими» ограничениями предполагает точное соответствие наблюдаемых событий и событий, имеющих в прецеденте, а также для каждого метрического ограничения C_{ij} в прецеденте и ограничения C_{ij}^* в наблюдаемой ситуации должны выполняться условия $C_{ij} \cap C_{ij}^* = C_{ij}$. При этом, для того, чтобы события в прецеденте и анализируемой ситуации соответствовали друг другу предполагается использовать для их нумерации числа, получаемые в результате сортировки по имени параметра и времени (см. пример на рис. 9). Среди достоинств этого метода – высокая скорость вывода и высокая точность результата. Метод также полезен для систем, в которых временные характеристики протекающих процессов достаточно стабильны.

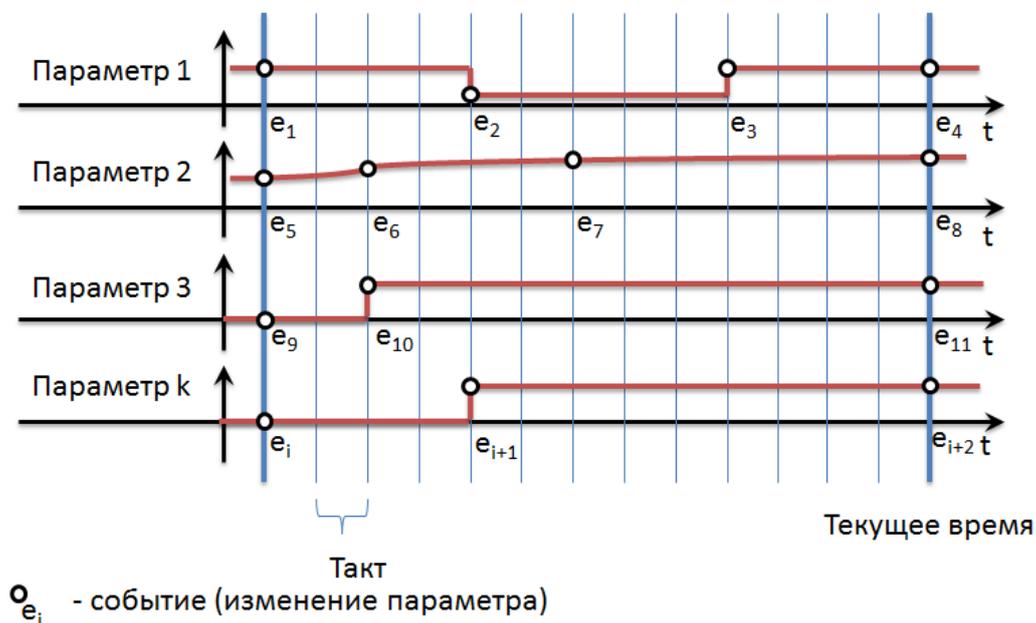


Рис. 9. Пример нумерации событий

Метод с «мягкими» ограничениями может рассматриваться в трех вариантах. В первом варианте условие $C_{ij} \cap C_{ij}^* = C_{ij}$ заменяется на более мягкое $C_{ij} \cap C_{ij}^* \neq \emptyset$ и $C_{ij}^* \subseteq C_{ij}$. Во втором - условие $C_{ij} \cap C_{ij}^* = C_{ij}$ заменяется на $C_{ij} \cap C_{ij}^* \neq \emptyset$; а в третьем оно заменяется на условие близости границ в ограничениях $C_{ij} = \{[a_1, b_1], \dots, [a_k, b_k]\}$ и $C_{ij}^* = \{[a_1^*, b_1^*], \dots, [a_k^*, b_k^*]\} : (|a_1 - a_1^*| < \varepsilon) \wedge \dots \wedge (|a_k - a_k^*| < \varepsilon) \wedge (|b_1 - b_1^*| < \varepsilon) \wedge \dots \wedge (|b_k - b_k^*| < \varepsilon)$, где $\varepsilon > 0$. Дальнейшее смягчение условий возможно при анализе степени сходства с учетом истории изменений каждого параметра по отдельности. Тут возможны разные стратегии (например, учитывающие с большим весом ограничения между событиями, связанными с изменением значений разных параметров, чем ограничения, наложенные на события полученные из-за изменения одного параметра).

Определение прецедента на основе выборки похожих ситуаций (обучение) в данном случае может быть выполнено на основе смягчения ограничений C_{ij} в одном из вариантов так, чтобы выполнялись условия его подобия другим ситуациям.

Формально ситуация определяется как $S = \langle V, C, P, \alpha \rangle$, где S – текущая ситуация; $V = \{V_1, V_2, \dots, V_m\}$ – конечное множество временных переменных, соответствующих моментам времени; $C = \{C_{ij}\}$ – конечное множество метрических временных ограничений, где C_{ij} – это ограничение для временных переменных V_i и V_j ; $P = \{P_1, \dots, P_k\}$ – набор параметров контролируемого объекта; $\alpha: V \rightarrow P$ – функция, сопоставляющая каждой временной переменной (событию) набор параметров, которые характеризовали состояние объекта или процесса в данный момент времени.

Темпоральный прецедент формально можно определить набором $\langle I, V, C, P, \alpha \rangle$, представляющим ситуацию, расширенную блоком описательных параметров I , куда могут входить статус прецедента (норма, ошибка, предупреждение), описание, ссылки, временной фактор, диагноз (решение) и т.д.

При извлечении прецедента для наблюдаемой ситуации используется алгоритм, учитывающий временные характеристики и состояние контролируемого объекта или процесса:

Алгоритм извлечения прецедента

Входные данные: $S = \langle V^s, C^s, P^s, \alpha \rangle$ – наблюдаемая ситуация,

$D = \{D_i\}$ – база прецедентов, где $D_i = \langle I^i, V^i, C^i, P^i, \alpha^i \rangle$.

Выходные данные: D_k – прецедент, соответствующий ситуации S или \emptyset

```

01: foreach ( $D_k = \langle I^k, V^k, C^k, P^k, \alpha^k \rangle \in D$ ) {
02:   bool isFailed ← false
03:   foreach ( $C_{ij}^s \in C^s$ ) // Сравнение временных характеристик
04:     if ( $C_{ij}^s \cap C_{ij}^k = \emptyset$ , где  $C_{ij}^k \in C^k$ ) { isFailed ← true; break; }
05:   if (!isFailed) {
06:     for ( $i = 0; i < |V^s|; i = i + 1$ )
07:       if (!СравнитьПодобие( $\alpha(V^s), \alpha(V^k)$ )) { isFailed ← true; break; }
08:   }
09:   if (!isFailed) return  $D_k$ 
10: }
11: return  $\emptyset$ 

```

В алгоритме для определения подобия (сходства) может быть применен подходящий для решаемой задачи метод.

Таким образом, организация темпоральных прецедентов на основе МЗСВО позволяет учитывать как последовательности событий, так и их длительности. В случае, если анализ длительностей не принципиален, а важен только порядок событий, можно воспользоваться качественной временной логикой – например, точечной временной логикой, для которой предложены быстродействующие алгоритмы вывода [Еремеев и др., 2007]. В этом случае для представления прецедента может использоваться точечная ЗСВО, а определение степени сходства наблюдаемой ситуации и прецедента может основываться на решении задачи логической эквивалентности соответствующих минимальных ЗСВО.

Программная реализация прототипа

Рассмотренные выше методы реализованы в прототипе модуля анализа нетиповых (нештатных) ситуаций для интеллектуальной системы (ИС), на примере ИС для управления крупными парковочными комплексами (ИС УП), построенной по модульному принципу с учетом требований расширяемости и простоты интеграции. Программная реализация выполнена на базе платформы Microsoft .NET Framework 4.0 на языке C# в среде разработки Microsoft Visual Studio 2010 и является кросс-платформенной. Прототип реализует следующие функции: поиск и выделение сбоев в работе парковочного комплекса, которые относятся к типовым проблемным ситуациям, а также определение нестандартных ситуаций, которые возникли в процессе эксплуатации ИС УП, но не были учтены при разработке управляющих правил. Созданный прототип может функционировать в трёх режимах:

- **режим обучения** – обучение ИС УП на основе данных из файлов и знаний экспертов;
- **“on-line” - режим** – события поступают напрямую на вход системы;
- **“off-line” - режим** – события читаются из протоколов (логов) работы системы.

Архитектура прототипа анализа нетиповых (нештатных) ситуаций для ИС УП приведена на рис. 10.

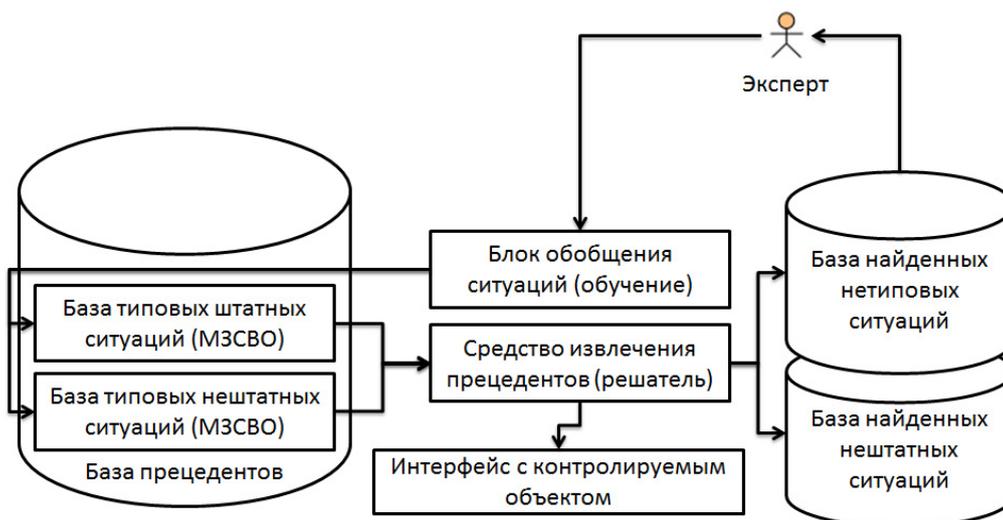


Рис. 10. Архитектура прототипа модуля анализа нетиповых (аномальных) ситуаций

Система состоит из блока обучения, блока вывода, базы прецедентов, и трёх интерфейсов: интерфейса взаимодействия с объектом, экспертом и пользователем. ИС содержит базу типовых ситуаций (как штатных, так и известных нестандартных). Решатель на основе прецедентов соотносит временные диаграммы в последовательностях операций за определенный период с моделями, хранящимися в этих базах. В случае, если выявляется типовая нестандартная ситуация, ее данные заносятся в базу нестандартных ситуаций. Если же выделяется ситуация, которая не описана ни в базе штатных типовых ситуаций, ни в базе нестандартных типовых ситуаций, то эта ситуация заносится в отдельную базу, содержимое которой анализируется разработчиками (экспертами) ИС УП.

Система соотносит временные диаграммы в последовательностях операций за определенный период с моделями, хранящимися в этих базах. В случае, если выявляется типовая нестандартная ситуация, то ее данные заносятся в базу нестандартных ситуаций. Если же выделяется ситуация, которая не описана ни в базе штатных типовых ситуаций, ни в базе нестандартных типовых ситуаций, то эта ситуация заносится в

отдельную базу, содержимое которой анализируется разработчиками ИС УП. Классификация осуществляется в три этапа: выборка событий из источников; выделение границ процессов; анализ каждой выделенной операции.

Прецеденты, включая темпоральные, хранящиеся в БП, представляют собой операции (ситуации). В общем случае каждый прецедент имеет имя, код (номер), описание (диагноз), последовательность действий (что необходимо предпринять при выявлении этого прецедента), а также структуру событий.

В блоке сопоставления для выявления типовых и нестандартных ситуаций применяются алгоритмы сопоставления. Алгоритм сопоставления ситуаций опирается на их представление. В зависимости от представления прецедентов строится та или иная модель прецедента.

Пример практического применения

Предлагаемые далее методы рассчитаны на применение в составе CBR-системы, использующейся для контроля нестандартных ситуаций в работе ИС УП [Eremeev et al., 2009], [Куриленко, 2009]. ИС УП, относящаяся к классу ИСППР РВ, предназначена для эффективного управления доступом и движением автотранспорта на современных парковках, представляющих собой территории, оборудованные программно-аппаратными комплексами, в которых задействовано значительное число автоматических и исполнительных устройств и датчиков (средства разграничения доступа, автоматизации оплаты, автоматические шлагбаумы, камеры видеонаблюдения и фото-идентификации, охранная и пожарная сигнализация и др.). В ИС УП в качестве объекта доступа выступает автомобиль, а исполнительными устройствами являются шлагбаумы и ворота, которые необходимо закрывать после въезда транспорта. ИС требуется контролировать движение автомобиля по территории. При этом осуществляется контроль проезда через точки доступа, узкие места, коридоры и пандусы, что требует учета временных характеристик.

Парковки со значительным числом машиномест (больше 1000) обычно содержат большое число точек доступа и оплаты, а также могут разбиваться на зоны с различным уровнем доступа и тарификацией. Проезд между зонами может осуществляться только через въездные и выездные точки доступа. Следует отметить, что каждая такая точка является автономной, а ее исполнительные устройства зачастую находятся на существенном расстоянии от сервера системы. На практике ИС УП должна реализовывать функции защиты не только от противоправных действий пользователей, но и от махинаций «нечистоплотного» персонала, заинтересованного в возможности присвоения денежных средств, поступающих в качестве оплаты за услуги парковки. Как следствие, посетители и обслуживающий персонал в некоторых случаях предпринимают попытки «обойти» контролирующие функции системы, имитировать сбой или информировать о ложных сбоях. Кроме того, любые ошибочные действия персонала могут интерпретироваться как сбой. В такой ситуации возникает задача контроля действий как персонала и пользователей, так и автоматических устройств, с целью отслеживания и недопущения нетиповых ситуаций. А если такие ситуации возникают, то следует их классифицировать и объяснить реальную причину наблюдаемых событий.

Благодаря наличию большого числа датчиков на точках доступа, возможна организация специального модуля, отслеживающего ход операций и расширяющего возможности ИС УП в плане реакции на нетиповые ситуации. При построении этого модуля может быть использован тот факт, что операции,

протекающие на точках доступа в штатном режиме, формируют стандартные последовательности событий, проиллюстрированные на примерах типового хода событий при операциях проезда (рис. 11).

Параметр		Время
Датчик наличия автомобиля перед шлагбаумом	есть	
	нет	
Датчик наличия автомобиля после шлагбаума	есть	
	нет	
Датчик верхнего положения стрелы шлагбаума (шлагбаум открыт)	1	
	0	
Датчик нижнего положения стрелы шлагбаума (шлагбаум закрыт)	1	
	0	
Датчик на выходе контроллера, управляющего шлагбаумом	открыть	
	закреть	
Датчик на выходе контроллера, управляющего светофором	зеленый	
	красный	
	желтый	
Кнопка запроса на въезд	нажата	
	отжата	
Датчик изъятия разового документа на въезд (билета)	1	
	0	

Рис. 11. Информация с датчиков, получаемая при проезде посетителя по разовому документу

Разработанный прототип повышает эффективность ИС УП за счет введения дополнительного модуля анализа типовых и аномальных (нештатных) ситуаций с использованием методов поиска решения на основе темпоральных прецедентов. В задачи этого модуля входит: выделение сбоев в работе ИС УП и подготовка экспертной оценки наблюдаемых сбоев по базе типовых проблемных ситуаций, определение нестандартных ситуаций, возникших в процессе эксплуатации ИС УП, но не учтенных при разработке управляющих правил, пресечение попыток противодействия обслуживающего персонала и посетителей.

Заключение

В работе рассмотрена проблема моделирования правдоподобных рассуждений (рассуждений «здравого смысла») на основе прецедентов. Описаны основные этапы, реализуемые в CBR-методах, а также рассмотрены способы представления прецедентов. Предложены методы моделирования временных зависимостей в CBR-системах на основе темпоральных прецедентов и соответствующие алгоритмы, которые могут применяться для расширения возможностей и повышения эффективности современных ИСППР реального времени, позволяя учитывать в процессе поиска решения (рассуждения) временные зависимости и динамику развития ситуации на объекте. Использование темпоральных прецедентов в CBR-системах, ориентированных на использование в составе ИСППР реального времени для мониторинга и управления различными сложными техническими объектами и процессами позволяет ЛПР (как это проиллюстрировано на примере ИС УП) своевременно выбрать необходимую стратегию поведения для устранения проблемной ситуации за счет сокращения поискового пространства и уменьшения времени реакции системы.

Благодарности

Работа поддержана Российским Фондом Фундаментальных Исследований (проекты 11-01-00140, 12-07-00508).

Список литературы

- [Вагин и др., 2008] Вагин В.Н., Головина Е.Ю., Загорянская А.А., Фомина М.В. Достоверный и правдоподобный вывод в интеллектуальных системах. 2-е изд. // Под редакцией В.Н. Вагина, Д.А. Поспелова. – М.: ФИЗМАТЛИТ, 2008.
- [Варшавский и др., 2009] Варшавский П.Р., Еремеев А.П. Моделирование рассуждений на основе прецедентов в интеллектуальных системах поддержки принятия решений // Искусственный интеллект и принятие решений, №2, 2009, с. 45-57.
- [Watson et al., 1994] Watson I.D., Marir F. Case-based reasoning: A review. The Knowledge Engineering Review, Vol. 9, No. 4, 1994. – Pp. 355-381.
- [Schank, Abelson, 1977] Schank R.C., Abelson R.P. Scripts, Plans, Goals and Understanding. Erlbaum, Hillsdale, New Jersey, US, 1977, - 248 p.
- [Варшавский и др., 2006] Варшавский П.Р., Еремеев А.П. Методы правдоподобных рассуждений на основе аналогий и прецедентов для интеллектуальных систем поддержки принятия решений // Новости искусственного интеллекта. 2006. №3. – С. 39-62.
- [Eremeev et al., 2009] Alexander Eremeev, Ivan Kurilenko, Pavel Varshavskiy. Application of Temporal Reasoning and Case-based Reasoning in Intelligent Decision Support Systems // International Book Series «Information science & computing», Number 10, Supplement to IJ «Information technologies & knowledge» Volume 3/2009 – 2009. – Pp. 9–16.
- [Еремеев и др., 2011] Еремеев А.П., Куриленко И.Е., Смирнова А.Е. Разработка темпорального расширения методов рассуждений на основе прецедентов // Труды конгресса по интеллектуальным системам и информационным технологиям (IS&IT'11). Научное издание в 4-х томах. - Т.1. – М.: Физматлит, 2011. - С.50-59.
- [Еремеев и др., 2007] Еремеев А.П., Куриленко И.Е. Реализация механизма временных рассуждений в современных интеллектуальных системах // Известия РАН. Теория и системы управления, 2007, № 2, с. 120–136.
- [Еремеев и др., 2010] Еремеев А.П., Куриленко И.Е. Применение темпоральных моделей в интеллектуальных системах / Интеллектуальные системы. Коллективная монография. Выпуск четвертый. / Под. Ред. В.М. Курейчика. – М.: Физматлит, 2010, 300 с., 222-252.
- [Куриленко, 2009] Куриленко И.Е. Применение временной логики при построении интеллектуальной системы управления крупными парковочными комплексами // Сб. док. научно-практ. конф. «Интегрированные модели и мягкие вычисления в искусственном интеллекте» 2009 в 2 т. – Т.2 – М.:ФизМатЛит, 2009. С. 171 -180.

Сведения об авторах

Еремеев Александр Павлович – д.т.н., проф., заведующий кафедрой Прикладной математики Московского энергетического института, 111250, Россия, Москва, Красноказарменная ул., 14; e-mail: eremeev@apmat.ru

Область научных интересов: искусственный интеллект, принятие решений, экспертные системы

Варшавский Павел Романович – к.т.н., доцент кафедры Прикладной математики Московского энергетического института, 111250, Россия, Москва, Красноказарменная ул., 14; e-mail: varp@apmat.ru

Область научных интересов: искусственный интеллект, принятие решений, методы правдоподобных рассуждений

Куриленко Иван Евгеньевич – к.т.н., доцент кафедры Прикладной математики Московского энергетического института, 111250, Россия, Москва, Красноказарменная ул., 14; e-mail: ivan@apmat.ru

Область научных интересов: искусственный интеллект, принятие решений, темпоральные логики

ПОСЛЕДОВАТЕЛЬНЫЙ АНАЛИЗ КОНЕЧНО МАЛЫХ ВОЗМУЩЕНИЙ ЛИНЕЙНЫХ МОДЕЛЕЙ ПРИ КОМПЬЮТЕРНОМ МОДЕЛИРОВАНИИ

Алексей Волошин, Владимир Кудин

Аннотация: Выделяются два уровня моделирования – математический и машинный (компьютерный). Обсуждается проблема адекватности линейной аппроксимации математических и машинных моделей. Исследуется влияние конечно малых возмущений в элементах моделей на качество локализации области решений модели на основе методологии последовательного анализа вариантов и результатов проведенного вычислительного эксперимента с учетом уровня обусловленности систем и организации вычислений (длины мантиссы в представлении чисел).

Abstract: Two design level are selected – mathematical and machine. Adequacy problem of linear approximation of mathematical and machine models comes into question. Influence certainly of small indignations is probed in the models elements on localization quality of model area decisions on the basis of successive analysis variants methodology and results of the conducted calculable experiment taking into account the conditionality systems level and calculations organization (mantisi lengths are in numbers presentation).

Keywords: mathematical and machine design, linear approximation, small indignations, localization, successive analysis variants, conditionality systems.

ACM Classification Keywords: H.4.2 Information Systems Applications: Types of Systems: Decision Support.

Введение

Линейные модели (ЛМ) являются лишь первым приближением, локальной аппроксимацией, при исследовании процессов реального мира, который является в принципе нелинейным. Для хорошо обусловленных задач (даже в «гладкой» ситуации, а в некоторых случаях - и в «негладкой» [Кларк, 1988]) существующий математический аппарат (анализ бесконечно малых) является основным инструментом при моделировании процессов «неживой и живой природы» и позволяет оценить влияние малых возмущений параметров на свойства модели лишь локально. Однако, во многих случаях аппроксимирующие линейные модели описываются в классе систем плохо обусловленных (некорректных), в частности, систем линейных алгебраических уравнений (СЛАУ) с квадратной и сильно заполненной матрицей ограничений [Самарский, 1989; Метьюз, 2001; Деммель, 2001]. Даже при небольшой размерности модели и при отсутствии структурных особенностей в матрице ограничений зависимости решения от возмущений аппроксимируются недостаточно адекватно (локализируются с помощью эллипсов с осями существенно различной длины или параллелепипедами с существенно разными границами на переменные). В частности, при исследовании экономических процессов во многих случаях результат моделирования представляется в виде системы линейных алгебраических неравенств (СЛАН) или как линейная оптимизационная модель. Известно, что категория плохой обусловленности

является определяющий в построении зависимостей решений (и областей локализации) от возмущений в элементах модели [Самарский, 1989].

Опыт решения практических задач показывает, что характерной для математического моделирования является неадекватность модели и реального моделируемого процесса (упрощение явления, неточность задания параметров), а для вычислительного эксперимента - неадекватность математической и машинной модели (неустраняемые ошибки, погрешности дискретизации, погрешности метода, округлений, усечений, потери значимых цифр, при выполнении операций). Округление чисел с точностью до фиксированного значения при вычислениях связано с приближенным представлением чисел с конечной (усеченной) разрядной сеткой (числа с фиксированной и плавающей запятой). «Классическая» непрерывность гарантирует существование локальной как угодно малой окрестности решения (проблемы необходимости определения решения с любой степенью точности сейчас не обсуждаются). В случае компьютерной (дискретной) модели окрестность решения не может быть меньше заранее заданного числа, а значит, не является корректной локальной аппроксимацией. Ошибки округления для таких вычислений могут характеризоваться относительной погрешностью. В ходе выполнения основных арифметических операций ошибки округлений могут накапливаться (в большей или меньшей мере). Существуют определенные отличия в схеме накопления ошибок при выполнении вычислений с фиксированными и плавающими числами. В совокупности они суммируются как погрешности вычислений [Самарский, 1989; Метьюз, 2001; Деммель, 2001].

