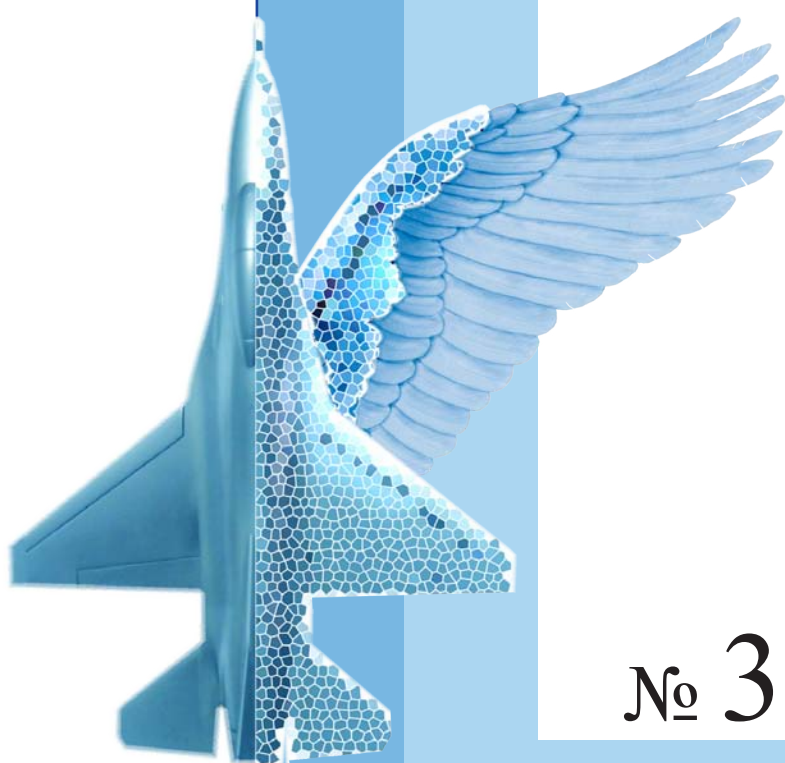


ОНТОЛОГИЯ ПРОЕКТИРОВАНИЯ

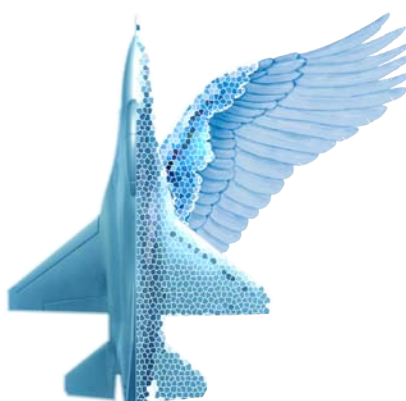


№ 3(13)/2014

ОНТОЛОГИЯ
ПРОЕКТИРОВАНИЯ

Научный журнал

№ 3(13)