Алгоритмы, которые реализуют конкретные методы, могут быть устойчивыми (погрешности в ходе вычислений накапливаются незначительно) и, соответственно, в противоположном случае - неустойчивыми. К алгоритмам, которые реализуют методы, можно предъявить две группы критериев: критерии адекватности моделей (дискретной машинной и математической непрерывной) и критерии сходимости вычислительных алгоритмов при увеличении числа уравнений в описании математической модели, быстродействия, точности вычислений. О проблемах неадекватности математической и машинной модели дает представление «феномен Рунге» [Метьюз, 2001], в котором при интерполировании Лагранжа погрешность аппроксимации растет при увеличении количества узлов (рис. 1).

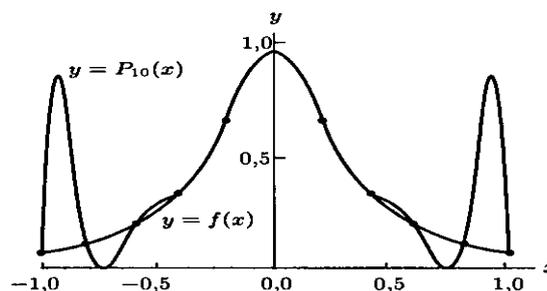


Рис. 1 Интерполяционный полином для функции $f(x) = 1/(1 + 12x^2)$, построенный по 11 равноудаленным узлам на интервале $[-1; 1]$

Со сходимостью алгоритмов связано такое понятие как корректность численного метода (непрерывная зависимость от входных данных, равномерная относительно количества уравнений). В частности,

корректность, как устойчивость при изменениях входных данных в линейных системах, характеризуется как устойчивость от входных данных правых частей (обусловленность) и от всех элементов модели. Мера корректности задачи количественно описывается числом обусловленности [Метьюз, 2001]. Следует отметить, что категория корректности линейной системы (обусловленность задачи) проявляется на каждом шаге итерационного процесса.

В непрерывном случае основным критерием сходимости является монотонность и ограниченность оценок решения, в случае компьютерной модели «численная» монотонность и ограниченность «компьютерного» времени не гарантирует отсутствия «выбросов» в «реальном» будущем. Так, для функции $f(t) = e^{1/e} - |t-T|^{1/e}, t \neq T, f(t) = e^{1/e}, t = T$ (e - константа Эйлера) при достаточно большом T (например, равно 10^{100}) «численная» монотонность на промежутке $[0; B]$ (в частности, на промежутке $[A; B]$, $A < B$, реализуется «машинный нуль» при вычислении значений функции) дает «практические» основания считать, что $f(t) \rightarrow 0$ при $t \rightarrow \infty$, хотя это «детерминировано» не так! (см. рис. 2, где e - константа Эйлера, $a = e^{1/e} - 1, b = e^{1/e}, A < B, B = T - e, C = T + e$).

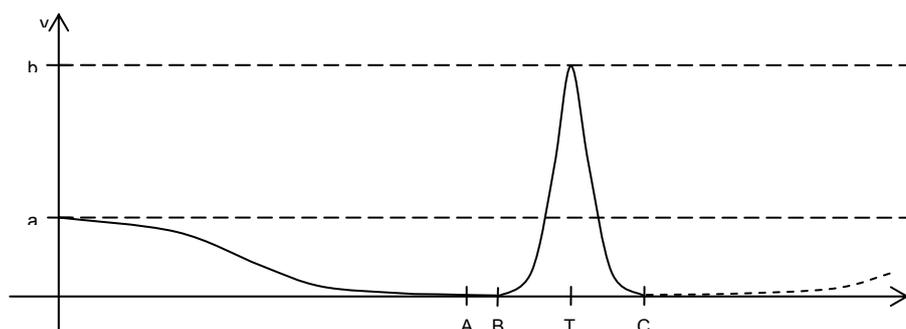


Рис. 2.

В условиях некорректности (за Адамаром) [Самарский, 1989] малые неточности представления математической модели, в частности, на уровне вычислительного алгоритма, могут существенно влиять на количественные и качественные характеристики получаемого решения при использовании конкретного метода (алгоритма) [Волошин, Кудин, 2010; Богаенко, Кудин, 2010]. Такие неточности чаще всего обусловлены ограниченностью длины мантииссы при представлении чисел с плавающей запятой (ошибки усечения, округления). Невзирая на наличие ЭВМ с эффективной организацией операции округления, в полной мере избежать их, или улучшить известные теоретические оценки, не удастся. Важно отметить, что проведение вычислений с низкой точностью ("грубо") сглаживает ("прячет") некоторые важные детали. В таких случаях создается иллюзия неединственности решения (неполный ранг) для СЛАУ или совместности области решений для СЛАН. Высокоточные вычисления раскрывают детали и чаще всего "дают" единственное решение (полный ранг матрицы) для СЛАУ и несовместимость для СЛАН. При проведении анализа вычислений с разной степенью точности важно согласовывать (согласовывать между собой) свойства грубых и высокоточных вычислений.

Метод базисных матриц

В данной работе проведено исследование зависимости качества локализации области принадлежности решений (или аппроксимации) от величины возмущений в элементах модели, от числа обусловленности (как определяющего фактора) и организации вычислений (от длины мантиссы в представлении чисел) для некоторых сравнительно простых задач. Акцентируется внимание на сложности построения аппроксимирующих (локализующих) множество решений. В основу исследований была положена методология последовательного анализа вариантов (ПАВ) [Волошин, 1987] и метод базисных матриц (МБМ), как его конкретная реализация [Волошин, Кудин, 2009, 2010]. МБМ позволяет учитывать влияние конечно малых возмущений на решения СЛАУ. В отличие от классических итерационных методов (например, симплекс-метода) МБМ находит решение в два этапа: 1) выделение базисной матрицы; 2) нахождение соответствующего выделенной базисной матрице решения СЛАУ и анализ обусловленности системы, что позволяет предусматривать ситуацию накопления ошибок.

В работе проводится:

- анализ свойств плохо обусловленных задач малой размерности;
- анализ типичных моделей средней размерности;
- исследование свойств алгоритмов на основе базисного метода;
- расчеты на основе вычислительных алгоритмов (без процедуры уточнения, с одно- и двухстадийной процедурой уточнения) при использовании разных типов данных (с плавающей запятой с мантиссой размерностью 64, 128, 256 бит) для СЛАУ с разными числами обусловленности;
- построены функциональные зависимости (интерполяционные многочлены) точности решения от используемых типов данных, алгоритмов и числа обусловленности системы.

Постановка задачи

Дана линейная модель:

$$Au = C \quad (1)$$

где матрица A (со строками a_1, a_2, \dots, a_m , $a_j = (a_{j1}, a_{j2}, \dots, a_{jm})$, $j = \overline{1, m}$) - квадратная размерностью $(m \times m)$, у которой вектор переменных $u = (u_1, u_2, \dots, u_m)^T$ и вектор ограничений $C = (c_1, c_2, \dots, c_m)^T$ имеют размерность m . Матрицы ограничений СЛАУ таких задач характеризуются в общем случае большой размерностью, сильной заполненностью и плохой обусловленностью.

В основу МБМ положена идея базисной матрицы. Базисные матрицы в процессе итераций изменяются последовательным замещением строк некоторой вспомогательной базисной матрицы (вспомогательной СЛАУ) строками ограничений модели (1).

Введем в рассмотрение векторы $a_{i1}, a_{i2}, \dots, a_{im}$ - нормали ограничений (будем называть их строками), $a_j u^T \leq c_j$, $j \in J_0$, которые образуют матрицу A_0 , где $J_0 = \{i_1, i_2, \dots, i_m\}$ - индексы ограничений.

Определение 1. Квадратную матрицу $A_{\bar{\sigma}}$, состоящую из m линейно независимых строк некоторой вспомогательной СЛАУ, будем называть искусственной базисной, а решение u_0 соответствующей ей системы уравнений $A_{\bar{\sigma}}u = C^0$, где $C^0 = (c_{i_1}, c_{i_2}, \dots, c_{i_m})^T$ - искусственным базисным решением.

Определение 2. Две базисные матрицы, отличающиеся одной строкой, будем называть смежными.

Пусть e_{ri} - элементы матрицы $A_{\bar{\sigma}}^{-1}$, обратной к $A_{\bar{\sigma}}$, $e_k = (A_{\bar{\sigma}}^{-1})_k - k$ - ый столбец обратной матрицы, $u_0 = (u_{01}, u_{02}, \dots, u_{0m})^T$ - базисное решение, $\alpha_r = (\alpha_{r1}, \alpha_{r2}, \dots, \alpha_{rm})$ - вектор разложения нормали ограничения $a_r u \leq c_r$ по строкам базисной матрицы $A_{\bar{\sigma}}$, $\Delta_r = a_r u_0 - c_r$ - невязка r -го ограничения для u_0 . Для строки a_l (нормаль ограничений $a_l u \leq c_l$, $l \notin J_{\bar{\sigma}}$), $\alpha_l = (\alpha_{l1}, \alpha_{l2}, \dots, \alpha_{lm})$ - вектор разложения строки a_l по строкам базисной матрицы $A_{\bar{\sigma}}$, вектор которой находится из соотношения $a_l = \alpha_l A_{\bar{\sigma}}$.

Между коэффициентами разложения нормалей ограничений (1) по строкам искусственной базисной матрицы, элементами обратных матриц, базисными решениями, невязками ограничений (1) в двух смежных базисных матрицах при замене k -ой строки в базисной матрице $A_{\bar{\sigma}}$ строкой a_l имеют место соотношения [Волошин, Кудин, 2010]:

$$\bar{\alpha}_{rk} = \frac{\alpha_{rk}}{\alpha_{lk}}, \quad \bar{\alpha}_{ri} = \alpha_{ri} - \frac{\alpha_{rk}}{\alpha_{lk}} \alpha_{li}, \quad r = \overline{1, m}; \quad i = \overline{1, m}; \quad i \neq k; \quad (2)$$

$$\bar{e}_{rk} = \frac{e_{rk}}{\alpha_{lk}}, \quad \bar{e}_{ri} = e_{ri} - \frac{e_{rk}}{\alpha_{lk}} \alpha_{li}, \quad r = \overline{1, m}; \quad i = \overline{1, m}; \quad i \neq k; \quad (3)$$

$$\bar{u}_{0j} = u_{0j} - \frac{e_{jk}}{\alpha_{lk}} \Delta_l, \quad j = \overline{1, m}, \quad (4)$$

$$\bar{\Delta}_k = -\frac{\Delta_l}{\alpha_{lk}}, \quad \bar{\Delta}_r = \Delta_r - \frac{\alpha_{rk}}{\alpha_{lk}} \Delta_l, \quad r = \overline{1, n}; \quad r \neq k; \quad (5)$$

причем условием невырожденности базисной матрицы при замещении строкой a_l k -ой строки базисной матрицы $A_{\bar{\sigma}}$ является выполнение неравенства $\alpha_{lk} \neq 0$.

Для существования единственного решения (1) необходимо и достаточно, чтобы $\alpha_{lk}^{(i)} \neq 0$, $i = \overline{1, m}$, где $\alpha_{lk}^{(i)}$ - ведущие элементы операции замещения строк базисной матрицы нормальными ограничениями (1).

Матрица A системы (1) является невырожденной, если $\alpha_{lk}^{(i)} \neq 0$, $i = \overline{1, m}$.

В векторном виде формулы (3) можно представить как $\bar{e}_k = \frac{e_k}{\alpha_{lk}}$, $\bar{e}_i = e_i - \frac{e_k}{\alpha_{lk}} \alpha_{li}$, $i = \overline{1, m}; \quad i \neq k$.

Ниже приведены основные стадии алгоритмической схемы нахождения величины машинного ранга, базисной матрицы и решения (1) на основе известных свойств тривиальной СЛАУ той же размерности:

- проводим симплексные итерации по замещению строк тривиальной базисной матрицы с известными элементами метода строками ограничений системы (1), согласно с соотношениями (2)-(5);
- проверяем выполнение условий невырожденности на каждой итерации;
- находим соответствующие элементы метода: вектора разложения по строкам базисных матриц ограничений системы (1), обратную базисную матрицу, невязки ограничений, искусственные базисные решения $u_0^{(r)}$, где k - номер итерации;
- контролируем количество итераций k замещения строк вспомогательной системы строками основной системы (1), для которых выполняются условия невырожденности. Если количество итераций, для которых $\alpha_{kk}^{(k)} \neq 0$, равно m , единственное решение находится из соотношения: $A_0^{-1} \cdot C = u^0$. Если нет, то при выполнении условия $k < m$ для СЛАУ (1) не выполняется условие единственности решения. В этом случае модель требует дальнейшего анализа.

Вычислительный процесс пересчета обратной матрицы представляет собой проведение двух этапов матричных операций:

- деления ведущего k -го столбца $e_k = (A_0^{-1})_k$ на значение ведущего элемента $\alpha_{kk}^{(k)} \neq 0$ $e_k^{k+1} = e_k^k / \alpha_{kk}^{(k)}$;
- расчета на k -й итерации для i -го столбца обратной матрицы ($i \neq k, i \in I$) нового столбца $e_i^{k+1} = e_i^k - e_k^{k+1} \times \alpha_{ki}$, при этом качество локализации области возмущений решений (или аппроксимации) существенно зависит от числа обусловленности (как определяющего фактора) и от организации вычислений (от длины мантиссы в подаче чисел).

Вычислительный эксперимент по анализу малых возмущений ЛМ малой размерности

Влияние малых возмущений в "малых" ЛМ на качество локализации решений иллюстрируется на тестовом исследовании, которое структурно было организовано как многошаговая алгоритмическая процедура: построения последовательности модельных задач малой размерности со свойством плохой обусловленности с некоторого шага; нахождения на основе МБМ элементов метода; вычисления чисел обусловленности; графическое представление основных функциональных зависимостей.

Эксперимент демонстрирует количественно-качественные связи основных параметров модели и МБМ.

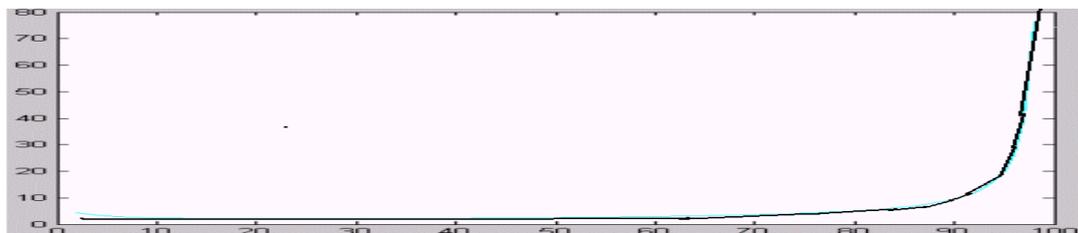


Рис. 3. Зависимость значений норм ведущих столбцов от номера итерации

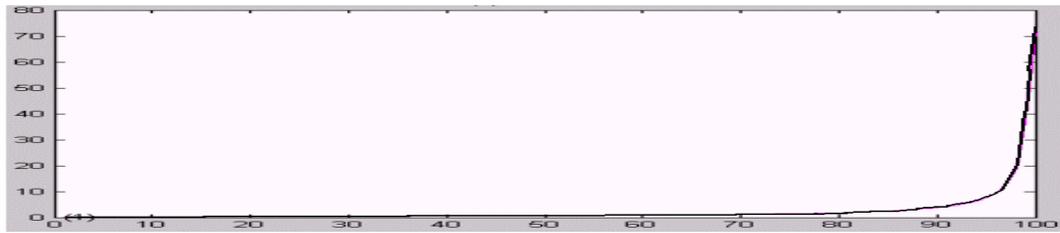


Рис. 4. Зависимость значений чисел обусловленности от номера итерации



Рис. 5. Зависимость значений ведущих элементов от номера итерации

Ниже приведены полученные в результате эксперимента графические образы множеств принадлежности решений (локализация).

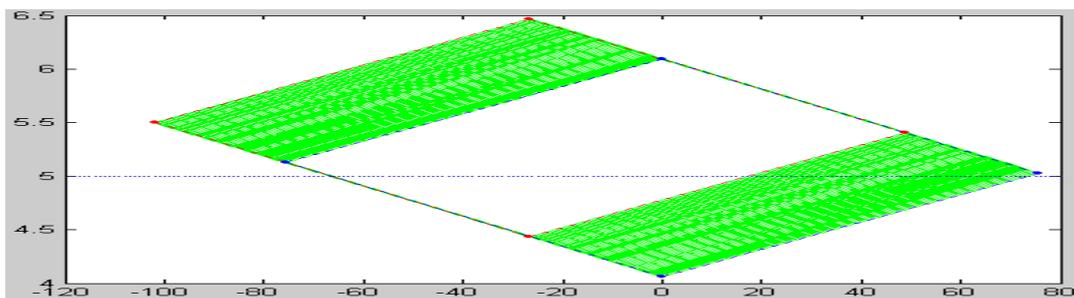


Рис. 6. Ромбовидные фигуры оценочных множеств принадлежности решений на итерациях метода (оси координат — компоненты вектора решений) при возмущениях типа „параллельный перенос”

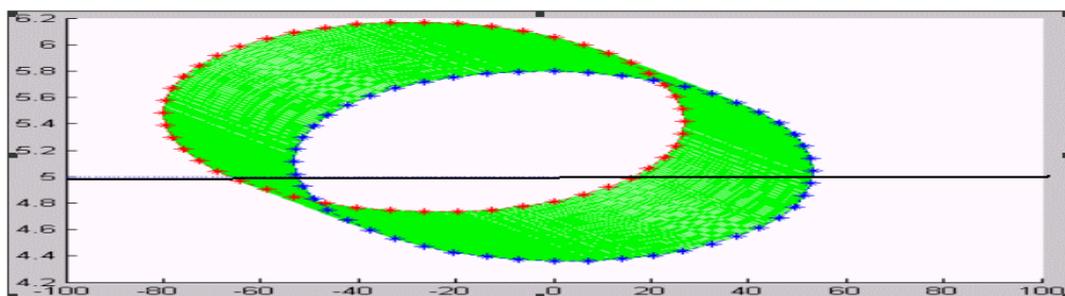


Рис. 7. Эллипсы принадлежности решений на итерациях метода при возмущениях вектора правых частей ограничений из сферы малого диаметра

Графические образы эллипсообразных фигур и ромбов (локализация) приведены в непропорциональном отношении минимума и максимума осей абсцисс и ординат. Эллипсообразные фигуры и ромбы образов возмущения множества решений (при малых возмущениях элементов модели) являются вытянутыми

относительно многообразия решений граничной системы неполного ранга. Фигуры структурно включают в себя "большую ось" (диагональ) и "малую". Большая ось отвечает большому собственному числу по оси абсцисс, малая (с малыми значениями диагонали по оси ординат) - отвечает малому собственному числу.

Рис.6, 7 в определенной степени подтверждает гипотезу, высказанную одним из соавторов статьи, о справедливости принципа неопределенности Гейзенберга при компьютерном исследовании влияния конечно малых возмущений – уменьшение неопределенности («улучшение» локализации) по одной переменной приводит к увеличению неопределенности («ухудшению» локализации) по другой.

Вычислительный эксперимент по анализу малых возмущений в ЛМ средней размерности

Для вычислительного эксперимента были избраны три процедуры МБМ: без уточнения решения (0), с одностадийным (1) и двух стадийным (2) уточнением [Волошин, Кудин, 2010], применялись разные типы данных: числа с плавающей запятой (двойной точности (Double), 128-битовые числа (Dd) и 256- битовые числа (Qd), а также вычисления в точных числах. Стоит отметить, что на одной и той же вычислительной платформе отношение быстродействия алгоритмов, что используют разные типы данных, будет постоянным. Данные вычислительных экспериментов показали, что на избранной для тестирования платформе, алгоритм, который использует 128-битовые числа, был в ~35 раз медленнее, чем алгоритм, который использует 64-битовые числа, а в случае 256-битовых чисел - в~450 раз медленнее. Такое существенное замедление объясняется тем, что операции с числами с плавающей запятой большого размера не реализованы аппаратно, в отличие от операций над 64-битовыми числами.

Для построения эвристических зависимостей между точностью решения и числом обусловленности была проведенная серия вычислительных экспериментов. Использованная для тестирования аппаратная платформа - процессор AMD Athlon64 с реальной тактовой частотой 1.8Ghz, 512Mb оперативной памяти. Проводилось решение СЛАУ размерностью 256x256 различными алгоритмами МБМ. В качестве критерия точности бралась точность машинного решения u_0 системы (1) по сравнению с аналитическим (точным) решением $u=(1,0,...,0)$, $\varepsilon = \|u_0 - (1,0,...,0)\|$. Полученные экспериментальные данные приведены ниже в табл.1. На основе этих данных (и данных при других размерностях) построены зависимости точности решения от числа обусловленности (Рис.9), 0, +1, +2 - количество итераций уточнения, $\log_{10} c$ - порядок числа обусловленности системы, $\log_{10} \varepsilon$ - порядок точности решения.

Таблица 1. Зависимость порядка точности решения ε ($\log_{10} \varepsilon$) от числа обусловленности ($m=256$)

$\log_{10} c$	6	7	8	9	10	11	12
double+0	15.81	12.29	8.89	2.75			
dd+0	46.16	42.45	39.07	34.66	32.50	27.71	22.33
qd+0	113.35	107.94	104.76	100.43	96.25	93.69	86.65
double+1	12.40	10.55	5.37	2.70			
dd+1	45.92	43.43	36.86	31.96	29.12	26.33	22.43
qd+1	109.58	108.57	103.25	95.93	93.22	91.00	87.99
double+2	10.85	8.55	6.75	2.27			
dd+2	44.41	42.15	36.89	32.16	31.15	26.44	21.90
qd+2	101.21	106.73	107.70	100.30	93.89	87.59	86.52

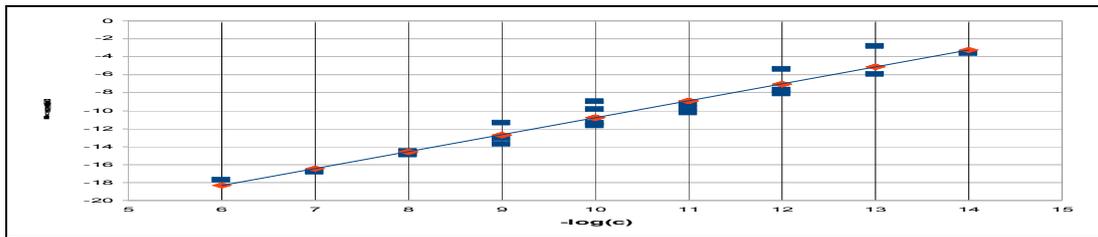


Рис. 8. Зависимость порядка точности ϵ ($\log_{10} \epsilon$) от числа обусловленности.

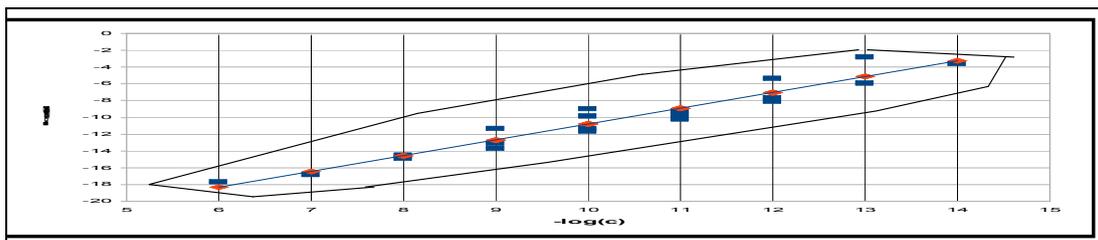


Рис. 9. Вариант кусочно-линейной аппроксимации (локализации) области значений порядка точности ϵ ($\log_{10} \epsilon$) в системе координат значений порядка точности и порядка числа обусловленности.

Вычислительный эксперимент установил близкие к линейной зависимости: 1) точности решения (и обращения матрицы) от числа обусловленности (при фиксированном типе данных и алгоритме) и размерности модели; 2) точности решения от типа данных (при фиксированном числе обусловленности и алгоритме), что позволяет строить: 1) интерполяционные многочлены - зависимости точности решения от обусловленности; 2) аппроксимирующие множества принадлежности точности решений (в том числе, эллипсы минимальной площади); 3) экстраполяционные зависимости (для точности решений алгоритмов) от числа обусловленности.