EDITORIAL BOARD – РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Anatoly I. Belousov	Белоусов Анатолий Иванович, д.т.н., профессор, СГАУ, г. Самара
Nikolay M. Borgest	Боргест Николай Михайлович, к.т.н., профессор СГАУ, член ИАОА, г. Самара
Vladimir V. Golenkov	Голенков Владимир Васильевич, д.т.н., профессор, БГУИР, г. Минск
Vladimir I. Gorodetsky	Городецкий Владимир Иванович, д.т.н., профессор, СПИИРАН, г. Санкт-Петербург
Yuri R. Valkman	Валькман Юрий Роландович, д.т.н., профессор, МНУЦ ИТиС НАН и МОН Украины, г. Киев
Stanislav N. Vasiliev	Васильев Станислав Николаевич, академик РАН, ИПУ РАН, г. Москва
Vladimir A. Wittich	Виттих Владимир Андреевич, д.т.н., профессор, ИПУСС РАН, г. Самара
Nicholay G. Zagoruiko	Загоруйко Николай Григорьевич, д.т.н., профессор, ИМ СО РАН, г. Новосибирск
Alexander S. Kleshchev	Клещёв Александр Сергеевич, д.ф.-м.н., профессор, ИАПУ ДВО РАН, г. Владивосток
Valery A. Komarov	Комаров Валерий Андреевич, д.т.н., профессор, СГАУ, г. Самара
Sergey M. Krylov	Крылов Сергей Михайлович, д.т.н., профессор, СамГТУ, г. Самара
Victor M. Kureichik	Курейчик Виктор Михайлович, д.т.н., профессор, Технологический институт ЮФУ, г. Таганрог
Lyudmila V. Massel	Массель Людмила Васильевна, д.т.н., профессор, ИСЭМ СО РАН, г. Иркутск
Dmitry A. Novikov	Новиков Дмитрий Александрович, член-корреспондент РАН, ИПУ РАН, г. Москва
Semyon A. Piyavsky	Пиявский Семён Авраамович, д.т.н., профессор, СГАСУ, г. Самара
George A. Rzevski	Ржевский Георгий Александрович, профессор, Открытый университет, г. Лондон
Peter O. Skobelev	Скобелев Петр Олегович, д.т.н., НПК «Разумные решения», г. Самара
Sergey V. Smirnov	Смирнов Сергей Викторович, д.т.н., ИПУСС РАН, г. Самара
Anatoly V. Sollogub	Соллогуб Анатолий Владимирович, д.т.н., профессор, РКЦ «Прогресс», г. Самара
Peter I. Sosnin	Соснин Петр Иванович, д.т.н., профессор, УлГТУ, г. Ульяновск
Dzhavdet S. Suleymanov	Сулейманов Джавдет Шевкетович, академик, вице-президент АН РТ, г. Казань
Robert I. Tuller	Таллер Роберт Израилевич, д.филос.н., профессор, СГАУ, г. Самара
Boris E. Fedunov	Федунов Борис Евгеньевич, д.т.н., профессор, ГосНИИ Авиационных систем, г. Москва
Altynbek Sharipbay	Шарипбай Алтынбек, д.т.н., профессор, Институт искусственного интеллекта, г. Астана
Boris Y. Shvedin	Шведин Борис Яковлевич, к.психол.н., ООО «Дан Роуз», член ИАОА, г. Ростов-на-Дону

Executive Editorial Board - Исполнительная редакция

Chief Editor	Smirnov S.V.	Главный редактор	Смирнов С.В.	директор ИПУСС РАН
Executive Editor	Borgest N.M.	Выпускающий редактор	Боргест Н.М.	директор изд-ва «Новая техника»
Editor	Kozlov D.M.	Редактор	Козлов Д.М.	профессор СГАУ
Technical Editor	Simonova A.U.	Технический редактор	Симонова А.Ю.	редактор изд-ва «Новая техника»
Translation Editor	Korovin M.D.	Редактор перевода	Коровин М.Д.	аспирант СГАУ
Proofreader	Shustova D.V.	Корректор	Шустова Д.В.	аспирант СГАУ

CONTACTS – КОНТАКТЫ

ИПУСС РАН

443020, Самара, ул. Садовая, 61.
тел.: +7 (846) 332 39 27, факс.: +7 (846) 333 27 70

Смирнов С.В.
smirnov@iccs.ru

СГАУ

443086, Самара, Московское шоссе 34, корп. 10, кафедра КиПЛА
тел.: +7 (846) 267 46 47, факс.: +7 (846) 267 46 46

Боргест Н.М.
borgest@yandex.ru

Издательство «Новая техника»

443010, Самара, ул.Фрунзе, 145, тел.: +7 (846) 332 67 84, факс: +7 (846) 332 67 81

Сайт журнала: http://agora.guru.ru/scientific_journal/

Журнал включен в Российский индекс научного цитирования (РИНЦ). Номер контракта 64-03/2012.



НАУЧНО-ПРОИЗВОДСТВЕННАЯ КОМПАНИЯ «РАЗУМНЫЕ РЕШЕНИЯ»
<http://www.smartsolutions-123.ru/>

Журнал зарегистрирован в Федеральной службе по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор). Свидетельство ПИ № ФС 77-46447 от 7.09.2011 г.



Отпечатано в издательстве «Новая техника»
Подписано в печать 09.09.2014. Тираж 300 экз.