Методология ПАВ [Волошин, 1987] направлена на проведение исследований по установлению статуса компонент модели, анализа включений (исключений) компонент модели на ограниченность и замкнутость (то есть на качество локализации). Как следует из результатов вычислительного эксперимента, даже незначительные количественные изменения в компонентах таких моделей могут качественно изменить статус компонент модели и, как следствие, структуру множества решений задачи. При исследовании устойчивости задачи важно выявить количественную меру изменений исходных данных (корректность) при которых это свойство хранится (теряется) и значения параметров, при которых происходят качественные изменения свойств системы. "Поймать и оценить" такие структурные зависимости качественных параметров от количественных (в малом) между практической задачей математическим и машинным описанием модели, например, на языке " $(\epsilon - \delta)$ " - оценки отклонения или изменения решений - от изменения данных в пределах (категория устойчивости) не только трудно, но, чаще всего, и невозможно даже для плохо обусловленных задач малой размерности. Определять и устанавливать количественное значение плохой обусловленности системы сама по себе тоже сложная задача [Волошин, Кудин, 2009, 2010].

В качестве вывода можно отметить необходимость включения в процесс проведения численного эксперимента лица, принимающего решение, для анализа влияния различных стратегий проведения

вычислений (типов данных, алгоритмов, уровня обусловленности системы) на основные параметры решения (точность самого решения и обращения матриц, быстродействие, объемы вычислений), разработки эффективных вычислительных схем методов и алгоритмов, оценивающих обусловленность системы в ходе эксперимента по отмеченным критериям на моделях различной размерности.

Заключение

Данная работа продолжает (и в определенном смысле обобщает) исследования влияния конечно малых возмущений линейных моделей [Волошин, Кудин, 2009, 2010] и акцентирует внимание на необходимости разработки «анализа конечно малых» как инструментария анализа компьютерных моделей.

Благодарности

Статья частично финансирована из проекта ITHEA XXI Института Информационных теорий и Приложений FOI ITHEA и Консорциума FOI Bulgaria (www.itea.org, www.foibg.com).

Список литературы

- [Кларк, 1988] Кларк Ф. Оптимизация и негладкий анализ.-М.:Наука, 1988.-280 с.
- [Самарский, 1989] Самарский А.А., Гулин А.Г. Численные методы.- М.: Наука, 1989. - 432с.
- [Метьюз, 2001] Метьюз Д.Г., Финк К.Д. Численные методы.-Москва-С.-Петербург-Киев: Вильямс, 2001.-703с.
- [Деммель, 2001] Деммель Дж. Вычислительная линейная алгебра. Теория и приложение - М.: Мир, 2001,-430с.
- [Волошин, 1987] Волошин А.Ф. Метод локализации области оптимума в задачах математического программирования // Доклады АН СССР, 1987, том 293, №3.-С.549-553.
- [Кудин, 2002] Кудин В.И. Застосування методу базисних матриць при дослідженні властивостей лінійної системи // Вісник Київського університету. Серія фіз.-мат. науки. - 2002. - 2. - С. 56-61.
- [Волошин, Кудин, 2009] Волошин А., Кудин В., Богаенко В. Анализ малых возмущений линейных экономико-математических моделей // Information Science & Computing, N 10, Vol. 3, SOFIA, 2009.- P. 67-73.
- [Волошин, Кудин, 2010] Волошин А, Кудин В., Кудин Г. Методы анализа малых возмущений линейных моделей // Natural and Artificial Intelligence, ITHEA, Sofia, 2010. - P. 41-47.
- [Кудин, 2010] Кудин В., Богаенко В. О принятии решений при анализе малых возмущений линейных моделей // Information Models of Knowledge, ITHEA, Kiev-Sofia, 2010. - P. 226-231.

Информация об авторах

Алексей Ф. Волошин, Киевский национальный университет имени Тараса Шевченко, факультет кибернетики, Украина, д. т. н., профессор, E-mail: olvoloshyn@ukr.net

Владимир И. Кудин, Киевский национальный университет имени Тараса Шевченко, факультет кибернетики, Украина, д. т. н., с. н. с., E-mail: V.I.Kudin@mail.ru

К ПОСТРОЕНИЮ МОДЕЛИ «ДЕМОКРАТИЧЕСКОГО КАПИТАЛИЗМА»

Алексей Волошин, Анна Сидорук

Аннотация: В работе предлагается экономико-математическая модель «демократического» капитализма, в которой все домохозяйства являются владельцами некоторого капитала. Предлагаются формализованные процедуры принятия решений для перехода к «демократическому» капитализму и методы, удовлетворяющие принципам «участия», «распределения» и «ограничения». Методы распределения затрат обобщаются на случай принятия решений в нечетких условиях..

Keywords: модель демократического капитализма, принципы участия, распределения, ограничения; модель смешанной экономики; нечеткие методы распределения затрат.

ACM Classification Keywords: 1.Computing Methodologies. 1.6. Simulation and modeling. 1.6.5. Model Development - Modeling Methodologies.

Введение

В 1958 году американским экономистом Луис Келсо в соавторстве с философом Мортимером Адлером была опубликована монография под названием "Манифест Капитализма", которая в 1959 году была переведена на русский язык и разослана по «специальному списку» [Келсо, 2004]. Авторы монографии утверждают, что экономика современных им Соединенных Штатов не является капиталистической, поскольку собственниками капитала (капиталистами) является только незначительная часть населения страны, остальные владеют лишь собственным трудом как средством производства.

Высокий уровень жизни в Соединенных Штатах авторы объясняют, в первую очередь, высокой производительностью капитала. Слабость «смешанной» капиталистической экономики, в которой рыночные механизмы дополняются механизмами государственного регулирования, заключается в том, что, с одной стороны, право собственности на часть капитала, который производит около 90 процентов богатства, сосредоточено в руках примерно 5 процентов семей, в то время, как более 70 процентов дохода, представляющего собой произведенное богатство, распределяется по труду.

Капиталистическое разрешение этих противоречий в смешанной экономике заключается в обеспечении сбалансированного участия домохозяйств в производстве за счет диффузии собственности на капитал. Так как бремя производства переходит от труда к капиталу, то и способ участия в производстве должен меняться от «трудового» к «капиталистическому».

Главной же причиной высококонцентрированной собственности на капитал является несоответствие между увеличением производительности капитала и почти постоянным уровнем производительности труда (за исключением административного и «высокоинтеллектуального»). Вторая причина представляет собой хорошо известное экономическое явление: чем выше доход, тем большая его доля инвестируется в дальнейшее формирование капитала.

Авторы указывают на необходимость перехода к новой модели экономики - «Демократическому капитализму» (ДК), при которой все домохозяйства являются владельцами некоторого капитала. Принципы, положенные в основу предлагаемой экономической системы, следующие:

- 1) участие всех домохозяйств в производстве путем владения определенным капиталом,
- 2) перераспределение капитала для преобразования всех потребителей в капиталистов,
- 3) ограничения на уровень владения капиталом.

В соответствии с вышеизложенными принципами, авторы предлагают следующие меры (математически не формализованные), способствующие их воплощению: 1) увеличение числа владельцев акций в промышленных предприятиях; 2) изменение ставок налога на наследство и дарованное имущество с целью поощрения создания новых капиталистов и препятствия чрезмерной концентрации капитала; 3) отмена налога на доходы корпораций; 4) принятие закона, обязывающего «зрелые» корпорации выплачивать своим акционерам 100 процентов чистого дохода; 5) признание правительством обязательства гарантировать всем семьям, занятым в экономике, возможность участия в производстве благ в масштабах, позволяющих получать жизнеобеспечивающий доход; 6) установление системы льгот для поощрения накопления капитала теми семьями, капитал которых не дает им жизнеобеспечивающего дохода; 7) использование финансовой системы государства для поощрения новых капиталобразований при условии, что таковые не способствуют высокой концентрации капитала.

Модель экономики Демократического Капитализма

За прошедшие 50 лет капиталистическая система не стала «более капиталистической». Особенно это касается возникших в последние 20 лет в результате распада Советского Союза рыночных экономических систем, которые отвечают «дикому» (олигархическому) капитализму. Поэтому построение математических моделей ДК актуально, как никогда.

В данной работе предпринимается попытка формализации в рамках математической модели «смешанной экономики» [Пономаренко, 2004] принципов и конкретных предложений авторов ДК. Выполнение принципов участия обеспечивается путем предоставления домохозяйствам субсидий в виде пакетов акций предприятий. Выполнение принципа перераспределения — путем установления системы налогов и субсидий и использования методов дележа. Выполнение принципа ограничения - путем определения уровня «монополистического» дохода. При этом разница между доходом потребителя после уплаты налогов и «монополистическим» уровнем дохода подлежит повторному налогообложению. Ставка повторного налогообложения определяется на основе балансового уравнения бюджета.

Особое внимание в данной работе уделяется проблемам дележа (распределения), поскольку сама по себе проблема дележа (земли, энергоресурсов, финансов, сфер влияния, квот на выбросы парниковых газов и т.п.) в настоящее время является одной из актуальных проблем человечества, лежащей в основе конфликтов (межгосударственных, внутригосударственных, межличностных и т.д.). В частности, проблема дележа возникает при исследовании рассматриваемой в данной работе модели ДК.

Следует отметить, что в известной математической модели Эрроу - Дебре [Пономаренко, 2004] наличие ненулевого начального капитала является одним из основных принципов (одной из основных аксиом).

В качестве базовой модели используется модель «дезагрегированной смешанной» экономики [Пономаренко, 2004] с N потребителями и E предприятиями. Модель модифицируется с целью удовлетворения основным принципам, которые предлагают авторы «Капиталистического Манифеста».

Принцип участия. Доходы домохозяйств формируются посредством участия в производстве в качестве акционеров: $K_n = \lambda \sum_{e=1}^E \alpha_{n,e} \pi_e$ - доход потребителя без учета налогов, тут $\alpha_{n,e}$ - часть акций потребителя n в предприятии e , $\sum_{n=1}^N \alpha_{n,e} = 1$, $(1 - \lambda)$ - ставка налогообложения, π_e - доход предприятия e , $\pi = \sum_{e=1}^E \pi_e$ - суммарный доход. Доходы корпораций не подлежат налогообложению и 100% чистой прибыли выплачивается владельцам акций в конце определенного периода (чаще всего – календарного года). Потребители, которые являются владельцами капитала, посредством покупки акций тех или иных корпораций «голосуют» за управленческую деятельность руководства компаний.

Принцип распределения. Определяется минимальный уровень дохода μd , где $d = \frac{\sum_{n=1}^N K_n}{N} = \frac{\lambda \sum_{i=1}^N \sum_{e=1}^E \alpha_{n,e} \pi_e}{N} = \frac{\lambda \pi}{N}$ - средняя величина дохода домохозяйства, и при $\mu d \geq K_n$ n -й потребитель получает субсидию размером $\mu d - K_n$. Путем начисления субсидий домохозяйствам, которые не принимают участия в производстве по причине отсутствия прав владения на капитал, будет обеспечиваться возможность их участия в производстве. При этом государство должно поощрять домохозяйства вкладывать эти средства в накопления. То есть, субсидия $\max\{0, \mu d - K_n\}$ расходуется для накопления капитала, который станет источником будущего дохода предприятия.

Принцип ограничения. Определяется уровень ld такой, что при $K_n \geq ld$ доход K_n считается «монополистическим», то есть лишаящим других потребителей возможности принимать участие в производстве и получать жизнеобеспечивающий доход. В этом случае разница $K_n - ld$ вторично облагается налогом по ставке λ_1 , $\lambda < \lambda_1 \leq 1$.

Таким образом, прибыль домохозяйства составляет:

$$I_n = K_n + \max\{0, \mu d - K_n\} - \lambda_1 \max\{0, K_n - ld\}.$$

Пусть c , $c \geq 0$, - расходы бюджета на обеспечение общественных благ (строительство дорог, образование, медицину), тогда параметры λ и λ_1 связаны соотношением:

$$\begin{aligned} c + \sum_{n=1}^N \max\{0, \mu d - K_n\} &= \sum_{n=1}^N \lambda \max\{0, K_n - ld\} + \sum_{n=1}^N (1 - \lambda) \sum_{e=1}^E \alpha_{n,e} \pi_e \\ c + \sum_{n=1}^N \max\{0, \mu \frac{\lambda \pi}{N} - \lambda \sum_{e=1}^E \alpha_{n,e} \pi_e\} &= \sum_{n=1}^N \lambda \max\{0, \lambda \sum_{e=1}^E \alpha_{n,e} \pi_e - l \frac{\lambda \pi}{N}\} + (1 - \lambda) \sum_{e=1}^E \pi_e \\ c + \lambda \sum_{n=1}^N \max\{0, \mu \frac{\pi}{N} - \sum_{e=1}^E \alpha_{n,e} \pi_e\} &= \lambda_1 \sum_{n=1}^N \max\{0, \sum_{e=1}^E \alpha_{n,e} \pi_e - l \frac{\pi}{N}\} + (1 - \lambda) \pi, \end{aligned}$$

из которого получаем значение λ_1 :

$$\lambda_1 = \frac{c + \lambda \sum_{n=1}^N \max\{0, \mu \frac{\pi}{N} - \sum_{e=1}^E \alpha_{n,e} \pi_e\} - (1 - \lambda) \pi}{\lambda \sum_{n=1}^N \max\{0, \sum_{e=1}^E \alpha_{n,e} \pi_e - l \frac{\pi}{N}\}}$$

Методы дележа

Определив основной налог, за счет которого формируется бюджет страны, модифицируем принципы дележа так, чтобы они удовлетворяли принципам экономической системы «справедливого» капитализма. Задача дележа формулируется следующим образом [Волошин, Мащенко, 2010]: коллективный объект имеет стоимость c , $c > 0$, денежных единиц и приносит доход b_i , $b_i \geq 0$, каждому пользователю $i = \overline{1, n}$.

Создание объекта эффективное: $\sum_{i=1}^n b_i > c$.

Задача состоит в справедливом распределении расходов между участниками. Различным понятиям принципа справедливости соответствуют различные методы дележа. При пропорциональном решении каждый участник платит налог, пропорциональный своему доходу $x_i = b_i \frac{c}{\sum_{j=1}^n b_j}$, $i = \overline{1, n}$.

Принцип равенства может воплощаться через механизмы выравнивания затрат или выравнивания чистой экономии на затратах. В первом случае каждый игрок платит $x_i = c/n$, во втором

$x_i = b_i - \left(\sum_{j=1}^n b_j - c \right) / n$. Очевидными недостатками двух последних методов является то, что решения,

полученные с их применением, могут не удовлетворять условию $0 \leq x_i \leq b_i$, $i = \overline{1, n}$.

Модификациями данных методов, которые удовлетворяют этим ограничениям, являются уровневый и подушный налоги, предложенные Янгом [Мулен, 1991].

Уровневый налог есть распределение (единственное) затрат (x_1, \dots, x_n) , являющееся решением задачи:

$$(x_1, \dots, x_n) \in A = \left\{ (y_1, \dots, y_n) \mid \sum_{i=1}^n y_i = c, 0 \leq y_i \leq b_i, \forall i \right\},$$

$$(b_1 - x_1, \dots, b_n - x_n) LM (b_1 - y_1, \dots, b_n - y_n), \forall y \in A,$$

где LM – лексиминный порядок на E^N .

Подушный налог есть распределение (единственное) затрат $(x_1, \dots, x_n) \in A$, являющееся решением задачи: $(x_1, \dots, x_n) LM (y_1, \dots, y_n)$, $\forall y \in A$.

Фазификация методов дележа

Необходимо отметить, что решение задач дележа приходится искать до того, как становятся известны «точные» доходы участников и коллективные затраты, поэтому в качестве исходных данных используются прогнозные значения или экспертные оценки. В данной работе они представляются с помощью нечетких треугольных чисел:

$$c = (c_1, \bar{c}, c_2), b_i = (b_{i1}, \bar{b}_i, b_{i2}), i = \overline{1, n}.$$

Пусть $\mu_{\geq}(x, y)$ - линейный порядок на числовой прямой, который представляется в виде нечеткого, рефлексивного, антисимметрического, транзитивного и связного бинарного отношения.

При реализации основных методов дележа, затраты каждого из участников будут составлять:

$$\text{Выравнивание затрат: } x_i = \frac{c}{n} = \left(\frac{c_1}{n}, \frac{\bar{c}}{n}, \frac{c_2}{n} \right), i = \overline{1, n};$$

$$\text{Выравнивание прибыли: } b_i - x_i = b_j - x_j = \frac{b-c}{n} = \left(\frac{b_1-c_2}{n}, \frac{\bar{b}-\bar{c}}{n}, \frac{b_2-c_1}{n} \right), i = \overline{1, n};$$

где b - сумарный доход всех участников, $b = (b_1, \bar{b}, b_2) = (\sum_{i=1}^n b_{i1}, \sum_{i=1}^n \bar{b}_i, \sum_{i=1}^n b_{i2})$.

$$\text{Пропорциональный дележ: } x_i = b_i \frac{c}{b} \approx \left(b_{i1} \cdot \frac{c_1}{b_1}, \bar{b}_i \cdot \frac{\bar{c}}{\bar{b}}, b_{i2} \cdot \frac{c_2}{b_2} \right), i = \overline{1, n}.$$

Чем ниже доход игрока, тем в большей степени на его благосостояние влияют отклонения от точки, на которой функция принадлежности принимает максимальное значение. Для ослабления этого влияния налагается условие: $|\bar{x}_i - x_{ij}| \leq \frac{\bar{b}_i}{100} \cdot \gamma \quad j = 1, 2; \quad 0 \leq \gamma \leq 100$

Тогда скорректированные значения затрат принимают вид:

$$x' = (\max\{x_{i1}, \bar{x}_i - \frac{\bar{b}_i}{100} \cdot \gamma\}, \bar{x}_i, \min\{x_{i2}, \bar{x}_i + \frac{\bar{b}_i}{100} \cdot \gamma\})$$

Рассмотрим уровневый и подушный налоги для случая нечетких входных данных. Допустимое множество решений задачи имеет вид:

$$Y(N, t, x) = \{y \in R^N \mid \sum_N y_i = c \wedge 0 \leq y_i \leq b_i \forall i\}$$

Здесь элементы множества $Y(N, t, x)$ – векторы (y_1, \dots, y_n) с компонентами, которые являются нечеткими треугольными числами: $y_i = (y_{i1}, \bar{y}_i, y_{i2})$.

Опишем функцию принадлежности вектора y нечеткому множеству допустимых решений $Y(N, t, x)$.

Для определения степени выполнения неравенств $0 \leq y_i \leq b_i, i = \overline{1, n}$, и равенства $\sum_{i=1}^n y_i = c$, запишем выражения, описывающие функции принадлежности (обобщенных по отношению $\mu_{\geq}(x, y)$) отношений предпочтения и равенства:

$$\eta_{\geq}(x, y) = \sup_{a, b \in R} \min\{\mu_x(a), \mu_y(b), \mu_{\geq}(a, b)\} \quad \eta_{=} (x, y) = \sup_{a, b \in R} \min\{\mu_x(a), \mu_y(b), \mu_{\geq}(a, b)\}$$

где $\mu_x(a)$ - функция принадлежности элементов нечеткому множеству X . Таким образом, функция принадлежности элементов нечеткому множеству допустимых альтернатив $Y(N, t, x)$ имеет вид:

$$\mu_Y(y) = \min \left\{ \eta_{=} \left(\sum_{i=1}^n y_i, c \right), \tilde{\eta}_{\geq}(y_1, 0), \eta_{\geq}(b_1, y_1), \dots, \tilde{\eta}_{\geq}(y_n, 0), \eta_{\geq}(b_n, y_n) \right\}$$

Следующим этапом в процессе решения задачи является построение лексиминного отношения предпочтения на множестве векторов из A . Упорядочить компоненты векторов за неубыванием позволяет следующее утверждение: если отношение предпочтения представляет собой линейный порядок на множестве Y , тогда индуцированное им отношение предпочтения является линейным порядком на классе всех нормальных нечетких подмножеств Y [Орловский, 1981].

Исходя из определения лексиминного порядка, вектор x лексиминно предпочтительней вектора y , если выполняется одно из следующих условий:

1) $x_{(1)} \geq y_{(1)}$ ($y_{(i)}$ – i -я компонента вектора y после упорядочения);

2) $x_{(1)} = y_{(1)}, x_{(2)} \geq y_{(2)}$;

n) $x_{(1)} = y_{(1)}, x_{(2)} = y_{(2)} \dots x_{(n-1)} = y_{(n-1)}, x_{(n)} \geq y_{(n)}$.

Лексиминный порядок на множестве векторов отображает тот факт, что каждая компонента вектора значительно важнее последующей. Для "ослабления" этого условия (это целесообразно сделать, так как операции выполняются над нечеткими числами) предлагается определять лексиминный порядок на множестве векторов из A . Таким образом: вектор x предпочтительнее вектора y по α -степенному лексиминному порядку с мерой β , если выполняется одно из следующих условий:

$$\eta_{=}(x_{(1)}, y_{(1)}) < \alpha, \eta_{\geq}(x_{(1)}, y_{(1)}) = \beta;$$

$$\eta_{=}(x_{(1)}, y_{(1)}) \geq \alpha, \eta_{=}(x_{(2)}, y_{(2)}) < \alpha, \eta_{\geq}(x_{(2)}, y_{(2)}) = \beta;$$

$$\eta_{=}(x_{(1)}, y_{(1)}) \geq \alpha, \eta_{=}(x_{(2)}, y_{(2)}) \geq \alpha, \dots$$

$$\dots, \eta_{=}(x_{(n-1)}, y_{(n-1)}) \geq \alpha, \eta_{=}(x_{(n)}, y_{(n)}) < \alpha, \eta_{\geq}(x_{(n)}, y_{(n)}) = \beta.$$

Учитывая, что индуцируемое сильно линейным отношением предпочтения $\mu_{\geq}(x, y)$, отношение $\eta_{\geq}(x, y)$ является сильно линейным, то есть $\max\{\eta_{\geq}(x, y), \eta_{\geq}(y, x)\} = 1, \forall x, y$, число α отражает степень «безразличия», при котором пренебрегается преимуществом i -ой компоненты вектора x над i -ой компонентой вектора y и осуществляется переход к сравнению компонент с номером $(i+1)$.

Так как множество допустимых альтернатив задачи описано нечетко, для того, чтобы сделать выбор, введем к рассмотрению отношение предпочтения $\xi_{\geq}(x, y)$, которое определяется следующим образом:

$$\xi_{\geq}(x, y) = \begin{cases} 1, & \mu_Y(x) \geq \mu_Y(y) \\ 0, & \mu_Y(x) < \mu_Y(y) \end{cases}$$

Считая равноценными отношения $LM_{\alpha}(x, y)$ и $\xi_{\geq}(x, y)$, агрегируем их в одно отношение:

$$\omega_{\geq}(x, y) = \min\{LM_{\alpha}(x, y), \xi_{\geq}(x, y)\}.$$

Функция принадлежности множества недоминированных за отношением $\omega_{\geq}(x, y)$ альтернатив имеет вид:

$$\vartheta^P(x) = 1 - \sup_{y \in Y} \{\omega_{\geq}(y, x) - \omega_{\geq}(x, y)\}$$

и оптимальными являются альтернативы, на которых функция принадлежности достигает максимума.

Отметим, что предлагаемые методы дележа продолжают исследования, опубликованные в [Волошин, 2009, 2010, Voloshin, 2010].

Выводы

В работе дается краткий обзор монографии Л. Келсо и М. Адлера «Манифест Капитализма», где авторы, отмечая недостатки существующей капиталистической системы, предлагают принципы, выполнение которых должно обеспечить «демократизацию» капитализма. В данной работе предлагается модификация экономико-математической модели «смешанного» капитализма [Пономаренко, 2004],

удовлетворяющая принципам «участия», «распределения» и «ограничения». Методы распределения затрат обобщаются на случай принятия решений в нечетких условиях.