© Все права принадлежат авторам публикуемых статей
© Издательство «Новая техника», 2011-2014

СОДЕРЖАНИЕ

От редакции О ПЕРВЫХ НАЧАЛАХ	5
В.И. Городецкий, О.Н. Тушканова ОНТОЛОГИИ И ПЕРСОНИФИКАЦИЯ ПРОФИЛЯ ПОЛЬЗОВАТЕЛЯ В РЕКОМЕНДУЮЩИХ СИСТЕМАХ ТРЕТЬЕГО ПОКОЛЕНИЯ	7
И.М. Зацман, П.С. Бунтман ПРОЕКТИРОВАНИЕ ИНДИКАТОРОВ МОНИТОРИНГА В СФЕРЕ НАУКИ: ТЕОРЕТИЧЕСКИЕ ОСНОВАНИЯ И МОДЕЛИ	32
Л.С. Глоба, М.В. Ковальский ПОВЫШЕНИЕ РЕЛЕВАНТНОСТИ ПРИ ПОИСКЕ ВЕБ-СЕРВИСОВ В МУЛЬТИОНТОЛОГИЧЕСКОЙ СРЕДЕ	52
В.М. Дёмкин, А.В. Соснин, С.С. Сусманова ОНТОЛОГИЧЕСКИЕ МОДЕЛИ СИТУАЦИЙ В ЗАДАЧАХ КОМПЬЮТЕРНОГО КОНТРОЛЯ ЗНАНИЙ ИНОСТРАННОГО ЯЗЫКА	63
В.И. Левин СТАБИЛЬНОСТЬ НЕОПРЕДЕЛЁННЫХ ЗАДАЧ ОПТИМИЗАЦИИ	77
С.А. Пиявский ПРОСТОЙ И УНИВЕРСАЛЬНЫЙ МЕТОД ПРИНЯТИЯ РЕШЕНИЙ В ПРОСТРАНСТВЕ КРИТЕРИЕВ «СТОИМОСТЬ–ЭФФЕКТИВНОСТЬ»	89
РЕКОМЕНДУЕМЫЕ ИЗДАНИЯ	103

CONTENT

From the Editors	
ABOUT THE FIRST PRINCIPLES	5
V.I. Gorodetsky, O.N. Tushkanova	7
ONTOLOGY-BASED USER PROFILE PERSONIFICATION IN 3G RECOMMENDER SYSTEMS	
I. M. Zatsman, P.S. Buntman	32
DESIGNING INDICATORS FOR MONITORING IN SCIENCE: THEORETICAL FRAMEWORK AND MODELS	
L.S. Globa, M.V. Kovalskyi	52
INCREASING WEB SERVICES DISCOVERY RELEVANCY IN THE MULTI-ONTOLOGICAL ENVIRONMENT	
V.M. Dyomkin, A.V. Sosnin, S.S. Susmanova	63
ONTOLOGICAL MODELS OF SITUATIONS APPLIED TO THE TASK OF COMPUTER-AIDED TESTING OF FOREIGN LANGUAGES PROFICIENCY	
V.I. Levin	77
STABILITY OF UNCERTAIN OPTIMIZATION PROBLEMS	
S.A. Piyavsky	89
A SIMPLE AND UNIVERSAL METHOD OF DECISION MAKING WITHIN THE SCOPE OF CRITERIA OF «COST AND EFFICIENCY»	
RECOMMENDED BOOKS	103



*Прежде всего, тебе я скажу о первых началах,
В коих возникло все то, что ныне мы зрим во вселенной:
Бурное Море, Земля, увлажнённый Воздух и с ними
Тот Титан — Эфир, что кругом облёг мирозданье.*

Эмпедокл, «Состав мироздания»
из поэмы «О природе»¹

Дорогой наш читатель, уважаемые авторы и члены редакционной коллегии!

Продолжая наши традиции, вспомним оригинального **Эмпедокла** (Empedocles, ок. 495, Агригент, о. Сицилия – 435 до н. э.) – греческого философа, поэта, мага и целителя. Оригинальность Эмпедокла состоит в его учении о четырёх элементах и в применении принципов Любви и Вражды для объяснения изменения. Он отверг монизм и рассматривал ход вещей как регулируемый скорее случайностью и необходимостью, чем целью. В этом отношении его философия была более научной, чем философия Парменида, Платона и Аристотеля².

Основатель школы красноречия в Сицилии, Эмпедокл первый изобрёл риторику, а свои философские воззрения изложил в поэмах «О природе» и «Очищения». Эмпедокл рассматривал сущностные вопросы бытия: состав мироздания, состав живой природы, жизненные процессы и др., тем самым создавая первые онтологические модели. В учении Эмпедокла в качестве начал были названы четыре традиционные стихии – огонь, воздух, вода и земля. Эти начала вечны, а все изменения видимого мира происходят благодаря их сочетанию друг с другом под воздействием двух сил притяжения (Любви, «Филии») и отталкивания (Вражды, «Нейкос»). Эмпедокл, как и Парменид, полагал, что переход из небытия к бытию невозможен, «рождение» и «гибель» – лишь механистическое «соединение» и «разъединение» вечных элементов.



Empedocle's.

Космогонический цикл по Эмпедоклу имеет 4 фазы: 1) «любовь» – все элементы слиты воедино, образуя однородный неподвижный «шар»; 2) «вражда» проникает в «шар» и вытесняет «любовь», разъединяя разнородные элементы; 3) «любовь» возвращается, соединяя разнородные элементы и разъединяя однородные; 4) фаза возникновения живых существ, которая распадается на 4 ступени: 4.1) отдельные члены, неспособные соединиться в организм; 4.2) неудачные соединения членов – монстры; 4.3) бисексуальные существа, неспособные к половому размножению; 4.4) полноценные животные.

Большое влияние на греческую философию (особенно, на атонизм) оказала теория ощущений Эмпедокла: от воспринимаемого объекта непрерывно отделяются материаль-

¹ Эллинские поэты VIII—III вв. до н.э. - М., Ладомир, 1999. H.Diels — W.Kranz. Die fragmente der Vorsokratiker. 10. Aufl., Berlin, 1961 / Перевод Г.И. Якубаниса в переработке М.Л. Гаспарова - <http://ancientrome.ru/antlittr/empeokles/nature.htm>

² Рассел, Б. История западной философии. Часть первая. ДОСОКРАТИКИ. Глава VI. Эмпедокл. Russell, B. A History of Western Philosophy. Page XI. Simon & Schuster, Inc., 1972 - <http://philosophy.ru/library/russell/01/01.html#6>

ные «истечения», проникающие в «поры» органов чувств. Теория «пор и истечений» имела у него универсальный характер и объясняла также физические и физиологические процессы.

«Воля Судьбы такова, что *присуща всем тварям разумность*» - эта крылатая фраза Эмпедокла в поэме «О природе» в разделе «Жизненные процессы» во многом определяет важное качество человеческой сущности, способной идентифицировать концепты, структурировать среду, классифицировать и атрибутировать находящиеся в ней сущности. Современная онтология в её информационном, технологическом статусе не способна пока воспроизвести важное качество, присущее живой сущности. Но накапливаемый экспериментальный и теоретический материал даёт надежду на близкий результат не только в научной, но и практической прикладной сфере.

...

Первые начала интеграционного научного направления онтологии проектирования были заложены, когда был поставлен учебно-научный курс для магистрантов в СГАУ (2010 год), и когда был создан наш с Вами журнал (2011 год). Эти начала лишь первые зёрна кристаллизации уже накопленных в разных областях знаний, которые, надеемся, послужат катализатором «объединения разнородных элементов», которые по метафорическому утверждению Эмпедокла способна и осуществляет Любовь.

Исследуя область онтологии проектирования с разных сторон научной направленности - с философской, психологической, методологической, информационно-технологической и прикладной нормализованной деятельности - журнал продолжает публиковать работы, формализующие важные процедуры, операции и опции в онтологии проектирования.

В этом номере в статье мэтра интеллектуальных систем, ведущего научного сотрудника Санкт-Петербургского института информатики и автоматизации РАН В.И. Городецкого и его аспирантки О.Н. Тушкановой рассматриваются онтологии и персонификация профиля пользователя в рекомендующих системах третьего поколения. К рекомендующим системам третьего поколения относят системы, которые вырабатывают решения на основе семантических моделей интересов и предпочтений пользователя, принимают во внимание мотивацию и причины, которые побуждают конкретного пользователя предпочитать то или иное решение, а также учитывают семантику контекста, сопутствующего принятию решений.

Семиотический подход к описанию итераций проектирования новых индикаторов мониторинга, включая верифицируемые его варианты, и генерации новых экспертных знаний в процессе проектирования исследуется в статье московских исследователей под руководством заведующего отделом Института проблем информатики РАН, доктора наук И.М. Зацмана и биолога П.С. Бунтмана.

Профессор из Пензы В.И. Левин продолжил серию статей в области оптимального проектирования в условиях неопределенности, а исследователи из Национального исследовательского университета «Высшая школа экономики» (Нижний Новгород) под руководством В.М. Дёмкина на пути к становлению новой парадигмы компьютерного тестирования сделали очередной шаг, связанный с «очеловечиванием» поведенческой природы холонов.