Библиография

- [Келсо, 2004] Л. Келсо, М. Адлер, Капиталистический Манифест. – Киев: Орианы, 2004. – 276 с.;
- [Волошин, 2010] Волошин О. Ф., Мащенко С. О. Моделі та методи прийняття рішень: Навчальний посібник для студентів вищих навчальних закладів. –Киев: ВПЦ "Київський університет", 2009.-340с.
- [Орловский, 1981] Орловский С. А. Проблемы принятия решений при нечеткой исходной информации. – Москва: Наука, 1981. – 206 с.
- [Мулен, 1991] Мулен Э. Кооперативное принятие решений: Аксиомы и модели: - Москва: Мир, 1991, 464 с.
- [Пономаренко, 2004] Пономаренко О.І., Перестюк М.О., Бурим В.М. Сучасний економічний аналіз. Навчальний посібник: у 2-х част. – К.: Вища школа, 2004. Ч.1. Мікроекономіка. – 262 с.
- [Волошин, 2009] Волошин А., Горицына И., Мащенко С. Методологические принципы распределения квот на выбросы парниковых газов; In: "Natural and Artificial Intelligence", ITHEA, Sofia, 2010. - P. 85-93.
- [Волошин, 2010] Волошин А. , Лавер В. Нечеткие обобщения моделей распределения затрат; In "Information Models of Knowledge", ITHEA, Kiev-Sofia, 2010.-P.215-219.
- [Voloshin, 2010] Voloshin O.; Maschenko S. Individually Optimal Principles of Distribution of Greenhouse Gas Emission Quotas; In "Information Models of Knowledge", ITHEA, Kiev-Sofia, 2010.-P.209-214.

Сведения об авторах

Алексей Волошин – доктор технических наук, профессор, Киевский национальный университет имени Тараса Шевченко, Украина, 01017 Киев, ул. Владимирская, 64; e-mail: olvoloshyn@ukr.net

Сфера научных интересов: принятие решений, системы поддержки принятия решений, математическая экономика, нечеткий анализ, экспертные системы, е-образование

Анна Сидорук – студент, Киевский национальный университет имени Тараса Шевченко, Украина, 03127 Киев, ул.Сеченова 6, общ.16; e-mail: anna.sydoruk@gmail.com

Сфера научных интересов: принятие решений, математическая экономика, нечеткий анализ

ИССЛЕДОВАНИЕ И МОДЕЛИРОВАНИЕ НЕЙРОСЕТЕВОГО МЕТОДА ОБНАРУЖЕНИЯ И КЛАССИФИКАЦИИ СЕТЕВЫХ АТАК

Адиль Тимофеев, Александр Браницкий

Аннотация: Рассматриваются возможности и перспективы применения нейросетевых технологий для распознавания сетевых атак. Значительное внимание уделяется методологии имитационного моделирования нейронных сетей для обнаружения и классификации сетевых атак и экспериментальным результатам.

Keywords: нейронные сети, методы обнаружения и классификации, сетевые атаки.

ACM Classification Keywords: E.4 CODING AND INFORMATION THEORY

Введение

Одним из наиболее эффективных средств массового распараллеливания и ускорения процессов обработки и передачи потоков данных в задачах обнаружения закономерностей, распознавания образов и классификации данных являются искусственные нейронные сети (НС). Естественным прототипом искусственных НС является биологический мозг и центральная нервная система человека и животных. Возможности искусственных и биологических НС могут значительно расшириться при коллективном (мультиагентном) решении сложных интеллектуальных задач (data mining, knowledge discovery и т.п.).

Высокая сложность и размерность многих задач обнаружения закономерностей, распознавания образов и классификации данных, а также часто возникающая необходимость их решения в реальном времени требуют массового параллелизма и самоорганизации распределённых вычислений на базе НС. С этой точки зрения особый интерес представляют самоорганизующиеся гомогенные и гетерогенные полиномиальные нейронные сети (ПНС) и их разновидности [1–9].

Основные идеи, математические модели, методы обучения и принципы самоорганизации гомогенных и гетерогенных НС были описаны и развиты в [2–9]. К гомогенным НС прежде всего относятся трёхслойные перцептроны и НС Хопфилда [1]. Особенности ПНС заключаются в следующем: архитектура ПНС гетерогенна и многослойна; наличие слоя полиномиальных нейронных элементов (П-нейронов); возможность и целесообразность самоорганизации архитектуры ПНС различных типов; детерминированные и вероятностные методы обучения и самоорганизации гетерогенных ПНС; принципы минимальной сложности и высокой экстраполяции гетерогенных ПНС; алгебраическое требование диофантовости (целочисленности синаптических весов) гетерогенных ПНС [2–9].

Архитектура НС представляет собой иерархическую последовательность нескольких однородных слоёв (непересекающихся подмножеств) параллельно работающих нейроэлементов (НЭ) различных типов. В различных слоях НС могут использоваться разные НЭ, но каждый слой (подмножество НЭ) является однородным (гомогенным). При этом обработка информации в каждом слое НЭ осуществляется параллельно.

Каналы связи между предыдущим и последующим слоями гетерогенной НС являются однонаправленными и имеют регулируемые веса (синаптические параметры). Эти веса каналов связи настраиваются в процессе обучения и самоорганизации архитектуры НС по имеющимся экспериментальным данным или прецедентам.

Традиционно гомогенные или гетерогенные НС используются для автономного принятия решений в задачах обнаружения закономерностей, распознавания образов, диагностики состояний, классификации данных и т.п. По существу эти НС являются обучаемыми интеллектуальными агентами, которые настраиваются на индивидуальное (одно-агентное) решение конкретных задач по обучающим базам данных (ОБД).

В то же время существует большой класс интеллектуальных задач, требующий не только индивидуальных (одно-агентных), но и коллективных (мульти-агентных) решений. Классическим примером этого могут служить особенно сложные и ответственные задачи медицинской диагностики, когда врачи вынуждены прибегать к помощи своих коллег для совместной постановки окончательного диагноза. При этом формируется "консилиум", т.е. профессиональная группа врачей, интегрирующая знания и опыт входящих в неё членов для коллективного принятия наиболее правильных и сбалансированных диагностических решений.

Другим примером сложных задач, требующих коллективных решений, являются глобальные задачи, допускающие естественную (например, иерархическую или мультифрактальную) декомпозицию на множество локальных задач. В этом случае решение сложной (глобальной) задачи может быть распределено между интеллектуальными НС-агентами, специализирующимися на решении N частных (локальных) задач. Параллельная работа N таких НС-агентов может значительно ускорить обработку информации и повысить надежность решения общей (глобальной) задачи.

Специальные агенты-координаторы могут принимать коллективные (мультиагентные) решения на основе локальных (одно-агентных) решений остальных НС-агентов с помощью мажоритарных принципов или процедур голосования (например, по "большинству голосов") [7–9]. При этом все локальные решения принимаются параллельно, что ускоряет принятие глобального (коллективного) решения в N раз.

В ряде случаев глобальная самоорганизация НС-агентов обеспечивается иерархической, фрактальной или мультифрактальной декомпозицией общей задачи на N подзадач. При этом степень внешнего (глобального) параллелизма в мульти-агентной нейросетевой системе определяется параметром N , характеризующем одновременную работу N локальных НС-агентов. Предлагаемые иерархические гетерогенные архитектуры и быстрые алгоритмы обучения ПНС разных типов обеспечивают высокий параллелизм и самоорганизацию нейровычислений в процессе решения интеллектуальных задач. Они успешно применялись для решения ряда прикладных задач (распознавание деталей на конвейере, классификация дорожных ситуаций, диагностика и оценка эффективности лечения артритов, векторная диагностика и расшифровка гастритов, прогнозирование исхода черепно-мозговых травм и т.д.) и нейросетевого представления генетического кода [1–9].

Аккумулируемые в гомогенных и гетерогенных НС с самоорганизующейся архитектурой "нейрообразы" и решающие (классифицирующие и идентифицирующие) правила обеспечивают массовый параллелизм, хорошую экстраполяцию и высокое быстродействие при принятии оптимальных или субоптимальных решений. Коллективное (мультиагентное) использование гетерогенных ПНС в качестве нейросетевых

агентов позволяет дополнительно распараллелить и распределить между локальными НС-агентами процессы решения сложных (глобальных) задач распознавания образов, анализа изображений и сцен, расширенной (векторной) диагностики состояний и адаптивной маршрутизации и классификации информационных потоков.

1. Проблемы и методы защиты информации

Бурное развитие компьютерных сетей и информационных технологий порождает множество проблем, связанных с безопасностью информационных ресурсов. В связи с несовершенством существующих методов защиты компьютерных систем от сетевых атак разработка новых методов защиты информации, позволяющих повысить уровень защищенности компьютерных систем от несанкционированного воздействия, является актуальной и востребованной [10].

Существует три основных подхода, используемых при обнаружении и классификации сетевых атак [10–13]:

- статистический анализ;
- экспертные системы;
- нейронные сети.

Кроме того, развиваются подходы, основанные на и генетических алгоритмах и иммуноклеточных методах.

Статистический анализ находит применение, как правило, при обнаружении аномального поведения. Отклонение от среднего значения (т. е. дисперсия) профиля нормального поведения дает сигнал администратору о том, что зафиксирована атака. Средние частоты и величины переменных вычисляются для каждого типа нормального поведения (например, количество входов в систему, количество отказов в доступе, время суток и т. д.). О возможных атаках сообщается, когда наблюдаемые значения выпадают из нормального диапазона, т. е. превышают заданный порог [10–13].

Экспертная система — это система, которая в контексте обнаружения атак принимает решение о принадлежности того или иного события к классу атак на основании имеющихся правил. Эти правила основаны на опыте специалистов и хранятся в специальном хранилище. В большинстве случаев правила экспертной системы опираются на так называемые сигнатуры, которые и ищутся в контролируемом пространстве [10].

Во введении к данной работе представлен краткий обзор основных типов НС и их особенностей при решении проблем интеллектуального характера. К этим проблемам относятся и задачи обнаружения и классификации сетевых атак. Значительный интерес представляют описываемые ниже исследования и средства их реализации и методом анализа сетевого трафика на наличие аномальных соединений с целью обнаружения и классификации атак на базе трёхслойных НС, обучаемых с помощью алгоритма обратного распространения ошибки.

2. Сетевые атаки и их классификация.

Удалённой сетевой атакой будем называть информационное разрушающее воздействие на распределённую компьютерную сеть, осуществляемое программно по доступным каналам связи. Конкретные разновидности сетевых атак представлены в базе данных (БД) KDD-99 [11]. В качестве обучающего множества выступает база KDD-99 [4]. Эта БД содержит около 5000000 записей о соединениях. Каждая запись представляет собой образ сетевого соединения, включает 41 параметр сетевого трафика и промаркирована как „атака" или „не атака". В базе представлены 22 типа атаки. При этом атаки делятся на 4 основные категории: DoS, U2R, R2L и Probe [10–12].

DoS атаки — это сетевые атаки, направленные на возникновение ситуации, когда на атакуемой системе происходит отказ в обслуживании. Данные атаки характеризуются генерацией большого объема трафика, что приводит к перегрузке и блокированию сервера. Выделяют шесть DoS атак: back, land, neptune, pod, smurf, teardrop.

U2R атаки предполагают получение зарегистрированным пользователем привилегий локального суперпользователя (сетевого администратора). Выделяют четыре типа U2R атак: buffer_overflow, loadmodule, perl, rootkit.

R2L атаки характеризуются получением доступа незарегистрированного пользователя к компьютеру со стороны удаленного компьютера. Выделяют восемь типов R2L атак: ftp_write, guess_passwd, imap, multihop, phf, spy, warezclient, warezmaster.

Probe атаки заключаются в сканировании сетевых портов с целью получения конфиденциальной информации. Выделяют четыре типа Probe атак: ipsweep, nmap, portsweep, satan. Согласно источнику [11] для обнаружения и классификации 9 из 22 типов атак достаточно 29 параметров, характеризующих сетевые соединения. Список этих параметров приведен в табл. 1.

Параметр	Описание параметра
1. duration	Продолжительность соединения
3. service	Служба
4. flag	Флаг терминального состояния IP-соединения
5. src_byte	Количество байт, переданных от источника к приемнику
6. dst_byte	Количество байт, переданных от приемника к источнику
7. land	Равенство порта отправителя порту получателя
8. wrong fragment	Количество отброшенных пакетов
9. urgent	Число пакетов с флагом URG
10. hot	Количество hot-индикаторов
11. count	Количество соединений между удаленным хостом и локальным хостом
12. srv_count	Количество соединений к локальной службе
13. serror rate	Процентное число соединений с ошибкой типа syn для данного

	хоста-источника
14. srv_serror_rate	Процентное число соединений с ошибкой типа SYN для данной службы источника
15. rerror_rate	Процентное число соединений с ошибкой типа REJ для данного хоста-источника
16. srv_rerror_rate	Процентное число соединений с ошибкой типа REJ для данной службы источника
17. same_srv_rate	Процентное число соединений к службе
18. diff_srv_rate	Процентное число соединений к различным службам
19. srv_diff_host_rate	Процентное число соединений к различным хостам
20. dst_host_count	Количество соединений к локальному хосту, установленных удаленной стороной
21. dst_host_srv_count	Количество соединений к локальному хосту, установленных удаленной стороной и использующих одну и ту же службу
22. dst_host_same_srv_rate	Процентное число соединений к локальному хосту, установленных удаленной стороной и использующих одну и ту же службу
23. dst_host_diff_srv_rate	Процентное число соединений к локальному хосту, установленных удаленной стороной и использующих различные службы
24. dst_host_same_src_port_rate	Процентное число соединений к данному хосту при текущем номере порта источника
25. dst_host_srv_diff_host_rate	Процентное число соединений к службе разных хостов
26. dst_host_serror_rate	Процентное число соединений с ошибкой типа syn для данного хоста-приемника
27. dst_host_srv_serror_rate	Процентное число соединений с ошибкой типа SYN для данной службы приемника
28. dst_host_rerror_rate	Процентное число соединений с ошибкой типа REJ для данного хоста-приемника
29. dst_host_srv_rerror_rate	Процентное число соединений с ошибкой типа REJ для данной службы приемника

Таблица 1. Параметры сетевых соединений.

3. Исследование нейросетевой технологии обнаружения и классификации сетевых атак

Обычно для обучения и тестирования НС имеющиеся экспериментальные данные разбиваются на обучающую БД и контрольную БД. В проведённых исследованиях в качестве обучающей БД с параметрами сетевых соединений, представляющая собой 10% subset KDD CUP. 99. Она представляет собой обычный текстовый файл, разбитый на несколько файлов, каждый из которых содержит набор

параметров определённого типа атаки или нормального соединения. Полученные файлы преобразуются в файлы, содержащие числовые эквиваленты, соответствующие определённому параметру соединения.

В целях ускорения обучения, тестирования и дальнейшего запуска применялся алгоритм, реализующий метод главных компонент. Вычисление вариационной матрицы главных компонент сводится к вычислению собственных векторов и собственных значений ковариационной матрицы исходных данных. Найденные собственные векторы представляют собой коэффициенты, задающие линейное преобразование для исходных векторов, а соответствующие им собственные числа выступают в роли критерия (меры) информативности новой системы. Из набора сжатых векторов с помощью генератора случайных чисел формируется обучающая БД, мощность которой составляет 10% от полной KDD CUP. 99.

Элементы этой БД поочередно подаются на вход НС для настройки весовых (синаптических) параметров. Весовые коэффициенты каждой НС сохраняются в соответствующих файлах для возможной их дальнейшей загрузки при очередном запуске системы обнаружения вторжений.

По завершении этих этапов НС полностью готовы к этапу запуска, который заключается в обнаружении и классификации активных аномальных и нормальных IP-соединений.

При эмулировании сетевых соединений данные считываются из полной базы KDD Cup 99, которая рассматривалась как контрольная БД. Записи из полной контрольной БД подавались на обученные НС. На основании результатов этого этапа определяется статистика и даётся анализ эффективности НС по критериям качества распознавания типов аномальных соединений и наличию ложных срабатываний (когда нормальное соединение принимается за атаку).

Процесс обучения и тестирования НС показан на рис. 1, а процесс запуска НС представлен на рис. 2.

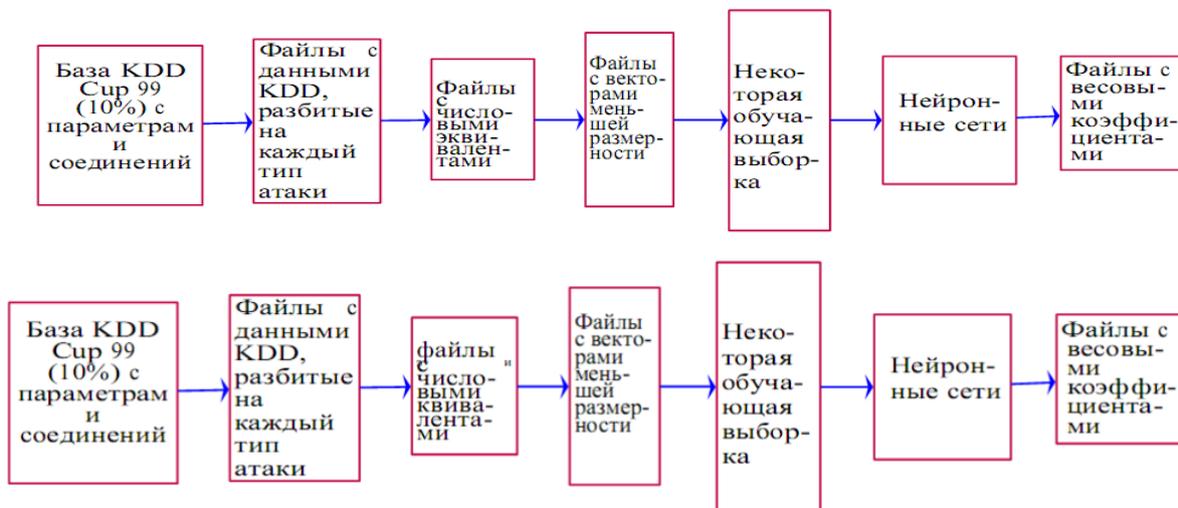


Рис. 1: Процесс обучения и тестирования НС



Рис. 2: Процесс запуска НС

Пакеты, поступающие на сетевую карту, перехватываются и обрабатываются анализатором пакетов (сниффером). Из поля данных и заголовка каждого пакета выделяются и вычисляются необходимые для классификации соединений параметры. Собираются некоторые статистические сведения об активных в данный момент соединениях. После сбора всех необходимых показаний о каждом запущенном соединении набор вычисленных параметров "подвергается" текстовой обработке: каждая отдельная составляющая набора параметров отображается в соответствующий ей числовой эквивалент согласно таблице, заданной в программе. Опционально (т. е. необязательно) полученный набор параметров сжимается по методу главных компонент.

Преобразованные векторы подаются на вход обученных НС, которые в свою очередь формируют линейную комбинацию каждого вектора со своими весовыми параметрами, выступающей в роли аргумента функции активации НС. На основании полученного значения принимается решение о принадлежности конкретного соединения тому или иному классу атак.

Следует отметить, что существенное значение в успешности обнаружения и определения злоупотреблений является правильно заданная планка, выступающая в роли "водораздела" между типичными атаками и нормальными соединениями. Слишком малое значение этого барьера означает возможность пропуска атак и принятия их за нормальные соединения, а слишком большое значение увеличивает число ложных срабатываний. Тем самым большое количество типично нормальных соединений будет приниматься за аномальные.

Каждая НС в отдельно созданном для нее потоке обучается параллельно с остальными для распознавания нормального соединения и одного типа атаки. Для их обучения используются равные по размеру выборки положительного и отрицательного трафиков. По завершении стадии обучения происходит барьерная синхронизация всех потоков (основной поток дожидается завершения выполнения созданных потоков).

Программная реализация НС имеет вид:

```
struct neural_network {
struct fann *ann;
enum type_attack attack;
pthread_t thread;
int index;
};
enum type_attack {back, neptune, pod, smurf, teardrop, ipsweep, nmap, portsweep, satan, normal};
```

Она состоит из следующих полей:

- указателя ann на структуру struct fann, которая определена в библиотеке FANN (Fast Artificial Neural Network) [13];
- переменной attack типа перечисления enum type_attack, задающей тип атаки, которую обучена распознавать данная нейросеть;
- переменной thread типа pthread_t, которая служит в роли идентификатора потока, в котором выполняется данная нейронная сеть;

• переменной `index` типа `int`, которая задает положение (индекс) данной переменной типа `struct neural_network` в статическом массиве.

Для обучения НС применяется алгоритм обратного распространения ошибки (итеративный градиентный алгоритм, который используется с целью минимизации ошибки работы многослойного перцептрона и получения желаемого выхода) с симметричной сигмоидальной функцией активации.

4. Результаты вычислительных экспериментов

НС, которые обнаружили подозрительную сетевую активность, сигнализируют об этом и записывают параметры аномального соединения в журнал регистрации.

При проведении этапа эмулирования сетевых атак были получены следующие результаты, представленные в табл. 2. Здесь левый столбец – тип эмулируемого соединения, верхняя строка – тип нейронной сети, их пересечение – процентное количество правильно распознанных атак соответствующего типа соответствующими НС.

	back	neptune	pod	smurf	teardrop	ipsweep	nmap	portsweep	satan
back	99.7%	0%	0%	0%	0%	0%	0%	0%	0.2%
neptune	7.9%	100%	0%	0%	0%	0%	80.8%	100%	19.1%
pod	0.4%	0.4%	99.6%	0%	32.1%	1.5%	1.9%	0.4%	1.5%
smurf	0%	0%	0%	99.8%	0%	0%	99.9%	0%	100%
teardrop	0%	0%	77.3%	0.1%	99.9%	0%	0.1%	0%	0.1%
ipsweep	0%	0%	0%	0%	0%	99.8%	92.2%	1.5%	1.7%
nmap	0%	44.6%	0%	0%	0%	44.1%	99%	44.6%	3%
portsweep	0.4%	9.9%	0%	0%	0%	0%	0.1%	99.6%	90.2%
satan	0%	88.6%	0%	0%	0.1%	0.1%	3.6%	88.7%	99.5%
normal	0.4%	0%	0%	0%	0%	0.5%	1.1%	0.1%	1.3%

Таблица 2: Показатели эффективности обнаружения атак НС

Заключение

Исследование и имитационное моделирование гомогенных НС на экспериментальных данных KDD Cup 99 в задачах обнаружения и классификации сетевых атак свидетельствует об эффективности нейросетевых технологий. Естественно ожидать, что при использовании гетерогенных НС полиномиального и диофантового типа, а также их коллективов результаты распознавания сетевых атак могут быть улучшены. Перспективным представляется также частично исследованный авторами метод обнаружения и классификации сетевых атак, основанный на иммуно-клеточном подходе.

Работа выполнена при частичной поддержке гранта по Программе № 14 Президиума РАН (GRID) и издательского гранта РФФИ № 12-08-07022-д.

Литература

- [1] С.Хайкин. Нейронные сети. Полный курс.– М.: Издательский дом “Вильямс”, 2006, 1104 с.
 [2] Каляев А. В., Тимофеев А. В. Методы обучения и минимизации сложности когнитивных нейро-модулей нейрокомпьютера с программируемой архитектурой. - Доклады АН, 1994, т. 237, с. 180-183.