Коллеги из Национального технического университета Украины «Киевский политехнический институт», профессор Г.С. Глоба и выпускник университета М.В. Ковальский предложили способ повышения релевантности при поиске веб-сервисов в мультионтологической среде, а известный самарский профессор С.А. Пиявский из архитектурно-строительного университета вводит в онтологию многокритериальной оптимизации новое понятие шансов оптимальности решения и предлагает, тем самым, простой и универсальный метод принятия решений в пространстве критериев «стоимость–эффективность».

Ждём ваших, читатель, откликов на публикации и ваших, авторы, новых результатов!

УДК 004.82

ОНТОЛОГИИ И ПЕРСОНИФИКАЦИЯ ПРОФИЛЯ ПОЛЬЗОВАТЕЛЯ В РЕКОМЕНДУЮЩИХ СИСТЕМАХ ТРЕТЬЕГО ПОКОЛЕНИЯ

В.И. Городецкий¹, О.Н. Тушканова²

Санкт-Петербургский институт информатики и автоматизации РАН, Санкт-Петербург, Россия

¹gor@iias.spb.su, ²tushkanova@iias.spb.su

Аннотация

Рекомендующие системы третьего поколения сейчас находятся в самом начале своего развития. К рекомендующим системам третьего поколения относят системы, которые вырабатывают решения на основе семантических моделей интересов и предпочтений пользователя, принимают во внимание мотивацию и причины, которые побуждают конкретного пользователя предпочитать то или иное решение, а также учитывают семантику контекста, сопутствующего принятию решений. Иначе говоря, основным новшеством таких систем является их ориентация на семантические модели представления и использования знаний, в частности, знаний о персональном профиле пользователя. В настоящее время наиболее естественным и наиболее разработанным способом формализации семантических категорий, которые обычно используются человеком в процессах выработки решений, является онтология. По этой причине в современных и будущих системах, которые называют рекомендующими системами третьего поколения, онтология рассматривается в качестве общей структуры для представления разнообразных и разнотипных знаний. Примерами таких знаний являются, например, знания о персональном профиле пользователя, о контексте принятия решений, об эмоциональном состоянии пользователя при принятии решения и т.п. Данная работа посвящена обзору публикаций, которые рассматривают подходы и модели персонального профиля пользователя в контексте онтологий предметной области рекомендаций, а также варианты их использования в рекомендующих системах. Эти публикации фактически подготовили современный взгляд на онтологию как единую модель концептуализации и представления компонент знаний, используемых в рекомендующих системах.

Ключевые слова: рекомендующие системы, онтология, профиль пользователя, персонификация, контекстно-зависимые рекомендации, фольклсономия, сетевой контекст, автоматизация разработки онтологий.

Введение

Концепция рекомендующих систем как систем прогнозирования решений человека при выборе им товара/услуги появилась в середине 1990-х годов. К настоящему времени эта концепция прошла стандартный путь развития, характерный для систем, основанных на знаниях. Она прошла путь от статистических и эвристических моделей обучения и принятия решений к моделям, в которых используются содержательно интерпретируемые знания, учитывается контекст принятия решений, а также персональные интересы и предпочтения пользователя, чье решение рекомендующая система пытается прогнозировать.

В настоящее время предметом перспективных исследований в этой области являются рекомендующие системы третьего поколения [1], которые акцентируют внимание на семантических моделях представления и использования всех компонент знаний, которые вовлекаются в процесс выработки рекомендаций. С функциональной точки зрения, рекомендующие системы третьего поколения должны вырабатывать решения на основе семантических моделей интересов и предпочтений пользователя, принимать во внимание мотивацию и причины, которые побуждают конкретного пользователя делать тот или иной выбор. Они должны так-

же учитывать семантику контекста, сопутствующего принятию решений пользователем. Системы третьего поколения должны быть способны давать мотивированные объяснения предлагаемым решениям. Они должны учитывать экономические, психологические и другие факторы, связанные с принятием решений [1].

Рекомендующие системы третьего поколения сейчас находятся в самом начале своего развития. Потребовалось около 15 лет для того, чтобы сформировать современную концепцию рекомендующих систем. В ней центральной идеей является использование семантических категорий естественного языка для концептуализации, представления и использования знаний, вовлекаемых в процессы выработки решений. В настоящее время наиболее естественной и наиболее разработанной моделью формализации семантических категорий, которые используются человеком, является модель онтологии. Поэтому в рекомендующих системах третьего поколения онтология рассматривается в качестве общей модели концептуализации и структуры для представления всех компонент знаний, например, знаний о предметной области рекомендаций, знаний о персональном профиле пользователя, о контексте принятия решений, об эмоциональном состоянии пользователя при принятии решения и т.п.