- [3] Тимофеев А.В. Методы синтеза диофантовых нейросетей минимальной сложности. - Доклады АН , 1995, т.301, № 3, с.1 106-1109.
- [4] Тимофеев А.В., Шибзухов З.М. Методы синтеза и минимизации сложности диофантовых нейронных сетей над конечным полем. Автоматика и телемеханика, 1997, № 4, с. 204-212.
- [5] Тимофеев А. В. Оптимальный синтез и минимизация сложности генно-нейронных сетей по енетическим базам данных. Нейрокомпьютеры: разработка и применение, № 5-6, 2002, с. 34-39.
- [6] Тимофеев А. В., Шибзухов З. М., Шеожев А. М. Синтез нейросетевых архитектур по многозначному дереву решений. – Нейрокомпьютеры:разработка и применение, № 5-6, 2002, с. 44-49.
- [7] Timofeev A. V. Polynomial Neural Network with Self-Organizing Architecture. - International Journal on Optical Memory and Neural Networks, 2004, N 2.
- [8] Timofeev A. V. Parallelism and Self-Organization in Polynomial Neural Networks for Image Recognition. - Pattern Recognition and Inage Analysis, 2005, vol. 15, No.1, pp. 97 - 100.
- [9] Тимофеев А.В. Иерархические самоорганизующиеся нейронные сети и алгоритмы мульти-агентного принятия решений – Материалы 4-й Всероссийской мульти-конференции по проблемам управления МКУ-2011 (Россия, Дивноморское, 28 сентября – 7 октября 2011 г.), т.1, с. 43–45.
- [10] А.В. Лукацкий. Обнаружение атак. 2-е изд., перераб. и доп. – СПб.: БХВ-Петербург, 2003.
- [11] KDD Cup 1999 Data <http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html>
- [12] Технологии обнаружения сетевых атак. <http://www.bstu.by/~opo/ru/uni/bstu/science/ids/>.
- [13] Fast Artificial Neural Network.<http://leenissen.dk/fann/up>

Сведения об авторах



Тимофеев Адиль Васильевич – заведующий лабораторией информационных технологий в управлении и робототехнике Санкт-Петербургского института информатики и автоматизации Российской академии наук, Профессор кафедры информатики математико-механического факультета Санкт-Петербургского государственного университета, доктор технических наук, профессор, Заслуженный деятель науки РФ, 199178, Россия, Санкт-Петербург, 14-я линия, д. 39, СПИИРАН, tav@ias.spb.su



Браницкий Александр Александрович – студент кафедры информатики Математико-механического факультета Санкт-Петербургского государственного университета, дипломник, alexander.branitskiy@gmail.com

МОДЕЛИРОВАНИЕ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ В ПРОЦЕССАХ ЭВОЛЮЦИОННОГО ОБУЧЕНИЯ.

Ирина Трегубенко

Аннотация: Предложена технология эволюционного обучения интеллектуальных систем. Сформулировано новое понимание дискретности процесса обучения с учетом принципа ограниченности жизненного цикла интеллектуальной системы. Предложена двухмодульная концепция моделирования интеллектуальных систем, введено новое понятие B-System.

Ключевые слова: интеллектуальные системы, эволюционное обучение.

ACM Classification Keywords: H. Information Systems, H.1 MODELS AND PRINCIPLES, H.1.1 Systems and Information Theory.

Анализ и постановка проблемы

Современные эволюционные методы в области искусственного интеллекта традиционно отталкиваются от физических законов и методов организации известной вселенной. Ученые концентрируют усилия на изучении физических составляющих материи, их взаимосвязях и превращениях. При моделировании и воспроизведении процессов эволюции, в самых разнообразных аспектах, в частности на базе когнитивного подхода [Трегубенко, 2010], исследователи базируются на изучении и моделировании реальных физических, физиологических процессах, проходящих в естественных системах. Процессы мышления, применительно к интеллектуальным системам рассматриваются как возможность использования знаний и выводов, базирующихся на этих знаниях, для формирования эффективного, рационального поведения системы [Трегубенко, 2010]. При этом, процессы анализа, систематизации и хранения информации не отделяются от физической платформы системы. Существует много теорий и методов воспроизведения, протекающих в естественных системах процессов мышления, например моделирование нейронов и нейронных систем, однако задача построения полноценного искусственного разума еще не решена.

При построении искусственных систем с использованием средств интеллектуализации наиболее существенными и проблематичными являются процессы обучения интеллектуальных составляющих, в частности модулей управления [Трегубенко, 2009]. Для проектирования таких подсистем могут быть использованы технологии интеллектуальных агентов, а для их адаптации к окружающей среде, наиболее эффективным и надежным представляется метод обучения с подкреплением [Tregubenko, 2010], построенный на базе теории обучения с подкреплением [Sutton, 1998], которая в свою очередь может рассматриваться как развитие теории адаптивного поведения. Однако возникают закономерные трудности, в связи со сложностью и многокритериальностью задачи обучения.

Цель

Необходимо предложить технологию, которая понизит сложность системы обучения, т.е. взаимосвязей и действующих факторов взаимодействия пары интеллектуальная система – среда обучения, и построить соответствующую модель интеллектуальной системы.

Основной материал.

Интеллектуальные системы ранее рассматривались автором, на примере интеллектуальных агентов, как совокупность нейросетевых модулей [Tregubenko, 2010], которые находятся в постоянном взаимодействии с постоянно изменяющейся внешней средой. Однако такой подход является чисто технологическим, привязанным к конкретной физической реализации и не дает значительного эффекта с точки зрения уменьшения сложности задачи. Безусловно, этот подход будет применяться в дальнейшем, однако на более низких уровнях абстракции процесса обучения.

Рассмотрим более подробно процесс обучения интеллектуальной системы. Наиболее значимая проблема организации эффективного и результативного обучения состоит в сложности корректной постановки задачи обучения.

Обучение более сложных систем занимает больше времени, чем обучение более простых систем. Кроме этого на время обучения так же влияет прикладная область, т.е. для выполнения каких именно функций обучаем интеллектуальную систему. Нельзя рассматривать обучение вне контекста к функциям, которые должна выполнять обучаемая система. Задачу обучения можем рассматривать исключительно в контексте ограничений по определенному набору функций, который должен уметь выполнять обучаемый.

Если принять во внимание, что ставится не абстрактная задача обучения без ограничений, то возникает проблема формулировки цели обучения и построение системы критериев, подтверждающих адекватность выполненного процесса обучения. Кроме того, необходимо учитывать, что так называемые самообучающиеся системы бессмысленны и могут стать опасными в случае бесконтрольности процесса обучения со стороны внешней среды и/или субъекта обучения.

В процессе обучения интеллектуальных систем можно выделить три составляющих (рис.1): объект, субъект и среда обучения. Под объектом обучения понимаем интеллектуальную систему (ИС), которая должна выполнить процесс обучения, т.е. освоить определенный объем знаний и приобрести соответствующий набор навыков. Под субъектом понимаем автора (А) системы. Без понимания участия и функций составляющей – автор обучающейся интеллектуальной системы, невозможно поставить цели обучения и собственно инициировать сам процесс. Безусловно, автор так же может иметь сложную структуру, быть отдельным независимым индивидом или объединением, системой, причем не существенно какой именно – искусственной или естественной.

Третья составляющая рассматриваемого процесса – среда обучения(С). В зависимости от соотношения автора и среды обучения, можно выделить три принципиально различных типа организации процессов обучения интеллектуальной системы (рис.1.). Вариант а) на рисунке 1 демонстрирует ситуацию, когда среда обучения отделена от автора. Вариант б) на рисунке 1 вторая демонстрирует ситуацию, когда среда обучения содержит в себе автора, либо полностью либо частично. Вариант в) на рисунке 1

представляет ситуацию, когда автор идентичен среде обучения, т.е. по сути, среда обучения есть автором. Последний вариант соответствует процессам самообучения интеллектуальной системы.

В области информационных технологий и их программных реализаций видно, что технологические детали организации платформы могут быть различны для одной и той же информационной системы. Очевидным есть факт того что, один и тот же алгоритм может существовать в различных записях, синтаксически и технологически различных. При этом, однако, алгоритм, по сути, не меняется. По аналогии, отделим понятие интеллекта от конкретной физической реализации системы, которая проходит процесс обучения.

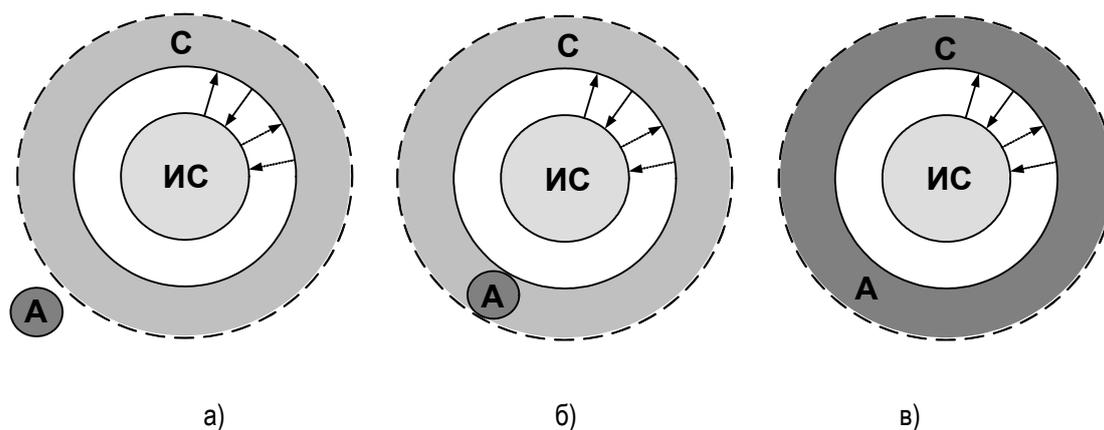


Рис. 1. Типы организации процессов обучения интеллектуальной системы: а) – внешний независимый автор; б) – автор, как часть среды обучения; в) – автор идентичен среде обучения.

Рассмотрим более подробно идею эволюционного обучения интеллектуальных систем. Современные методы обучения интеллектуальных систем разнообразны, однако обычно привязаны к физическим характеристикам, структуре и свойствам объекта обучения и окружающей его среды. Абстрагируемся от технической платформы и физической реализации конкретной интеллектуальной системы. Для упрощения задачи, на первом этапе, отбрасываем внутреннюю структуру, конкретные конструкционные решения и ставим задачу представить и описать технологию обучения интеллектуальной системы на уровне абстракции.

Если посмотреть на организацию естественного мира нам известного, можно утверждать что универсальной интеллектуальной системы не существует. Каждая из интеллектуальных систем находится на определенном уровне сознания и способна выполнять ограниченный набор функций. Можно утверждать так же, что жизненный цикл любой системы ограничен и конечен. Безусловно, за ограниченный жизненный цикл интеллектуальной системы она может обучиться ограниченному набору функций и приобрести конечный набор знаний. Следовательно, необходимо рассмотреть вопрос увеличения жизненного цикла интеллектуальных систем, сделав его условно бесконечным. Невозможно построить интеллектуальную систему с одним непрерывным бесконечным жизненным циклом, т.е. условное бесконечное существование невозможно обеспечить на базе ограниченной физической реальности.

Если принять за исходное утверждение, что за один жизненный цикл существования интеллектуальной системы, она может обучиться определенному ограниченному набору функций и навыков. Тогда, для

продолжения обучения, система должна получить новые исходные данные и задачу – новый комплект осваиваемых функций. Таким образом, задача обучения интеллектуальной системы, сводится к созданию интерактивного процесса обучения на разных иерархических уровнях функционирования, с учетом необходимости сохранения части знаний и умения выполнять определенные функции при переходе на новый иерархический уровень, с потерей не нужных на следующем уровне свойств и способностей (технологических интерфейсных функций), которые ранее использовались для приобретения знаний и умения выполнять определенные функции, с приобретением новых свойств, способностей (технологических функций) и исходных данных, позволяющих оперировать с новыми навыками в новой среде.

Исходя из выше сказанного, проведем упрощение сложности процесса обучения, представив его в виде эволюционного процесса, который состоит из последовательных дискретных вложенных процессов с динамически изменяемыми входящими данными и исходными состояниями.

Процесс обучения интеллектуальной системы разделяем на конечные иерархические итерационные этапы (рис.2). В рамках каждого этапа будет проходить процесс обучения, привязанный к конкретным физическим параметрам среды и ограниченным задачам обучения, что и будет обуславливать конечность эволюционного этапа обучения на данном итерационно-иерархическом уровне.

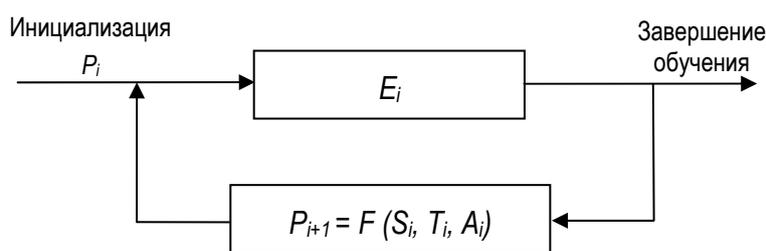


Рис.2. Итерационный этап эволюционного обучения.

За эволюционный этап обучения принимаем цепочку: инициализация – обучение – завершение обучения. Во время первой процедуры инициализации, проводится моделирование первого эволюционного этапа, определяются задачи обучения, формируются текущие технологии адаптации к физической среде. Далее запускается процесс обучения, т.е. в общем случае выполняется i -ый эволюционный этап обучения (E_i). На этой стадии вполне применимы технологии обучения с подтверждением [Tregubenko, 2010].

После окончания текущего i -ого эволюционного этапа обучения (E_i) предусматривается три основные процедуры:

1. Сохранение результатов (S_i) обучения интеллектуальной системы, т.е. оптимизацию и сохранность полученных в результате обучения знаний в независимой от среды и физ. платформы форме.
2. Отбрасывания ненужных технических данных и служебных функций (T_i), с помощью которых происходила адаптация системы под конкретную физическую реализацию и конкретные задачи конкретного этапа обучения.
3. Анализ (A_i), т.е. оценка соответствия результатов обучения поставленным целям, принятие решения о переходе на следующий уровень обучения, с возможностью интерактивного моделирования этого уровня, постановкой новых задач на обучение и формированием текущих технологий адаптации к новой физической среде. Следует обратить внимание, что следующим уровнем может быть выбран, как и

новый $i+1$ -ый эволюционный этап, так и текущий i -ый этап, что будет соответствовать повторному обучению. В общем случае следующим уровнем, на котором будет проводиться обучение интеллектуальной системы, может быть выбран любой иерархический уровень.

Построим модель интеллектуальной системы с учетом рассмотренных выше технологий эволюционного обучения. Разделим интеллектуальную систему на две составляющие. Одна составляющая должна нести в себе интеллектуальные функции и сохраняться сколь угодно долго. Другая составляющая должна быть достаточно гибкой и обеспечивать приспособляемость системы к изменяющейся окружающей среде. Эта составляющая не обязательно должна быть постоянной и вполне может быть сменной, заменяемой, с коротким периодом существования. Модель эволюционно обучающейся системы представлена в виде двух принципиально различных модулей: интеллектуальный модуль и модуль адаптации к окружающей среде (рис. 3).

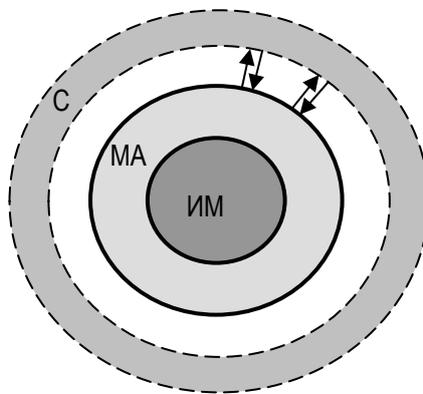


Рис.3. Модульная структура интеллектуальной системы.

Интеллектуальный модуль (ИМ) рассматривается как постоянная составляющая и является сутью системы и собственно, самой системой в нашем понимании, время существования которой в принципе неограниченно. Модуль адаптации (МА), является изменяемой частью и обеспечивает функционирование интеллектуального модуля в текущих физических реалиях. Время существования и модификации модуля адаптации определяется физическими свойствами окружающей среды и/или выполнением задач обучения на конкретном эволюционном этапе. Двухмодульное представление о интеллектуальных системах, которые в дальнейшем будем называть B-System, соответствует эволюционному пониманию процесса обучения. Ниже представлен механизм приспособляемости таких систем к новым физическим параметрам среды при переходе на новый итерационный эволюционный этап обучения (рис.4).

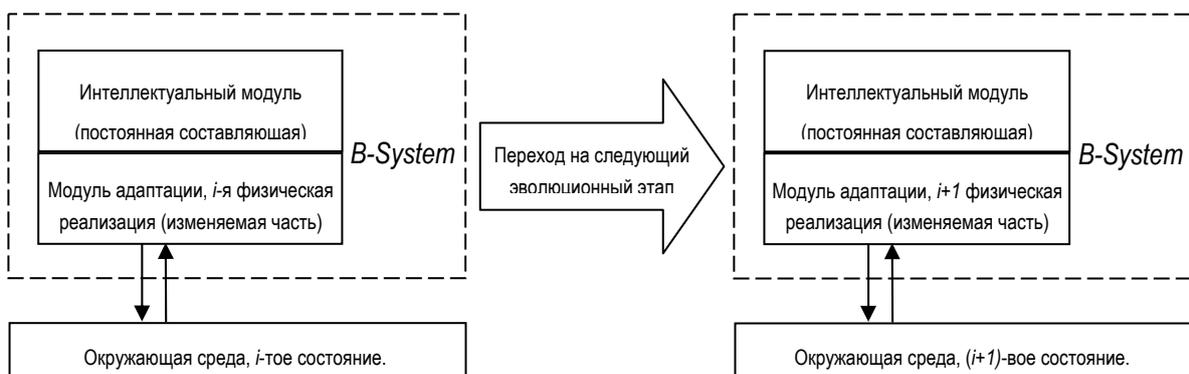


Рис. 4. Механизм приспособляемости B-System к новым физическим параметрам среды.

Выполним формализацию задачи обучения интеллектуальной системы на i -ом эволюционном этапе итерационно-иерархического процесса обучения.

Во время процедуры инициализации на i -ом эволюционном этапе выполняется дискретизация задач обучения, определяется конечный набор функций, которые должна освоить интеллектуальная система. Таким образом, на i -ом эволюционном этапе интеллектуальная система должна выполнить M процессов обучения. Интеллектуальной системе выделяется соответствующее множество ресурсов R^i для выполнения каждого процесса на i -ом эволюционном этапе обучения:

$$R^i = \sum_{k=1}^M r_k^i \quad (1)$$

Для выполнения одного k -ого процесса обучения необходимо освоить ограниченное количество N_k^i ресурсов. Условие выполненного процесса обучения, возможно определить через условие полного освоения выделенного ресурса R^i на i -ом эволюционном этапе обучения:

$$\sum_{k=1}^M \sum_{j=1}^{N_k^i} r_{kj}^i N_k^i = R^i \quad (2)$$

где, r_{jk}^i - количество j -ого ресурса переработанного k -им процессом обучения,

Целевая функция обучения на i -ом эволюционном этапе обучения может быть представлена в виде:

$$F^i = \sum_{k=1}^M \sum_{j=1}^{N_k^i} z_{kj}^i N_k^i + \sum_{k=1}^M \varphi_k^i(N_k^i, N_k^{ii}) \quad (3)$$

где z_{kj}^i - затраты j -ого ресурса на k -ом процессе i -ого эволюционного этапа обучения интеллектуальной системе.

Понятно, что N_k^i всегда меньше чем N_k^{ii} -максимально возможное количество реализованных процессов обучения интеллектуальной системы, т.е. освоенных ресурсов на на i -ом эволюционном этапе.

φ_k^i - штрафная функция, за отклонения количества выполненных процессов обучения от максимально возможного количества выполненных процессов обучения на i -ом эволюционном этапе.

Таким образом, задача обучения интеллектуальной системы на i -ом эволюционном этапе может быть сформулирована следующим образом: минимизировать целевую функцию (3) при выполнении ограничения (2).

Выводы.

Организация процессов обучения интеллектуальных систем на базе эволюционных методов может повысить эффективность обучения. Разработана идея и технология эволюционного обучения интеллектуальных систем. Сформулировано новое понимание дискретности процесса обучения с учетом принципа ограниченности жизненного цикла интеллектуальной системы. Предложена двухмодульная концепция моделирования интеллектуальных систем, введено новое понятие B-System, состоящих из двух взаимосвязанных модулей, имеющих различные свойства и жизненный цикл. На этой базе и

предполагается развивать методы построения и организации процессов эволюционного обучения интеллектуальных систем.

Библиография

[Трегубенко, 2009] Трегубенко І.Б. Концепція інтелектуального управління в складних розподілених системах. // Матеріали ІХ міжнародної наукової конференції ім.Т.А.Таран „Інтелектуальний аналіз інформації ІАІ-2009” – К. : ПРОСВІТА, 2009. - 464 с.:ил - С. 391-393

[Трегубенко, 2010] Трегубенко І.Б. Адаптивна поведінка інтелектуальних агентів захисту на базі теорії когнітивної еволюції. // 17 міжнародна конференція з автоматичного управління "Автоматика - 2010". Тези доповідей. Том2. - Харків: ХНУРЕ, 2010. - 306 с.- С.202-203

[Tregubenko, 2010] Tregubenko I.B. Reinforcement learning intellectual agent of protection for adapting to surrounding environment // SIN'10: Proceedings of the 3rd international conference on Security of information and networks/ - ACM New York, NY, USA ©2010 - pp: 110-112. ISBN: 978-1-4503-0234-0 DOI=10.1145/1854099.1854122.

[Sutton, 1998] Richard S. Sutton , Andrew G. Barto, Introduction to Reinforcement Learning, 1st edition, MIT Press, Cambridge, MA, USA ©1998 P.342. ISBN: 0262193981

Информация про автора



Ирина Трегубенко – кафедра программного обеспечения автоматизированных систем, Черкасский государственный технологический университет, Черкассы, кандидат технических наук, доцент. Научные интересы: теория эволюционных систем, искусственный интеллект, эволюционные алгоритмы, интеллектуальные технологии в информационной безопасности и управлении; e-mail: irttri@rambler.ru

ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ ПРИМЕНЕНИЯ СЕМАНТИЧЕСКИ ОРИЕНТИРОВАННЫХ МЕТОДОВ КЛАССИФИКАЦИИ ЗАДАЧИ OPINION MINING

Нина Хайрова, Наталья Шаронова

Аннотация: Как правило, для того чтобы принять верное бизнес решение, необходимо прогнозировать изменения на рынке и в обществе в целом. Интеллектуальный анализ знаний, содержащихся в текстах Web-ресурсов пользовательского контента, сегодня в основном, осуществляется с помощью «ручной» работы аналитиков, скорость работы которых не соответствует темпам социально-экономических процессов общества. Хорошим дополнительным средством решения проблемы интеллектуального анализа слабоструктурированных текстов Web-ресурсов может стать отслеживание общественного мнения по отношению к тому или иному объекту или субъекту анализа, с использованием методов sentiment классификации. В работе предлагается система, реализующая методы opinion mining, позволяющая извлекать полезные знания из пользовательского Web-контента, использующая информационные технологии семантической классификации и идентификации многоязыковых семантических эквивалентов, и базирующаяся на специально-структурированном тезаурусе. Система использует механизм классификации мнений, позволяющий добавлять оценочные правила, представляющие позитивные, негативные и нейтральные оценки характеристик слов. В качестве основания классификации слабоструктурированной информации используются отношения содержащихся в ней аргументов к исследуемому объекту. Соответствие полного текста классу позитивного или негативного мнения определяется статистико-семантическими методами, основанными на атрибутивной положительной или отрицательной оценке объектов (субъектов) предложений. Кроме того, в работе показана специальная структура тезауруса, с явно заданными отношениями между концептами, используемыми для решения задач Opinion Mining, и идентификации семантических эквивалентов. Экспериментальные результаты на полнотекстовой базе пользовательского Web-контента текстов на английском и немецком языках показали достаточно высокую эффективность использования системы.

Ключевые слова: интеллектуальный анализ знаний, opinion mining, семантическая классификация, онтологии, тезаурусы.

ACM Classification Keywords : H.3.3 .Information Search and Retrieval

Введение

Для того чтобы принять верное бизнес решение, необходимо прогнозировать изменения на рынке и в обществе в целом. Правильный прогноз во многом базируется на отслеживании общественного мнения по отношению к тому или иному объекту инвестирования. Сегодня для правильного принятия финансово значимых экономических решений необходимо анализировать большое количество информации,

находящейся как в специализированных экономических источниках (например, информационные источники рейтинговых агентств Fitch, Moody's, Standard@Poor's, A.M.Best), так и в источниках общего характера.