Представление профиля пользователя в виде структур на множестве понятий предметной онтологии становится особенно продуктивным и экономным, если эти понятия являются *причинами*, которые определяют его отношение к выбору того или иного товара/услуги. Такие структуры знаний однозначно и естественным образом задают механизмы принятия решений пользователя, представленные в декларативной форме. Их принято называть активными структурами знаний, или, другими словами, знаниями, которые *способны генерировать действия* (англ. *actionable knowledge*¹) [2].

Данная работа посвящена обзору публикаций, в которых предлагаются онтологические модели интересов пользователя. Эти публикации фактически подготовили современный взгляд на онтологию как единую модель концептуализации и представления компонент знаний, используемых в рекомендующих системах. В разделе 1 даются краткие сведения о рекомендующих системах в контексте их развития, а также сведения о «классических» методах прогнозирования решений пользователя. Раздел 2 поясняет понятие профиля пользователя и его связь с онтологией предметной области рекомендаций. Раздел 3 описывает те модели представления профиля пользователя, которые непосредственно предшествовали онтологическим моделям, однако во многом способствовали принятию этой концепции для перспективных рекомендующих систем. В разделе 4 рассматриваются модели профиля пользователя, которые уже полностью базируются на использовании онтологий.

Следует заметить, что разработка онтологической модели предметной области рекомендаций и, в частности, онтологической модели профиля пользователя требует решения ряда новых специфических проблем, с которыми разработчики онтологий ранее не сталкивались. Это связано с тем, что при построении модели предметной области рекомендаций и профиля пользователя приходится использовать множество разнородных источников информации и данных, в которых, так или иначе, отражаются «следы» присутствия или деятельности пользователя. Эти источники обычно распределены по сети Интернет и могут содержать весьма специфичную информацию. К такой информации относятся, например, тэги (англ. *tags*). Тэги – это семантически осмысленные метки. Ими сам пользователь или другие пользователи обычно метят продукты/услуги, которые могут быть или уже были *объектами выбора* (англ. *items*) тех или иных пользователей. Информация о тэгах обычно является достаточно информативной и семантически богатой. Она особенно информативна для задачи персонификации интересов и предпочтений пользователя. Тэги как элементы естественного языка обычно с

¹ “Propositions that are actionable are those that actors can use to implement effectively their intentions. Actionable knowledge requires propositions that make explicit the causal processes required to produce action. Causality is the key in implementation” [3].

трудом отображаются на понятия, которые имеются в стандартных словарях понятий. Поэтому категоризация тэгов – это весьма нетривиальная задача, которую, однако, нужно решать. Обычно тэги связываются с одним или несколькими понятиями и категориями в специальных словарях, которые принято называть *фольксономиями*. Проблема построения профиля пользователя на основе тэгов рассматривается в разделе 5, в котором также даются краткие сведения о существующих подходах к построению онтологий на основе тэгов. Еще одна специфическая проблема возникает при построении онтологического профиля пользователя в случае рекомендаций, формируемых в контексте социальных сетей. Особенности построения онтологического профиля пользователя и возможные способы учета «сетевого контекста» демонстрируются на примере в разделе 6. Во многих случаях онтологию предметной области рекомендаций приходится строить с использованием множества источников данных. Краткие сведения о предложенных подходах к построению онтологий и профиля пользователя для таких случаев приведены в разделе 7.

В заключении по работе приводится краткая сводка преимуществ онтологических моделей профиля пользователя, а также перечисляются некоторые тенденции в этой области. Кроме того, дополнительно дается краткий перечень проблем, не затронутых в данном обзоре, которые, однако, важны и имеют свою специфику применительно к построению профиля пользователя в рекомендующих системах и требуют специального рассмотрения.

1 Краткие сведения о рекомендующих системах

Рекомендующими системами² называют класс систем принятия решений, которые используют знания об интересах и предпочтениях человека для оценки/прогнозирования его реакции на рекомендацию купить некоторый товар или воспользоваться некоторой услугой [4]. Потенциальная реакция человека оценивается системой с помощью некоторой метрики, которая измеряется обычно в порядковой шкале, например, в целочисленной шкале. Эта метрика характеризует меру уверенности системы в том, что человек отреагирует на ее рекомендацию положительно. Ее принято называть *рейтингом* товара/услуги.