Традиционно инвесторы, принимая решения на фондовом рынке о входе на рынок (покупка акций) или выходе из него (продажа акций), обычно используют технический анализ, основанный на анализе временных рядов цен акций, с привлечением различных индикаторов, а также методы фундаментального анализа, базирующиеся на отчетной информации о текущем финансовом состоянии компаний эмитентов акций.

Однако, в связи с тем, что объемы неструктурированной информации увеличиваются с каждым годом, и сегодня, по данным агентства Gartner, неструктурированные документы составляют более 80% корпоративных данных, а количество внешних источников (интернет-ресурсов, блогов, форумов, СМИ) исчисляется миллионами, мощным дополнительным средством прогнозирования ожидаемых изменений на фондовых рынках может стать информация, извлекаемая из текстов соответствующей проблематики. Интеллектуальный анализ знаний, содержащихся в текстах Web-ресурсов, электронных журналах, электронной корреспонденции, социальных сетях и другом пользовательском контенте (user generated content), на сегодня, по большей части, осуществляется с помощью «ручной» работы аналитиков, скорость работы которых не соответствует интенсивности работы фондового рынка. Поэтому все чаще основным вопросом эффективности принятия бизнес решений становится скорость извлечения информации и постоянный ее мониторинг.

Актуальность исследования

Различные системы, использующие Text Mining, могут предоставить бизнес аналитику возможность работать с большими объемами исходных данных за счет автоматизации процесса извлечения нужной информации [Gupta, 2009]. Но, несмотря на наличие большого количества подобных систем, к настоящему моменту нет систем, осуществляющих полную автоматическую обработку слабоструктурированных текстов, представленных в корпоративных и глобальных информационных системах. Это связано, прежде всего, с тем, что для принятия бизнес решения необходимо быстрое получение полной и релевантной информации, а коэффициенты полноты и точности существующих гипертекстовых систем составляют в среднем соответственно 0,9, и 0,8, а документальных систем - даже 0,7 [Gupta, 2009], что представляется довольно низким для принятия правильных бизнес решений.

Используемые технологии поиска и классификации слабоструктурированной информации базируются, в своем большинстве, на традиционных статистико-вероятностных подходах, не используя или слабо используя семантический, т.е. смысловой анализ [Нгок, 2012]. Это поиск по ключевым словам, основанный на техниках двоичного поиска, ранжированного поиска, вероятностных моделях поиска и т. д. Такой поиск не учитывает ни семантические связи между понятиями, ни наличие семантических эквивалентов слов и словосочетаний [Напу, 2011].

Информационное обеспечение процессов принятия бизнес решений может стать более качественным благодаря использованию, при решении общей задачи получения и обработки информации, моделей и методов Opinion Mining [Ding, 2008], в которых, в качестве основания классификации слабоструктурированной информации пользовательского контента, используется отношение (например,

положительное или отрицательное, поддерживающее или отклоняющее) содержащихся в них аргументов к исследуемому объекту [Esuli, 2006].

Дальнейшее развитие методов Text Mining в направлении информационного обеспечения бизнес решений сдерживается не только отсутствием адекватных моделей и методов семантического анализа, семантического поиска и смысловой классификации, но и отсутствием стройной системы лингвистического обеспечения [Оробинская, 2010]. Недостаточность онтологий и тезаурусов широких предметных областей приводит к невозможности учитывать особенности смысловой организации текстов по конкретной тематике, в частности, особенности смысловой организации экономических текстов.

Общая постановка задачи

Для получения полной информационной картины по определенному объекту (например, акции) или субъекту (например, банк) необходимо использовать методы Opinion Mining, представляющие инструменты для автоматического извлечения из текстов «субъективной» информации, используемой для автоматической оценки (позитивной, негативной, нейтральной) новостных событий, продуктов, персоналий, организаций, стран мира и т.д., поступающих в режиме реального времени из сообщений электронных средств массовой информации, сообщений блоггеров, дискуссионных форумов и т.д.

Инструменты Opinion mining включены во многие системы, например, Saliency Engine (Lexalytics, Inc., Boston, USA); SentiMetrix (University of Maryland, Institute for Advanced Computer Sciences), Twitter Sentiment (Stanford University), J.D.Power Text Analyst (J.D. Power and Associates, USA), RavenPack News Analyst (RavenPack International S.L.) и др. Для решения данной задачи используются алгоритмы опорных векторов и нейронных сетей, семантические сети, ассоциативные и оценочные правила, наивный байесовский метод, CRF модели и др. [Kobayashi, 2007]. Однако реального практически используемого результата, позволяющего включить данную технологию в систему информационного обеспечения процессов бизнес решений, на сегодняшний день не существует.

Задача осложняется наличием такой неотъемлемой составляющей естественных языков, как синонимия. И хотя сегодня существуют методы автоматического выделения смысловых эквивалентов из текста, а также синонимические связи учитываются онтологиями и тезаурусами, в практической работе для учета семантической эквивалентности пользователю часто приходится повторять запросы с каждым из «вручную» определенных смысловых эквивалентов термина.

Таким образом, на основе проведенного анализа существующих моделей и методов решения задачи Opinion mining в работе предлагается рассмотреть задачу Opinion Mining и как часть общей задачи получения и обработки информации, и как задачу автоматизированного построения классификатора слабоструктурированных текстов, представляющих так называемый «user generated content». Кроме того, для получения релевантного результата при осуществлении Social Media Monitoring необходимо учесть поступление информации на различных языках и использование в текстовых сообщениях семантических (смысловых) эквивалентов.

В настоящей работе категория мнения рассматривается как субъективные заявления людей, отражающих настроения или представления о сущностях или событиях, а информационная технология решения задачи Opinion Mining рассматривается, как поэтапное решение следующих задач:

-
- идентификации объекта или субъекта мнения;
 - идентификации сущностей, выражающих мнения;
 - идентификации полярности мнения;
 - выявления семантических эквивалентов.
-

Структурное описание модели

Проведенный обзор показывает, что 28% информации, значимой для принятия бизнес решений, находится в слабоструктурированных текстах. Информационное обеспечение, необходимое для обработки данной информации, должно включать в себя как информационный поиск, содержащий мультилингвистическую семантическую задачу, напрямую связанную с релевантностью результатов, — выявления семантических эквивалентов, так и методы Text Mining. Метод Opinion Mining, позволяющий осуществлять отнесение текстов, по мнению авторов, к определенной проблеме, появился сравнительно недавно, и во многом использует подходы Text Mining и Information Retrieval (рис. 1).

Тексты экономически значимых документов, в подавляющем большинстве, оформлены в официально-деловом стиле, лексика которого характерна логической связностью и насыщенностью экономическими и специальными общенаучными терминами. Такого рода термины представляют собой эмоционально-нейтральные слова (словосочетания), передающие название точно определяемых понятий, относящихся к определенной области науки. Использование в текстах такой лексики дает возможность наиболее четко, точно и экономно излагать содержание данного предмета и обеспечивает правильное понимание существа трактуемого вопроса. В экономически значимых текстах, которые и являются информационной базой для принятия серьезных бизнес решений, термины несут основную семантическую нагрузку, занимая главное место среди прочих общелитературных и служебных слов.

Такое использование экономически значимой научной и «околонаучной» терминологии позволяет довольно эффективно использовать для анализа подобных текстов электронные толковые словари, тезаурусы, онтологии соответствующих предметных областей, на которых обоснованно должны базироваться задачи Text Mining, Opinion Mining и Information Retrieval (см. рис. 1)

Сегодня задачи выявления семантических эквивалентов и Opinion mining решаются изолированно, с использованием, в подавляющем большинстве, статистических и, в редких случаях, комбинированных методов. Использование семантических методов хотя и позволяет получать более точные результаты, но развитие подобных моделей, с использованием нейронных или семантических сетей, представляют довольно сложный и трудоемкий процесс.

Предлагается для решения задачи Opinion mining использовать семантически ориентированные методы классификации (sentiment classification), с добавлением оценочных правил, представляющих позитивные, негативные и нейтральные оценки характеристик слов (рис. 2). Используется комбинированный семантико-статистический метод, семантическая составляющая в котором основывается на использовании положительной и отрицательной связи аргументов (слов и словосочетаний) с объектом анализа [Osgood, 1957].

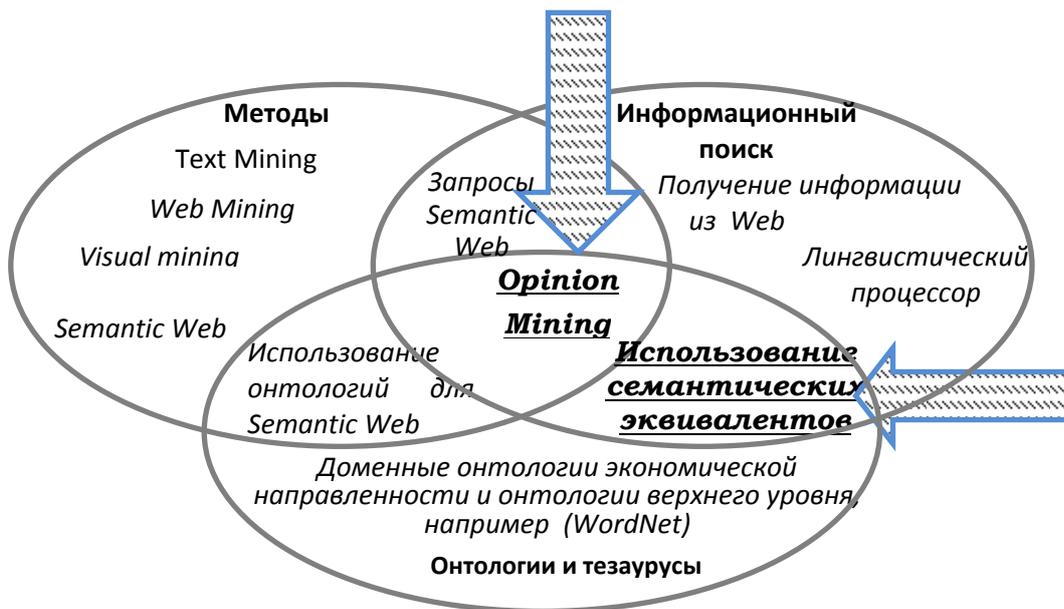


Рис. 1. Общая структура формирования информационного обеспечения задачи принятия бизнес решений при обработке слабоструктурированной экономически значимой информации.

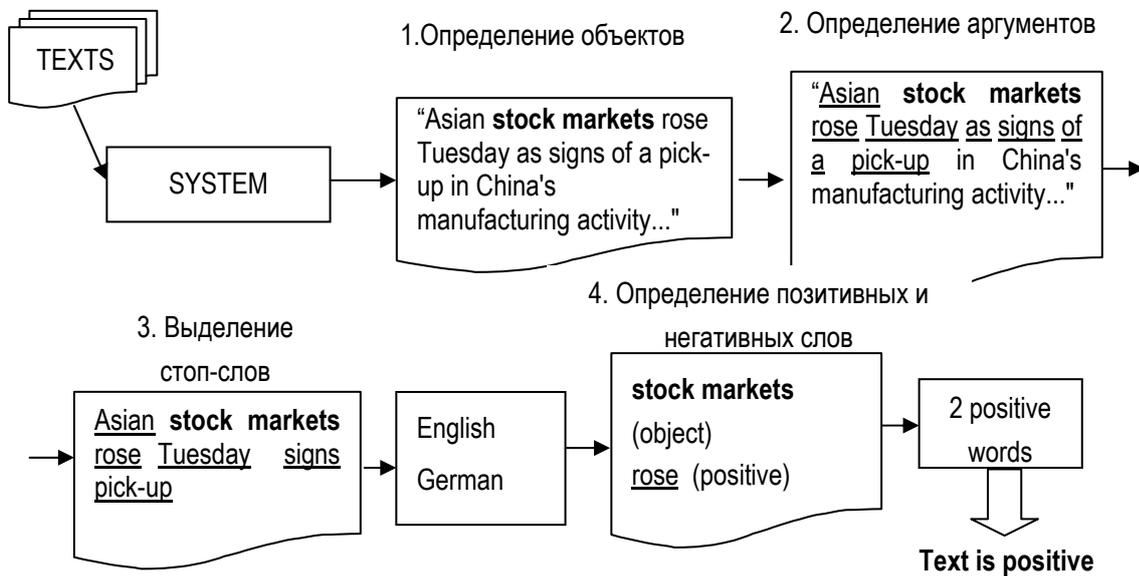


Рис. 2 Структурная схема семантически-ориентированной задачи классификации Opinion mining

Для выявления семантической силы положительного или отрицательного мнения по отношению к тексту определяется атрибутивная оценка объектов (субъектов) небольших фрагментов текста (фраз или предложений). Определение объектов мнений и аргументов, выражающих классификационные признаки мнений, а также их эквивалентов базируется на предварительном осуществлении предлингвистического анализа, стеминга и формального синтаксического анализа [Хайрова, 2010], в результате которых осуществляется нормализация лексем. Нормализация, или приведение лексем к канонической форме, осуществляется для устранения несущественных различий между последовательностями символов, в результате чего множество словоформ предложения преобразуется в список терминов.

На следующем этапе для повышения степени релевантности результата используемой технологии обрабатываются отрицательные слова и стоп слова языка. Отрицательные слова представляют собой четко очерченное множество шумовых слов, подобных словам 'I', 'me', 'itself', 'being', 'after', 'about', 'can' и т.п. английского языка или подобные слова любого естественного языка, которые исключаются из списка аргументов.

Обработка отрицаний предполагает использование шаблона регулярных выражений для таких форм как 'can't', 'mustn't' и др. и отдельную обработку слова 'not':

if ('not' before arg (positive)) \Rightarrow arg (negative) and
if ('not' before arg (negative)) \Rightarrow arg (positive)

Для определения позитивности (аналогично негативности) предложений текста, по отношению к объекту принимаемого бизнес решения, определяется принадлежность нормализованных лексем слов, находящихся на расстоянии k от объекта, аргументами или характеристиками данного слова. Основываясь на работе [Osgood, 1957], в качестве оптимального значения выбрано $k \leq 4$.

Разработанный тезаурус, базирующийся на свободно распространяемом словаре английских слов Minqing Hu and Bing Liu [Bing, 2005], включает 6800 позитивных и негативных английских существительных, глаголов (3 основных используемых формы: инфинитив, past и 3-я форма единственного числа), прилагательных и наречий.

Целостная оценка мнения текста по анализируемому объекту вычисляется статистически с помощью определения веса в тексте соответственно позитивных или негативных, по отношению к запрашиваемому объекту, предложений:

$$O_p = \frac{\sum_{i=1}^k S_i^p}{\sum_{j=1}^n S_j} \quad (1)$$

где O_p — вес позитивной оценки текста, S_j — любые предложения текста, содержащие объект исследования, n — число фраз или предложений текста, содержащих объект исследования, S_i — предложения текста, положительный аргумент в которых находится на расстоянии ≤ 4 от исследуемого объекта оценки, k — число таких фраз или предложений в анализируемом тексте. Аналогичным образом вычисляется вес негативной оценки текста.

Структура используемого тезауруса

В наиболее полных открытых мультиязычных онтологиях типа STW Thesaurus for Economics (by Information Center for Economics), WordNet (by Princeton University), Finance ontology (by Eddy Vanderlinden), Ruthes (Russian Center for information researches), Gemet (by EEA and Eionet), SentiWordNet используются отношения, покрывающие довольно большой класс семантических задач: symmetrical, associativity, homonymy, synonymy, hyponymy, hyperonymy, measuring, symmetrical, association, equivalence, antonymy, troponymy, asymmetrical, value judgments [Соловьев, 2006].

Для решения задачи выявления характеристик текстов, передающих субъективное отношение говорящего (пишущего) к содержанию или адресату речи, с учетом возможности разноязычных высказываний с использованием смысловых эквивалентов, используется тезаурус, в котором явно определены следующие связи между концептами:

1. Синонимия (synonymy) — отношение, связывающее термины, являющиеся синонимами
2. Эквиваленты (equivalence) — отношение, связывающее переводные эквиваленты одинаковых концептов, на разных языках.
3. Ценностное суждение (value judgments) — отношение между объектами и словами, которые описывают определенный эмоциональный оттенок объекта.
4. Антоним (antonymy) — отношение, связывающее слова и словосочетания с противоположным значением.

Под концептом понимается абстрактная группа, коллекция или выборка сущностей, которая может включать отдельных представителей, другие классы или и то, и другое. На рисунке 2 приведен структурный фрагмент тезауруса, для концепта «economic area» и возможных позитивных, негативных и нейтральных суждений для английских и немецких семантических эквивалентов. На рисунке представлены следующие отношения:

1. Synonymy (economic area \Rightarrow economic geography, economic area \Rightarrow industrial area);
2. Equivalence (economic area \Rightarrow Wirtschaftsgebiet (de), economic area \Rightarrow Wirtschaftsraum (de), economic area \Rightarrow Wirtschaftsregion (de));
3. Antonym (small \Rightarrow big, expensive \Rightarrow cheap,...);
4. Value judgment: positive (economic area \Rightarrow big); negative (economic area \Rightarrow unproductive); neutral (economic area \Rightarrow interesting).

Онтология представляет собой формальное описание знаний в виде множества концептов в предметной области и связей между этими концептами, используемое для решения большого класса задач, в том числе онтология может быть использована для описания и вывода сущностей предметной области. Среди множества языков, используемых в настоящее время для описания онтологий (CycL, F-Logic, KL-ONE, LOOM, OCML, DAML-OIL, Web Ontology Language (OWL), RDF (Schema), SQL) наиболее популярными по-прежнему остаются OWL, RDF, SQL.

Информационное обеспечение задачи Opinion Mining, использующее мультиязыковые семантические эквиваленты, представлено тезаурусом, реализованным в виде реляционной базы данных. Выбор языка SQL обоснован возможностью описания определенных выше отношений между концептами на разных языках, простотой использования и наличием инструментов перехода от языка SQL к языку RDF и наоборот.

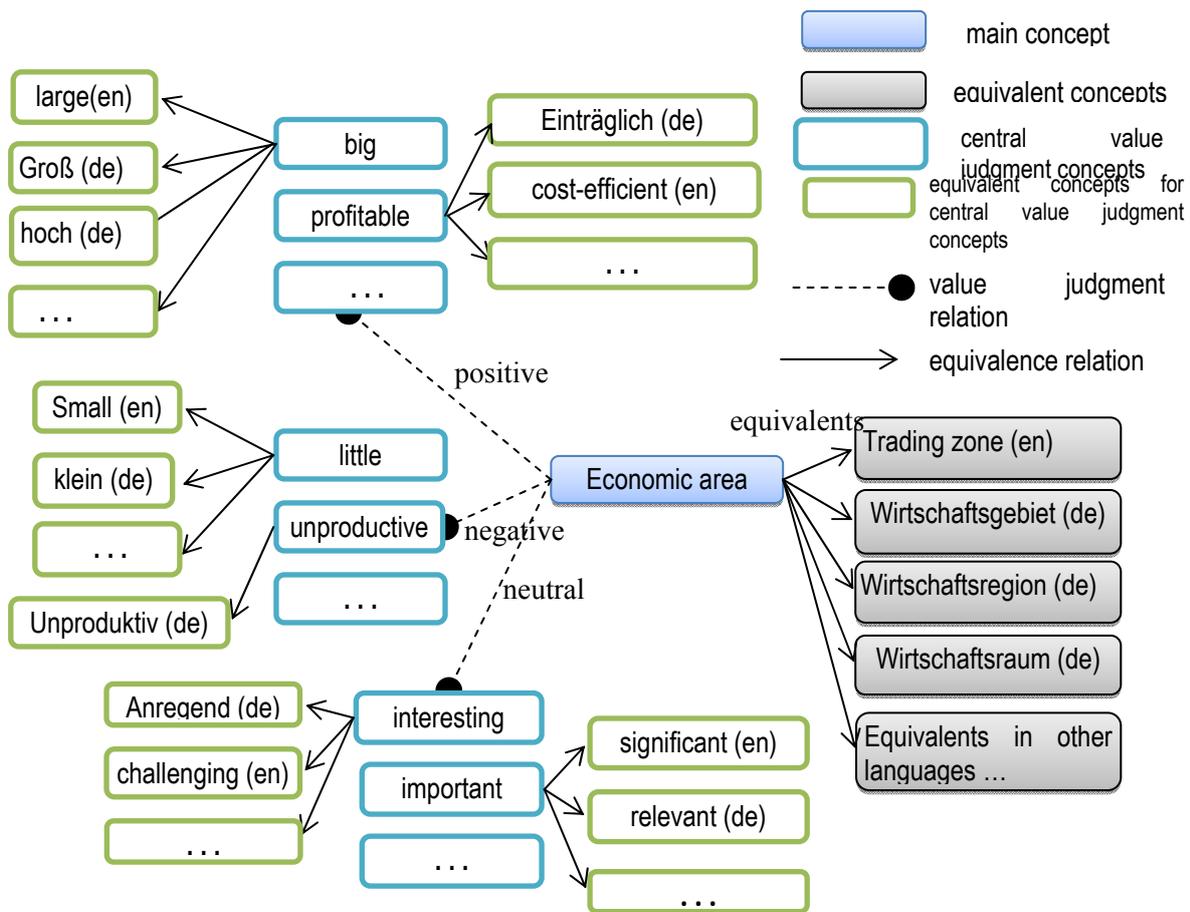


Рис. 2. Структурный фрагмент тезауруса, используемого для решения задачи Opinion Mining, с выделением семантических эквивалентов

Программная реализация модели

Для тестирования алгоритма будет использован инструмент Natural Language Toolkit, обладающий преимуществом свободного распространения и включением большого количества встроенных функций обработки языков.

Для проверки работы алгоритма были протестированы несколько десятков текстов различной экономической направленности – RSS - рассылок на английском и немецком языках. В подавляющем большинстве случаев (87%) классификация текста рассмотренной системой, использующей средства opinion mining, совпала с мнением эксперта о положительном (отрицательном) мнении автора текста об анализируемом субъекте (объекте).

Выводы

Результатом данного исследования является разработка информационной технологии, позволяющей осуществлять классификацию текстов Web-ресурсов пользовательского контента с применением в качестве основания классификации извлеченных знаний о семантически ориентированном мнении, содержащемся в тексте по отношению к тому или иному объекту или субъекту анализа. Предложенная

технология включает в себя методы Opinion Mining, позволяющие автоматизировать извлечение из текста знаний об отношении автора (позитивном, негативном, нейтральном) к рассматриваемому предмету анализа и методы идентификации многоязычных семантических эквивалентов. Применение данных методов базируется на специально разработанном тезаурусе, структура которого в явном виде реализует отношения синонимии, эквивалентности, антонимии и ценностных суждений между концептами. Тезаурус, представляющий информационное обеспечение задачи Opinion Mining, использующей мультиязычные семантические эквиваленты, реализован в виде реляционной базы данных.

Программная реализация модели показала возможность ее использования при классификации английских и немецких текстов экономической направленности по отношению к определенному объекту или субъекту анализа.

Литература:

[Bing, 2005] Liu, Minqing Hu and Junsheng Cheng. "Opinion Observer: Analyzing and Comparing Opinions on the Web" Proceedings of the 14th international World Wide Web conference (WWW-2005), May 10-14, 2005, in Chiba, Japan.

[Ding, 2008] Ding Xiaowen, Liu Bing and Philip S Yu. . A Holistic Lexicon-Based Approach to Opinion Mining. Proceedings of the first ACM International Conference on Web search and Data Mining (WSDM'08), 2008.

[Esuli, 2006] Andrea Esuli, Fabrizio Sebastiani. SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining. In Proceedings of the 5th Conference on Language Resources and Evaluation (LREC 2006), Genova, IT, 2006, pp. 417-422.

[Gupta, 2009] Vishal Gupta, Gurpreet S. Lehal A Survey of Text Mining Techniques and Applications / Journal of Emerging Technologies in Web Intelligence, Vol 1, No 1 (2009), 60-76, Aug 2009.

[Hany, 2011] Hany M. Harb, Khaled M. Fouad, Nagdy M. Nagdy Semantic Retrieval Approach for Web Documents / International Journal of Advanced Computer Science and Applications (JACSA), Vol. 2, No. 9, 2011 — pp.67-76.

[Kobayashi, 2007] Nozomi Kobayashi, Kentaro Inui and Matsumoto Yuji. Opinion Mining from Web Documents: Extraction and Structurization. Journal of Japanese society for artificial intelligence, Vol.22 No.2, special issue on data mining and statistical science, pages 227-238, 2007.

[Osgood, 1957] Osgood, Charles E., Suci, George J., and Tannenbaum, Percy H. 1957. The Measurement of Meaning. Urbana: university of Illinois Press.

[Нгюк, 2012] Нгюен Ба Нгюк, Тузовский А.Ф. Обзор подходов семантического поиска / Доклады ТУСУРа. Управление, вычислительная техника и информатика. № 2 (22), часть 2, 2010. — С. 234-237.

[Оробинская, 2010] Оробинская Е.А., Кочуева З.А. Технологии TEXT MINING: обзор методов и задач обработки смысловой информации // Информационные технологии / Вестник ХНТУ №2(38), 2010. с. 348-353.

[Соловьев, 2006] Соловьев В. Д., Добров Б.В., Иванов В.В., Лукашевич Н.В. Онтологии и тезаурусы. Учебное пособие. Казань, Москва. 2006. — 157с.

[Хайрова, 2010] Хайрова Н. Ф., Тарловский В. А. Использование семантико-ориентированного лингвистического процессора для добытия новых знаний из потока документов корпоративной информационной системы / Вісник Національного технічного університету «ХПІ». Збірник наукових праць. Тематичний випуск «Системний аналіз, управління та інформаційні технології». — Х.: НТУ «ХПІ». — 2010. — № 67. — С. 132-138.

Сведения об авторах

Хайрова Нина – доцент кафедры интеллектуальных компьютерных систем Национального технического университета «Харьковский политехнический институт», ул. Фрунзе, 21, Харьков, 61002, Украина e-mail: nina_khajrova@yahoo.com

Научные интересы: искусственный интеллект, обработка знаний, автоматическая обработка текстов

Шаронова Наталья – профессор, заведующий кафедрой интеллектуальных компьютерных систем Национального технического университета «Харьковский политехнический институт», ул. Фрунзе, 21, Харьков, 61002, Украина e-mail: nvsharounova@mail.ru

Научные интересы: искусственный интеллект, математическое моделирование, автоматизированные библиотечные системы

КОНЦЕПЦИЯ РЕФОРМЫ СИСТЕМЫ НАЦИОНАЛЬНОГО ОБРАЗОВАНИЯ ОСНОВАННАЯ НА МЫШЛЕНИИ

**Александр Сосницкий, Виталий Хоменко, Денис Кравченко, Мария Сосницкая,
Наталья Сосницкая, Наталья Кравченко**

Резюме: Статья посвящена проблеме современного классического Образования, основанного на знаниях, принципиально исчерпавшего свой потенциал вследствие ограниченной биологической способности людей воспринимать современные объемы знаний, что в перспективе ведет к множественным социальным проблемам и конфликтам. Очевидно, что дальнейшее развитие системы Образования должно основываться на иной принципиальной основе, в первую очередь активирующей ресурсы мышления людей. Однако для расчета таких процессов отсутствует адекватная формальная теория живых (интеллектуальных) явлений. В работе показывается способ решения этих проблем применением Универсальной Модели для обоснованного вывода понятия Образования и абсолютизации всей системы производных абстрактов. Предлагается реформа национальной системы Образования из 4-х последовательно развивающихся этапов, конечный из которых направлен на гениализацию людей достижением предельного мышления посредством приближения к абсолютным свойствам Вселенной. Выведены основные принципы гениального Воспитания и Образования, которые при профессиональной педагогической конкретизации уже сегодня могут ложиться в основу новой организации учебных процессов и дидактических материалов к ним с использованием имеющихся вполне доступных электронных сред и средств обучения.

Ключевые слова: реформа Образования, система национального Образования, электронные технологии Образования, Мировые стандарты Образования, технология гениального Воспитания и Образования.

ACM Classification Keywords: *H.0 Information systems – General*

Универсальная Модель для Решения Сложных Фундаментальных Проблем

Современная Наука оказалась в относительном скрытом кризисе, проявляющемся в сочетании сильного развития сравнительно простых областей и хронического застоя сложных областей. Проницательный анализ показывает, что первые области относятся к неживой природе, хорошо изученной классической Наукой, а вторые - к живой природе, для которой фактически отсутствует адекватная теория.

Причина проблемы лежит в принципиальной ограниченности современной научной базы для формализации сложных Вселенских явлений и совершенно очевидно, что этого должна произойти очередная научная революция, подобная ранее прошедшим фундаментальным реформам в физике, математике и др. Эта реформа сродни известным попыткам создания объединяющей теории физических полей и должна решать несравненно более сложную проблему создания единой теории неживой и живой природы и, более того, объединить все научные области в единую универсальную Модель Природы.

Очевидно, что такой объединяющей теорией должна стать некая Универсальная Модель, из которой некая Универсальная Методология, как часть этой Модели, должна выводить все сущее нашей Вселенной. Создание такой Модели и Методологии является грандиозной задачей, но Наука обречена на ее решение для устранения актуальных глобальных проблем, количество которых бесконечно растет с каждым временем.

В Украине впервые разработаны концептуальные основы такой Модели и Методологии, универсальные в полном смысле этого слова, представленные в работах [Sosnitsky, 2007-2009], [Сосницкий, 2011]. Они претендуют на роль единого научного стандарта, и имеют приложения во всех научных и прикладных областях. Фактически они являются высшей мета-теорией, позволяющей (пока вручную) выводить вторичные концепции прикладных теорий отдельных областей как основы последующей формализации их отдельных сущих. В данной статье она применяется для актуальной проблемы Образования людей.

Проблема Классического Образования, Основанного на Знаниях

Образование людей является важной третьей составляющей Науки как копии Вселенной в Человечестве (рис. 1), которая должна удовлетворять трем основным требованиям:

- внешняя непротиворечивость Вселенной;
- внутренняя непротиворечивость копии;
- доступность всем субъектам Человечества.

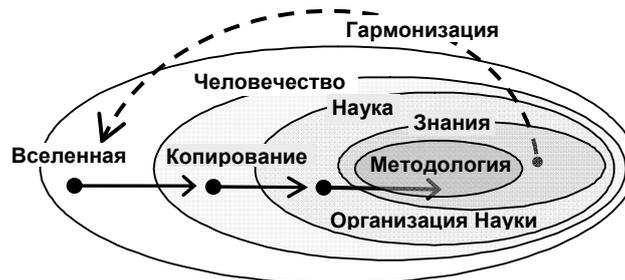


Рисунок 1. Общая схема Науки

Поскольку Вселенная, Человечество и Наука состоят из отдельных взаимно подобных сущих, то это относится к каждому из них. При этом указанная триада взаимосвязана так, что каждое достижение в Науке порождает прогресс в соответствующей области Человечества и наоборот, каждый недостаток в Науке порождает проблему в соответствующей области Человечества (рис. 2). Поскольку господствует правило невидимости субъекту неведомого ему сущего, проблемы трудно идентифицируются в Науке и легко в существовании, где их всегда много во все времена.

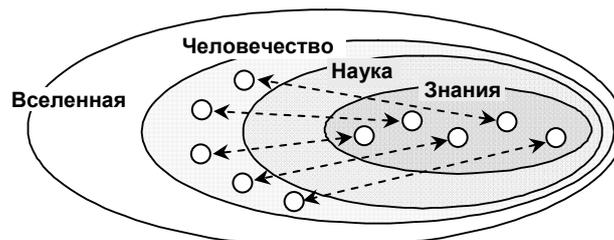


Рисунок 2. Иллюстрация взаимного соответствия недостатков (белые пятна) Науки и Человечества

Последнее есть хороший идентификатор научных недостатков. Если в области естественных наук достигнуты удивительные результаты вплоть до границ существования нашей Вселенной, то в гуманитарной области обычны вполне очевидные отклонения на всех уровнях от личности до международных отношений, которые являются платой за наши незнания.

В данной статье из всего спектра недостатков Науки и Образования ограничимся центральной проблемой исчерпания ресурса классической схемы Образования, основанной на знаниях, вследствие принципиальной ограниченности человека усваивать знания свыше некоторого присущего данному биологическому виду объема и темпа, которые в растущем объеме производит современная Наука.

Вследствие этого недостатка Человечество вынуждено 1) увеличивает сроки Образования вплоть до непрерывного пожизненного, 2) повышает интенсивность обучения, 2) применяет более эффективные дидактические средства, 3) ограничивает другие естественные потребности людей, 4) сокращает производительный период людей, 5) повышает пенсионный возраст, 6) занимается продлением жизни, и др., что перенапрягает его биологическую способность и порождает многие социальные проблемы. Очевидно, что 1) все попытки косметического реформирования классического Образования лишь консервируют кризис и 2) проблема может быть решена только в рамках иной парадигмы Образования.

Проблема Образования имеет Всемирный характер и обусловлена следующими системными факторами:

- Ускоряется рост объема и уровня мировой Науки, расширяющий спектр образованности людей от нуля и до высших уровней (рис. 3). Все большая часть людей вследствие биологической и социальной недостаточности не может достигать высших уровней и «выпадает» из растущих требований цивилизации, увеличивая относительно необразованную часть общества как источник различных форм криминала (дословный смысл этого слова – неправильное мышление) вплоть до терроризма, что повсеместно в мировом масштабе наблюдается уже сейчас (рис. 4).

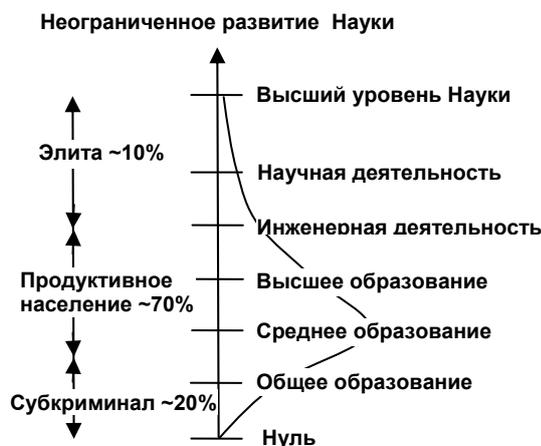


Рисунок 3. Характерное распределение Образования в современном обществе

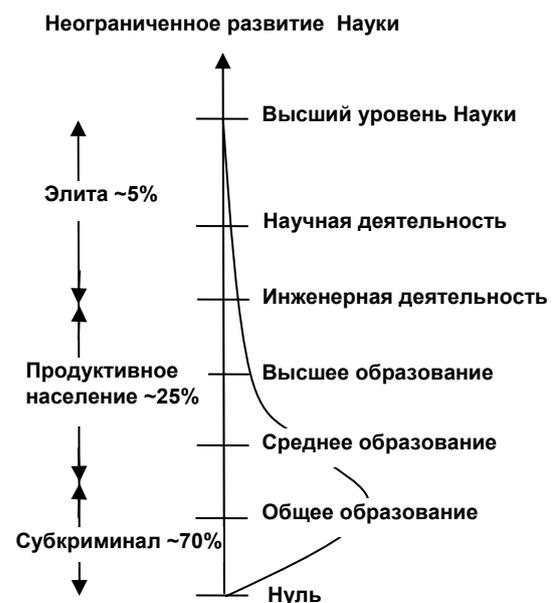


Рисунок 4. Тенденция распределения Образования в недалеком будущем

– Быстрый рост Знаний уже с давних времен делает невозможным их полноценное освоение. Сегодня наметилась тенденция сокращения учебных дисциплин в пользу введения новейших актуальных непрерывно возникающих прогрессивных теорий. Приближается предел «насыщения» фактологической памяти людей, которая все более замещается компьютерными информационными ресурсами, несколько смягчающими, но не решающими проблему.

Идея Образования, Основанного на Мышлении

Наша Вселенная устроена так, что, открывая проблему, она всегда предоставляет средства ее решения, которые предстоит выявить и применить. Идея данной работы состоит в представлении людей Вселенскими сущими и выводе из Универсальной Модели полноценной концепции их предельного Образования с привлечением всех возможных ресурсов Вселенной (рис. 5). Для этого надо предварительно иметь такую Модель и, в ее рамках, формальные определения всех релевантных этому выводу понятий. Основы такого подхода впервые заложены в вышеуказанных работах.



Рисунок 5. Общая схема вывода понятия Образования

Полученная концепция фактически станет методологической основой предельного (гениального) Воспитания, позволяющая коренным образом изменить людей, семьи, коллективы, государства и Человечество в целом и при достижении абсолютного уровня использования Вселенских ресурсов решить все известные и будущие проблемы Цивилизации. Наиболее полный эффект от новой методологии получится на уровне наций, которые получат радикальное развитие и преимущество во Всемирной Цивилизации.

Центральной частью этой идеи есть переход от обучения знаниям к обучению мышлению, позволяющему сократить объем передаваемых знаний за счет их получения логическим выводом (как части знаний) внутри субъектов (рис. 6). Проблема мышления всегда была, есть и будет и определяет в конечном итоге все внутренние и Мировые проблемы.

При рефлексорном обучении (знаниям) человек получает неполные копии взаимодействующих с ним сущих (информацию). Однако имеются дополнительные стадии обучения, следующие из Универсальной Модели и неполно используемые в классическом обучении, такие как понимание и исследование.

Переход к этим стадиям посредством Универсальной Методологии открывает гигантский дополнительный ресурс Образования людей без чрезмерных материальных затрат и уже сейчас доступный каждой национальной системе Образования даже скромных государств. Его проблема большей частью состоит в коррекции субъективной (организационной) как наиболее инерционной части такой системы.



Рисунок 6. Схема сокращения получаемых знаний посредством виртуализации

Принципы Образовательной Реформы

Реализация образовательной реформы возможна постепенным последовательно-параллельным (в рамках каждой учебной дисциплины, класса, учебного заведения, региона, Министерства) переводом современной классической бумаго-ориентированной системы Образования на новую техническую, информационную, организационную, методологическую и дидактическую базу, включающую четыре компоненты (этапа):

- электронизация Образования;
- демократизация Образования;
- всемирная стандартизация Образования;
- гениализация (абсолютизация) Образования.

Компонента Электронизации Образования

Целью этой образовательной компоненты есть полный перевод бумаго-ориентированных образовательных технологий на электронные формы для автоматизации всего комплекса образовательных действий. Он выполнен посредством общего популярного компьютерного обеспечения, и удовлетворяет основной части образовательных требований.

Впервые за последние годы компьютерные средства стали всемирно массово доступными по всем релевантным всеобщему Образованию показателям:

- приемлемая стоимость для большей части учащихся и учебных заведений;
- наличие бесплатной информации и программного обеспечения;

-
- широкое распространение мобильных дешевых глобальных и локальных сетей;
 - апробированность средств дистанционного обучения;
 - всеобщая компьютерная грамотность учащихся и преподавателей.

Электронизация впервые поддерживает:

- бесконспектные электронные формы обучения;
- полную доступность учебных программ и дидактических материалов;
- оперативное совершенствование индивидуальных материалов;
- компоновку материалов из авторских оригиналов (евронорма);
- комплексирование материалов (текст, графика, движение, цвет, видео, звук);
- всемирные и национальные базы дидактических материалов;
- поддержание высокого стандарта качества учебы;
- дистанционное обучение (независимость от местоположения учащегося);
- непрерывную доступность материалов учащимся по принципу 24/7/12/365;
- оперативный контроль учебного процесса;
- непрерывную оперативную связь учащихся и учителей;
- непрерывную многоуровневую интеграцию обучения;
- оперативную учебную отчетность;
- Всемирную стандартизацию Образования;
- технологию гениального Воспитания и Образования.
- Электронизация комплексно улучшает все важные учебные показатели:
- в среднем вдвое повышает качество обучения;
- контролирует по принципу «доступно все, возможно понятное»;
- повышает активность и самостоятельность мышления учащихся;
- развивает профессионализм и компетенцию;
- универсализирует учебные дисциплины и глобализирует Образование в целом;
- вызывает взаимное уважение участников и стабилизирует учебный процесс;
- облегчает работу всех участников учебного процесса;
- легко тиражирует учебные технологии;
- обеспечивает полную прозрачность и контроль обучения на всех уровнях;
- инициирует сквозную систему национального Образования.

Компонента Демократизации Образования

Целью этой образовательной компоненты есть создание эффективных гибких адаптивных прогрессивных сквозных общенациональных структур управления Образованием, во главе которых ставится триада «эксперт - учитель - ученик».

Радикально меняется расстановка участников образовательного процесса. Во главе указанной триады как наиболее компетентные ставятся эксперты и учителя. Министерство координирует деятельность этой триады на всех уровнях для удовлетворения потребности нации в Образовании.

Ведущие специалисты для надежности объединены в специализированные экспертные советы под контролем Министерства и несут полную личную ответственность за качество национальной системы Образования. Деятельность спецсоветов контролируется центральным ревизионным советом. Все советы полностью прозрачны и обязаны обоснованно реагировать на все без исключения вопросы в пределах своей компетенции с использованием административного аппарата Министерства.

Экспертные советы по специальностям:

- определяют квалификационные требования к этапам Образования;
- определяют содержание Образования;
- определяют траектории сквозного Образования в системе учебных заведений;
- разрабатывают образовательные стандарты специальностей и учебных заведений;
- разрабатывают типовые программы, методики и учебники дисциплин;
- разрабатывают сертификаты качества преподавателей и учебных заведений;
- контролируют результаты Образования в учебных заведениях.

Министерство координирует все национальные образовательные процессы:

- определяет перспективные потребности государства в специалистах;
- ведет классификаторы специальностей;
- координирует экспертные советы по специальностям;
- утверждает сквозные учебные планы специальностей;
- ведет электронные базы контрольных материалов и сертификатов учащихся;
- материально обслуживает систему Образования.

Региональные компоненты Министерства:

- утверждают рабочие программы и дидактические материалы согласно установленным нормативам.

Учебные заведения (учителя):

- разрабатывают на основе типовых программ рабочие программы, дидактические и контрольные материалы в соответствии с реальными условиями учебных заведений, которые заносятся в базу данных Министерства;
- реализуют индивидуальные учебные процессы с тестом через базу Министерства.

Учащиеся:

- изучают и выбирают траектории получения Образования по специальностям;
- выполняют учебный процесс в учебных заведениях и сдают контрольные материалы и результаты итогового контроля в базу данных Министерства;
- получают индивидуальные электронные сертификаты.

Отменяются все лишние травмирующие формы контроля знаний учащихся кроме как заинтересованными учебными организациями, несущими ответственность за качество добавляемых ими знаний. Принимающие учебные заведения проверяют оценки индивидуальных электронных сертификатов повторным подтверждающим контролем по релевантным учебным дисциплинам со сравнением в базе контрольных материалов Министерства.

Система Образования становится демократической рыночной, и учащиеся самостоятельно выбирают объем, качество и стоимость Образования в соответствии со своими желаниями и возможностями и имеют на это полное право в наш ограниченный ресурсами век как ведущие участники образовательного процесса при поддержке государства.

Компонента Всемирной Стандартизации Образования

Целью этой образовательной компоненты есть совместимость национальной системы Образования с мировыми образовательными стандартами, поскольку ни одна страна не способна конкурировать против всего мира. Совместимость достигается участием в разработке и применении известных специальных и общих международных стандартов национального Образования и позволяет сильнее интегрировать и

демократизировать мировую Науку и Образование. Всемирная стандартизация открывает возможность свободного Образования в любых мировых научных и учебных заведениях.

Компонента Гениализации (Абсолютизации) Образования

Целью этой образовательной компоненты есть предельное использование Вселенских ресурсов для наивысшего Образования в границах возможностей людей. Вселенная и ее ресурсы бесконечны, но каждый из них в отдельности, как и их сумма, в каждой ситуации ограничены. Поэтому надо искать, определять объемы и использовать ресурсы в конкретных условиях каждого человека.

Проблема Образования непрерывно решается с момента возникновения Человечества до сих пор, что само по себе свидетельствует о неполноте достигнутого качества образовательных технологий, использующих лишь часть образовательных ресурсов. Для более полного использования ресурсов надо проводить специальные исследования, невозможные в рамках современной Науки вследствие чрезвычайной сложности явления Образования, непостижимого эмпирическими методами. Такие ресурсы могут быть лишь выведены из Универсальной Модели, что показывается в данной работе.

Гениализация является конечной центральной целевой компонентой образовательной реформы, в которой мышление становится главным ресурсом повышения Образования людей, позволяющего радикально согласовать его с требованиями современной Цивилизации.

Реализация Образовательной Реформы

Реформа предполагает немедленную мягкую эволюцию из текущего в целевое состояние с сосуществованием обеих форм Образования по мере развития всех четырех компонентов на следующих региональных уровнях:

- личный (воспитатель, учитель, преподаватель);
- подразделение (детская группа, класс, кафедра, факультет (институт));
- учебное заведение (детский садик, школа, лицей, колледж, университет);
- национальный регион (село, район, город, область);
- Министерство (государство).

Реформа потребляет материальные ресурсы, однако большей частью в рамках государственных бюджетов в сочетании с вполне доступными накопительными затратами учащихся, родителей и спонсоров со сроком технического обновления порядка 5-ти лет. Наиболее проблемной есть консервативная субъективная часть – эксперты, педагоги, администрация, система подготовки и переподготовки педагогических кадров.

Релевантная Образованию Система Исходных Вселенских Понятий

Приведем исходную релевантную онтологию последней целевой компоненты. Гениализация следует из концепции Вселенной, происходящей из единственного исходного абстрактного сущего Ничто, производящего все бесконечное многообразие нашей Вселенной в виде полной совокупности прямо или косвенно связанных сущих, что впервые принципиально позволяет получать абсолютно точные модели частных сущих, одной из которых есть Образование. Сущее есть часть Вселенной, выделяемая некоторой связью как единое целое (рис. 7).

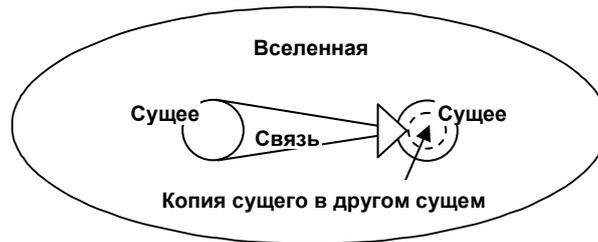


Рисунок 7. Общая схема сущего как единого целого, выделяемого некоторой связью

Связь (существование) есть (обычно частичная) копия одного сущего (объект) в другом сущем (субъект). Сущие обычно имеют множественные входящие и исходящие связи, включая их абстрактные и реальные компоненты (рис. 8). Свойство есть структура связей сущих. Таким образом, сущее есть не более, чем бесконечная вложенная вглубь структура свойств (связей).

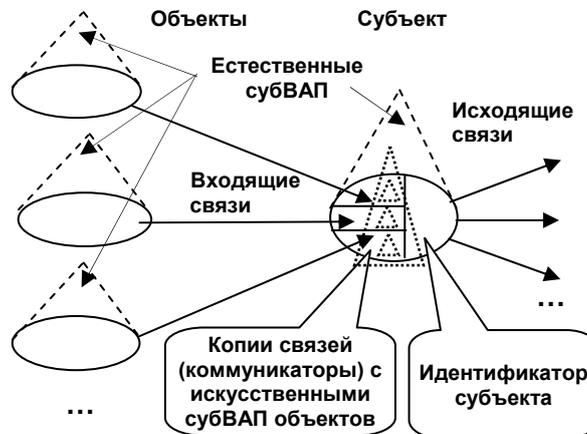


Рисунок 8. Общая схема сущего

Разные комбинации повторения связей порождают следующую классификацию копий (рис. 9):

- полная (повторяются все связи оригинала);
- частичная (повторяется часть связей оригинала, но могут присутствовать и чуждые оригиналу связи);
- белая (повторяется только часть связей оригинала);
- серая (вместе с белыми появляются связи, отсутствующие в оригинале);
- черная (все связи в копии отсутствуют в оригинале).