Исследования в области рекомендующих систем начались в середине 1990-х годов. Появление и развитие систем такого типа в значительной степени было обусловлено задачами маркетинга. В настоящее время интерес к практическому использованию рекомендующих систем постоянно возрастает, и главной, хотя и не единственной причиной этого интереса является повышение относительной доли Интернет-торговли в общем объеме продаж товаров и услуг. И это понятно, поскольку потенциальный покупатель товара или потребитель услуги в Интернет-магазине не видит сам товар. При отсутствии рекомендаций, соответствующих его интересам и предпочтениям, он вынужден был бы искать нужное в толстых каталогах или базах данных о товарах или услугах конкретного магазина, отслеживать появление новых товаров и т.д. Очевидно, что ни один нормальный покупатель товара или потребитель услуги не стал бы на это тратить так много времени, а пошел бы в обычный магазин, где все товары выставлены, или в офис, в котором всю информацию о потребной услуге можно быстро получить.

К настоящему времени рекомендующие системы получили уже значительное теоретическое развитие, и нашли широкое практическое применение. Рекомендующие системы первого поколения (1995-2005+), которые и сейчас имеют широкое хождение, использовали модель, основу которой составляли три матрицы [1]. Первая из них, матрица $X = \{x_{ij}\}_{i=1, j=1}^{M, P}$,

² В научной литературе на русском языке равноправно используются два термина: *рекомендующие* (англ. *recommending*) системы и *рекомендательные* (англ. *recommender*) системы. Читатель вправе сделать контекстную замену первого термина, который и используется в работе, на второй, если ему так кажется лучше.

задает по строкам множество имен пользователей $C = (c_1, \dots, c_M)$, а по столбцам — множество имен свойств (атрибутов) пользователей $X = (x_1, x_2, \dots, x_P)$, так что элемент x_{ij} матрицы X задает значение атрибута x_j для пользователя c_i . Вторая матрица, $Y = \{y_{ij} \mid i=1, j=1\}^{N,Q}$, задает по строкам множество имен S_i товаров/услуг $S = (s_1, \dots, s_N)$, а по столбцам — множество $Y = (y_1, \dots, y_Q)$ атрибутов (свойств) y_j этих товаров/услуг. Наконец, третья матрица $R = \{r_{ij} \mid i=1, j=1\}^{M,N}$, которая называется *матрицей рейтингов*, имеет в качестве строк имена пользователей $C = (c_1, \dots, c_M)$, в качестве столбцов — имена товаров/услуг $S = (s_1, \dots, s_N)$. Ее элементом r_{ij} является значение рейтинга товара/услуги s_j , присвоенное ему пользователем c_i . Обычно матрица рейтингов R бывает слабо заполненной, поскольку общее число товаров/услуг всегда значительно больше, чем то их число, которым пользователь реально мог воспользоваться и оценить с помощью рейтинга. Задача рекомендующей системы состоит в том, чтобы спрогнозировать значения рейтинга, который будет присвоен пользователем новому для него товару/услуге в матрице R . Обратим внимание на то, что интересы пользователя здесь представлены в матричной форме, а в модели знаний рекомендующей системы онтология не используется.

В рекомендующих системах первого поколения используются разные методы принятия решений, однако базовыми являются три из них [1]:

- Методы, основанные на фильтрации контента (множества свойств) $Y = (y_1, \dots, y_Q)$, описывающего продукты/услуги, которые потенциально могут быть рекомендованы пользователю. В этих методах прогноз пользовательского рейтинга нового товара/услуги формируется по его сходству с другими товарами/услугами, которые пользователь ранее оценил с помощью рейтинга («скажи мне, какие рейтинги данный пользователь присваивал раньше похожим товарам/услугам, и я скажу тебе значение рейтинга нового товара для него» [5]). Для алгоритмической реализации этого метода должна быть задана форма метрики, по которой вычисляется сходство товаров/услуг. Обычно для этих целей используются различные метрики сходства [6].
- Методы коллаборативной фильтрации, в которых полагается, что текущие предпочтения пользователя будут сходны с предпочтениями других пользователей с похожими интересами по отношению к предлагаемому товару или услуге. В этом случае решающую роль играет сходство пользователей по своим предпочтениям («скажи мне, какой рейтинг присвоили этому товару другие пользователи, интересы которых сходны с интересами данного пользователя, и я скажу тебе, какой рейтинг он присвоит этому товару/услуге» [5]).
- Гибридные методы, которые комбинируют оба названных выше подхода и, возможно, используют еще какие-то новшества.