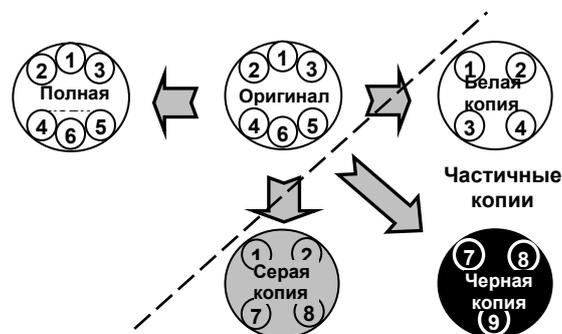


Рисунок 9. Схема классификации копий

Сущие (ниже ПВМ-Комплекса) имеют обобщенную внутреннюю структуру из 8-ми классов компонентов (рис. 10): 1-3) Комплексы Пространства, Времени и Материи (ПВМ-Комплекс), 4) вложенные сущие, 5) внутренние связи, 6) движение, 7) законы и 8) цель.

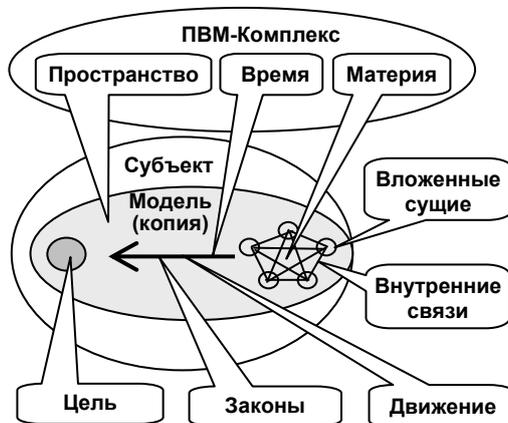


Рисунок 10. Внутренняя классификация сущего

Понятие связи производит понятие Познания как циклическое во Времени развитие копии объекта в субъекте (рис. 11). Каждый цикл состоит из 4-х стадий продвижения копии объекта к субъекту: 1) информация, 2) знание, 3) понимание и 4) исследование. Информация есть копия объекта на границе с субъектом, знание есть копия объекта внутри субъекта, понимание есть всевозможное согласование (гармонизация) знания объекта с другими знаниями внутри субъекта, исследование есть получение дополнительных знаний об объекте согласно субъективной цели развития познания.

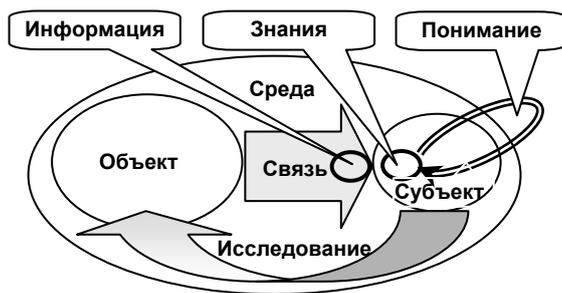


Рисунок 11. Общая схема Познания

Понятие Познания производит понятие Образования как полной совокупности всех знаний субъекта об объектах окружающего Мира (рис. 12). Знания наследуют историю познания субъектом и обычно получают неоднородными и разделяются на относительно малосвязные домены. Гармонизация таких доменов является одним из образовательных ресурсов.

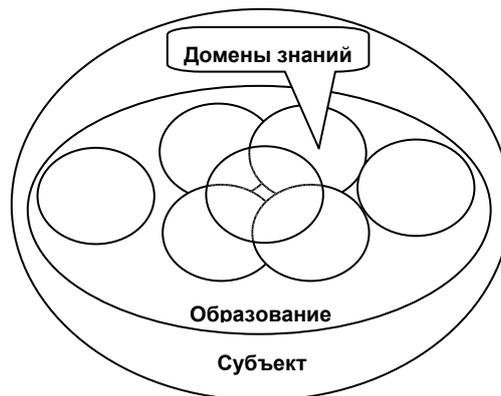


Рисунок 12. Общая схема Образования

Познание еще выполняется по частной от исследования схеме обучения посредством косвенного копирования копий объектов через промежуточные субъекты (учители) (рис. 13).

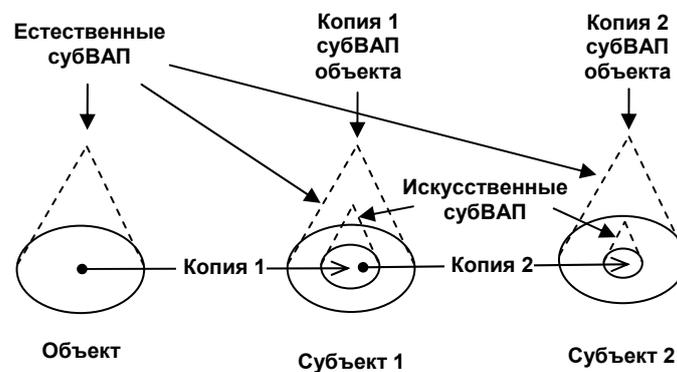


Рисунок 13. Общая схема обучения

Вселенная есть система со временем, разделенным на три части: Прошлое, Настоящее и Будущее. Все связи строго реализуются в движущемся из Прошлого в Будущее Настоящем. Интеллект (живые существа) позволяет посредством мышления (запоминания и прогнозирования) устанавливать виртуальные связи на некотором временном интервале вокруг движущейся точки Настоящего (радиус мышления) с последующей реализацией в этой точке (рис. 14).

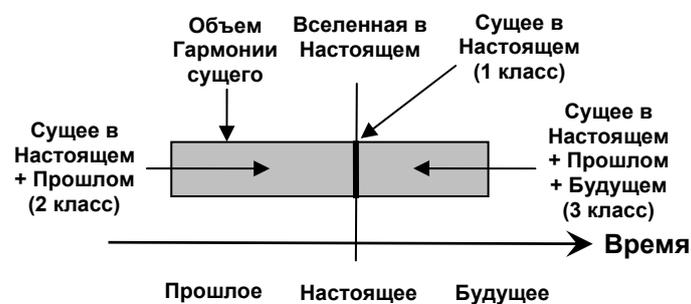


Рисунок 14. Общая схема Интеллекта

Механизм (основная процедура действия) мышления непрерывно циклически работает по схеме на рис. 15. Его характерным признаком есть обратный ход (возврат в реальное) субъективного времени.

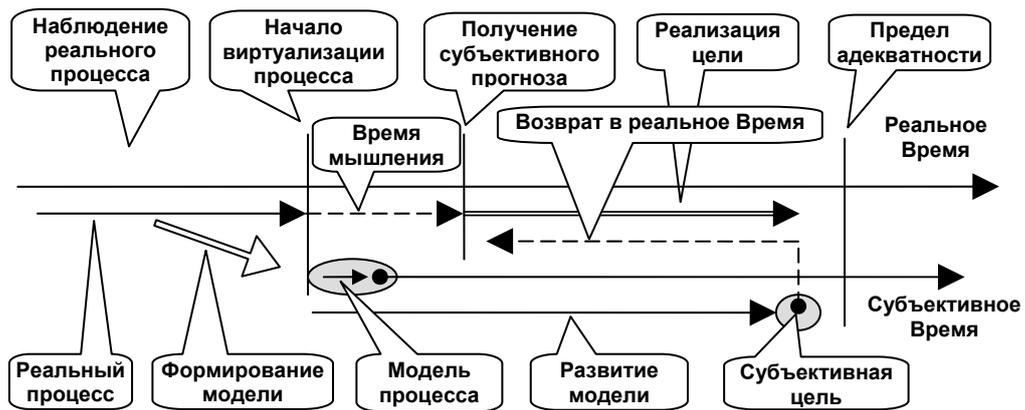


Рисунок 15. Схема акта мышления

Понятие Познания производит понятие существования как циклическое во времени развитие субъектов согласно их Образованию, гармонизирующему субъекты с окружающим Миром (рис. 16). В зависимости от качества управления процесса субъектом существование объекта может получать состояния эволюции, проблемы, противоречия, кризиса, революции или катастрофы, также имеющие свои концептуальные определения.

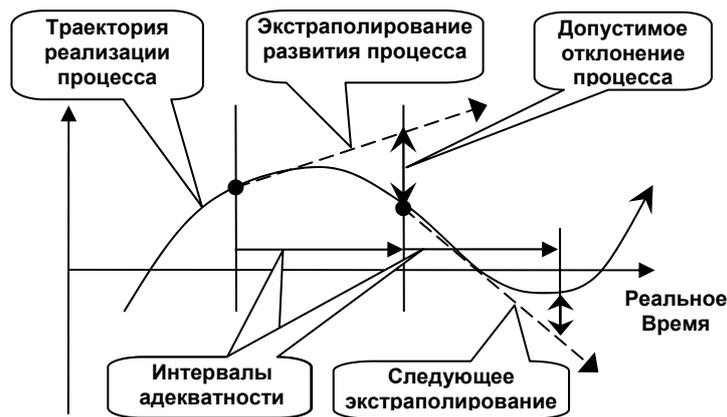


Рисунок 16. Схема процесса существования объекта

Каждое временное сущее имеет структуру процесса, начинающего от определения и завершающегося его разрушением при исчерпании имеющегося гармонического ресурса с порождением производных более гармонизированных сущих-процессов (рис. 17).

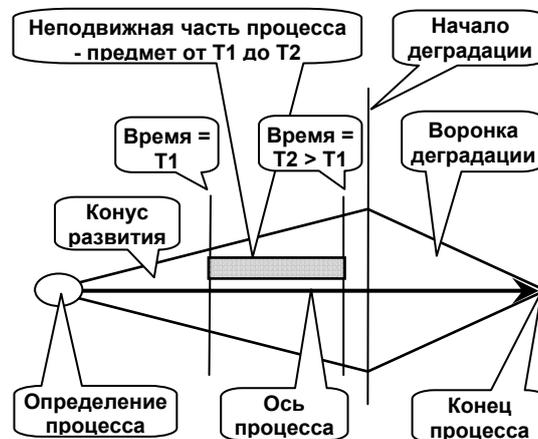


Рисунок 17. Общая схема сущего как процесса

Приведенная онтология следует из четырех принципов конструкции Вселенной: Абсолютность, Абстрактность, Гармоничность и Познаваемость, в свою очередь следующих из вышеуказанной концепции происхождения Вселенной из Ничто.

Абсолютность есть свойство единства и единственности Вселенной, все сущие которой есть ее (обычно частичные) копии. Это обеспечивает наличие, наследование, общность и фундаментальную непротиворечивость абсолютных (точных) свойств сущих.

Абстрактность есть свойство разделения Вселенной на абстрактную и части в форме Вселенской Абстрактной Пирамиды (ВАП), начинающуюся с исходного абстракта Ничто и последовательно производящего систему Абстрактных Миров (АМ), завершающуюся Реальным Миром (РМ) (рис. 18). Абстракты последовательно взаимно встраиваются сверху-вниз в ВАП и невидимы на производных нижних уровнях, поскольку не существуют в чистом виде и требуют выделения по образцу (характерным признакам) для идентификации. Каждое сущее имеет собственную субВАП как часть ВАП.

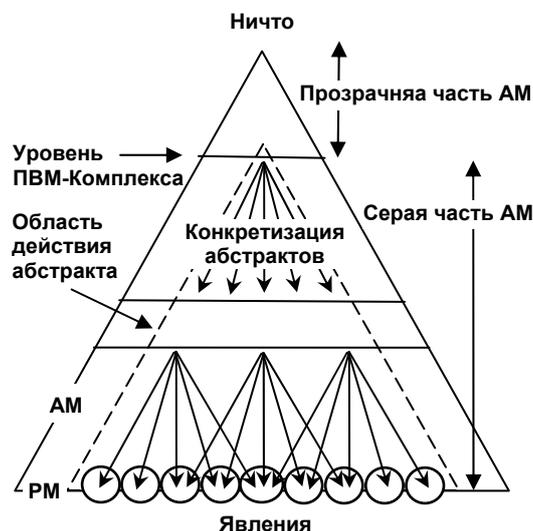


Рисунок 18. Абстрактная схема Вселенной

Гармоничность есть стремление каждого сущего образовывать и развивать всевозможные связи (копии других сущих) согласно высшему СверхЗакону Вселенной: каждое сущее, предоставленное само себе, стремится в состояние максимальной Гармонии, под которым (приближенно) понимается количество постоянных связей сущего.

Это состояние определяет цель сущего, которая есть (обычно неточная) проекция СверхЗакона на уровень сущего посредством других косвенных сущих. Поскольку связи в общем случае подвижны, вместе с ними перемещается и Гармония как некоторая Вселенская величина, имеющая качественное и количественное измерение.

Познаваемость есть способность каждого сущего в меру его текущих свойств неограниченно познавать другие сущие, начиная с любого стартового сущего вплоть до полного познания Вселенной неограниченным развитием свойств познающего сущего.

Принципы Гениального Воспитания и Образования

Универсальная Модель должна стать новым научным и образовательным стандартом, приближение (абсолютизация) к которому позволяет радикально улучшить качество существования за счет привлечения более точных Вселенских понятий. Модель производит новые принципы организации учебно-воспитательного процесса и дидактических материалов учебных дисциплин, универсальность которых следует из сверхвысокого уровня абстрагирования и замкнутости (наличие единого исходного центра) применяемых понятий, обеспечивавших внешнюю и внутреннюю непротиворечивость педагогической теории.

- Абсолютность как принадлежность сущих к Абсолюту (высшая степень текущей Гармонии). Символизирует точность и полноту сущих.
- Абстрактность как принадлежность сущих к ВАП. Символизирует иерархию и онтологию (происхождение) сущих.
- Абстрагирование как выделение высших абстрактов сущих. Символизирует генеалогию (онтологию) и централизацию сущих.
- Конкретизация как выделение производных в ВАП сущих от некоторого сущего. Символизирует подчиненную часть ВАП (субВАП) сущих.
- Гармоничность как связность сущих. Символизирует наличие трех частей сущих: происхождение (генеалогия, онтология, история, причина), определение (суть) и наследование (следствие, результат).
- Гармонизация как СверхЗакон стремления сущих к неограниченному Образованию и развитию связей.
- Гармония как Вселенская величина, определяющая потенциал (возможность) гармонизации сущих.
- Познаваемость как возможность неограниченного полного копирования сущих от нулевого начального уровня развития субъектов. Символизирует неограниченность связности и развития субъектов.
- Структурные (качественные) как исходные приоритетные методы познания. Символизирует вторичность других методов познания, поскольку сущие есть не более, чем вложенные структуры связей.
- Сущее как часть Вселенной, выделяемое некоторой связью как единое целое. Символизирует вариативность сущих в пределах внутреннего разделения Вселенной.
- Связь как копия (и копирование) одного сущего в другом сущем. Символизирует вариативность сущих как частичных копий других сущих.
- Свойство (понятие, качество) как структура связей сущих. Символизирует вариативность свойств сущих в пределах их определений.
- Направленность есть несимметричность свойств от одних сущих к другим. Символизирует необратимость процессов сущих.
- Сущее как бесконечная вложенная структура свойств. Символизирует вариативность сущих в зависимости от их внутренних связей.
- Прозрачность (полнота) связей как мера абсолютности сущих. Символизирует степень абсолютности сущих.
- Чистота (белизна) связей как мера внутренней непротиворечивости сущих. Символизирует степень устойчивости развития (гармонизации) сущих.
- Гомотропность сущих как мера единства свойств. Символизирует частичную однородность (одинаковость) Вселенских сущих и их копий (описаний, формализаций).
- Конечность абстрактных сущих. Символизирует возможность их полного познания конечными субъектами.
- ПВМ-структура временных сущих (ниже ПВМ-Комплекса) из 3(4)-х Комплексов: Пространство, Время и Материя (Предметы и Процессы). Символизирует соответствующее происхождение и структуру сущих.

- Классность компонентной структуры временных сущих. Символизирует единую форму временных сущих из 8-ми классов компонентов.
- Бесконечная вложенность временных сущих. Символизирует принципиальную возможность абсолютного познания Вселенной каждым субъектом.
- Движение как изменение (развитие) сущих. Символизирует принципиальную возможность достижения абсолютной Гармонии сущих.
- Закономерность как направленность движения сущих. Символизирует наличие механизмов достижения абсолютной Гармонии сущих.
- Целенаправленность как генеральное направление движения сущих. Символизирует наличие искомого конечного состояния движения.
- Процессность как структура временных сущих во Времени. Символизирует единую структуру процессов временных сущих.
- Центральность сущего в форме (абстрактного) определения. Символизирует наличие абстрактного начала каждого сущего.
- Динамичность как наличие действующих на сущие (внешних и внутренних) сил (гармонизирующих и разгармонизирующих). Символизирует основной механизм изменения Гармонии сущих.
- Детерминизм (пассивность) как зависимость сущих от входных внешних связей. Символизирует влияние окружающей среды.
- Свобода (активность, независимость) как зависимость сущих от внутренних связей. Символизирует самостоятельность сущих.
- Регулярность (однородность, цикличность, повторяемость) как равенство копий в (ПВМ) Комплексах. Символизирует эквивалентность результатов гармонизации сущих в некоторых частях Комплексов.
- Неточность как неполнота (не белое копирование) копий. Символизирует расхождение копии с оригиналом.
- Виртуальность как превышение в копиях копируемых оригинальных свойств сущих. Символизирует преодоление существующих проблем на копиях этих проблем.
- Замкнутость как ограниченность сущих, их предков и потомков некоторой частью ВАП. Символизирует целостность (точность, абсолютность) сущих в соответствующих замыканиях ВАП (субВАП сущих).

Подобные принципы бесконечно генерируются Универсальной Моделью и излагаются в высокоабстрактной интерпретации для поддержки совместимости с Моделью, но имеют и вполне прикладные выражения, которые должны формулироваться отвечающей за это педагогической Наукой [Лазарев, 2011], [Кравченко, 2011].

Например, несоблюдение принципа замкнутости объектов ведет к неоднородному познанию и малосвязной доменной структуре знаний (рис. 19), которые ограничивают распространение логических выводов границами доменов и принципиально снижают гармонизирующую способность субъекта.

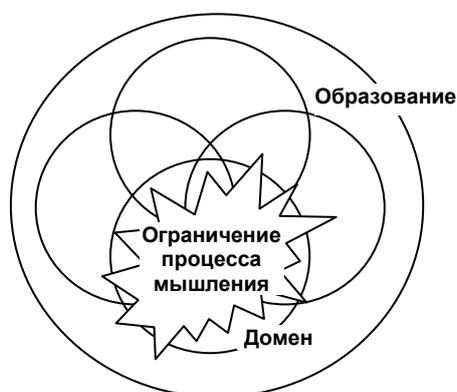


Рисунок 19. Схема ограничения процессов мышления неоднородностью Образования

Получается ограниченная система неточных понятий, ведущая при многошаговых логических процессах к отклонениям от абсолютности вплоть до противоположных решений, примерами которых переполнена наша действительность.

Заключение

1. Современная национальная система Образования, основанная на знаниях, принципиально исчерпала себя вследствие наличия ограниченной биологической способности людей воспринимать современные знания, что в перспективе ведет к множественным социальным проблемам и конфликтам.
2. Дальнейшее развитие национального Образования должно основываться на иной принципиальной основе, в первую очередь активирующей ресурсы мышления людей. Однако для расчета таких процессов отсутствует адекватная формальная теория живых (интеллектуальных) явлений.
3. В данной работе показывается способ решения этих проблем применением Универсальной Модели для обоснованного вывода понятия Образования и абсолютизации (адекватизации) всей системы производных абстрактов.
4. Предлагается реформа национальной системы Образования из 4-х последовательно развивающихся этапов, конечный из которых направлен на гениализацию людей достижением предельного мышления посредством приближения к абсолютным свойствам Вселенной.
5. Выведены основные принципы гениального Воспитания и Образования, которые при профессиональной педагогической конкретизации уже сегодня могут ложиться в основу принципиально новой организации учебных процессов и дидактических материалов к ним с использованием доступных электронных сред и средств обучения.

Библиография

- [Sosnitsky, 2007] Sosnitsky A. V. Conceptual programming: program as a copy of subject domain // Theoretical and applied aspects of program systems development. – Kiev, 2007, pp. 305-311.
- [Sosnitsky, 2008] Sosnitsky A. V. The Conception of Abstract Programming // Sino-European Engineering Research Forum, Glasgow, UK, - 2008, Volume 1. pp. 34-40.
- [Sosnitsky, 2009] Sosnitsky A. V. Harmonious Foundations of Intelligence // Communications of SIWN, Vol. 7, May 2009. pp. 66-72.
- [Сосницкий, 2011] Сосницкий А. В. Искусственный интеллект и радикальная реформа современной Науки // Искусственный интеллект. – Донецк. 2011, №1, с. 91-105.
- [Сосницкий, 2011] Сосницкий А. В. Искусственный интеллект и Универсальная гармоническая Методология Познания // Искусственный интеллект. – Донецк. 2011, №2, с. 70-83.
- [Лазарев, 2011] Лазарев М. І., Хоменко В. Г., Чуприна Г.П. Теоретичні та методичні засади навчання програмних засобів захисту інформації на основі подвійного дидактичного узагальнення. – Донецьк: ЛАНДОН-XXI, 2011. 173с.

[Кравченко, 2011] Кравченко Н. В. Интернет технології як інструмент розвитку творчого потенціалу інженера-педагога. / Н.В. Кравченко, Д.В. Кравченко // Збірник наукових праць Бердянського державного педагогічного університету (Педагогічні науки). - №2. – Бердянськ: БДПУ, 2011. – с. 138 - 142.

Информация Авторов



Александр В. Сосницкий – доцент, Институт Образовательных Инженерно-Педагогических Технологий, Бердянский Университет, Бердянск, 71100, Украина, электронная почта: sosnitsky.ukr@yandex.ua, тел. +380665093093

Главные Области Научного Исследования: Абсолютная Теория Вселенной, Реформа Науки и Образования, новая Научная Методология, Искусственный Интеллект, Компьютерные Технологии



Виталий Г. Хоменко – доцент, директор Института Образовательных Инженерно-Педагогических Технологий, Бердянский Университет, Бердянск, 71100, Украина, электронная почта: v_g_homenko@ukr.net, тел. +380505907432

Главные Области Научного Исследования: Автоматизированные Системы Управления в Обучении и Технологических Процессах



Денис В. Кравченко - директор компании "Интеллект - компьютер", Бердянск, 71100, Украина, электронная почта: kdv@icom.com.ua, тел. +380615345065

Главные Области Научного Исследования: Современные Образовательные Технологии, Искусственный Интеллект



Мария А. Сосницкая – аспирантка Института филологии, Бердянский Университет, Бердянск, 71100, Украина, электронная почта: sosnitsky.maria@yandex.ua, тел. +380992634900

Главные Области Научного Исследования: Теория литературы, Информационные методы в филологии, Компьютерные технологии в Образовании.



Наталья В. Кравченко – доцент, Институт Образовательных Инженерно-Педагогических Технологий, Бердянский Университет, Бердянск, 71100, Украина, электронная почта: nvkravchenko@mail.ru, тел. +380615345065

Главные Области Научного Исследования: Математическая Теория Управления, Теория Стабильности, Современные Образовательные Теории

TABLE OF CONTENT

Формализация проблемы извлечения знаний из естественно языковых текстов	
Александр Палагин, Сергей Крывый, Николай Петренко, Дмитрий Бибилов	203
Компромиссное решение задач условной оптимизации	
Альберт Воронин	219
Моделирование временных зависимостей в интеллектуальных системах поддержки принятия решений на основе прецедентов	
Александр Еремеев, Павел Варшавский, Иван Куриленко	227
Последовательный анализ конечно малых возмущений линейных моделей при компьютерном моделировании	
Алексей Волошин, Владимир Кудин.....	240
К построению модели «демократического капитализма»	
Алексей Волошин, Анна Сидорук	250
Исследование и моделирование нейросетевого метода обнаружения и классификации сетевых атак	
Адил Тимоев, Александр Браницкий.....	257
Моделирование интеллектуальных систем в процессах эволюционного обучения.	
Ирина Трегубенко	266
Информационная технология применения семантически ориентированных методов классификации задачи opinion mining	
Нина Хайрова, Наталья Шаронова.....	273
Концепция реформы системы национального образования основанная на мышлении	
Александр Сосницкий, Виталий Хоменко, Денис Кравченко, Мария Сосницкая, Наталья Сосницкая, Наталья Кравченко	283
Table of content	300