Важно отметить, что по своей сути все эти методы являются чисто статистическими и весьма слабо учитывают персональные предпочтения конкретного пользователя. Пользователь в них характеризуется выборкой примеров товаров/услуг, которым он присвоил значения рейтинга в той или иной шкале. По этой причине предсказательная сила рекомендующих систем первого поколения оказалась достаточно слабой.

Отличительной чертой рекомендующих систем второго поколения (2003-2014+) является, в первую очередь, учет контекста, который сопутствует рекомендациям. В качестве главных измерений (атрибутов) контекста рассматриваются обычно время, место и социальный контекст. Например, если пользователю рекомендуется фильм, то следует учитывать желаемое время просмотра (рабочий день или выходной день и т.п.), хочет ли пользователь смот-

реть фильм дома или в кинотеатре и, наконец, в каком коллективе он планирует просмотр (с родителями, с друзьями и т.п.). Детальный анализ сущности контекста, моделей его формального представления и учета в процессах выработки рекомендаций можно найти в [7–9], а также в [4].

Рекомендующие системы второго поколения характеризуются и другими возможностями, которыми их предшественники не обладали. К ним относится, например, использование многокритериальных рейтингов и рейтингов, формируемых сразу для групп пользователей. Для оценки товаров/услуг в них можно использовать не только рейтинг, но и простое упорядочение рекомендуемых вариантов выбора. Они способны вырабатывать рекомендации в контексте социальных сетей, использовать тэги в качестве источника знаний и профиле пользователя и т.п. Обобщая сказанное, можно утверждать, что в системах второго поколения появился акцент на использование *знаний об интересах* пользователя и о *контексте конкретной задачи* выработки рекомендаций. Заметим, что для формализации знаний о профиле пользователя в рекомендующих системах второго поколения начали использовать их представление в терминах онтологий³.

В настоящее время наиболее передовые *практические* разработки в области рекомендующих систем соответствуют второму их поколению, и активные научные исследования в интересах развития таких систем продолжают. Однако для исследователей основной научный интерес представляют уже рекомендующие системы третьего поколения [1]. Такие системы, прежде всего, ориентируются на семантические аспекты моделей интересов и профиля пользователя. В них большое значение придается мотивационным аспектам того или иного выбора пользователя и причинам, которые определяют его выбор. Полагается, что рекомендующие системы третьего поколения должны быть способными давать мотивированные объяснения решениям, которые они предлагают. Считается, что рекомендующие системы третьего поколения должны учитывать экономические, психологические и другие факторы, определяющие выбор пользователя [1]. Очевидно, что большинство перечисленных свойств носит семантический характер, акцентирует внимание разработчика на представлении модели интересов и предпочтений пользователя в терминах семантических категорий, обычно используемых человеком в естественном языке.

Такая тенденция наметилась еще в период развития систем второго поколения, однако в системах третьего поколения эта идея становится основной. Поэтому онтологии стали постепенно рассматриваться в них в качестве семантической концепции и структуры для представления всех компонент знаний. В этой концепции онтология должна объединять в себе знания о предметной области рекомендаций и о персональном многомерном профиле пользователя, поддерживать процессы выработки кросс-доменных рекомендаций (например, имея информацию об интересах пользователя в области музыки, формировать рекомендации в области кинофильмов или художественной литературы). Контекстная зависимость профиля интересов пользователя в онтологической концепции рекомендующих систем также должна быть представлена в онтологической модели знаний.

Динамика формирования этой концепции, предложенные подходы к представлению в ней персонального профиля пользователя и ее современное состояние рассматриваются в последующем материале работы.

2 Профиль пользователя

Построение персонального профиля пользователя в течение последнего десятилетия является одной из самых «горячих» тем исследований и разработок в области рекомендующих

³ Естественно, что провести *точные* границы между рекомендующими системами разных поколений невозможно.

систем, и она остается такой для рекомендующих систем третьего поколения. Базовая идея онтологической модели формализации профиля пользователя состоит в том, что понятие *интерес пользователя* рассматривается как экземпляр некоторой категории в иерархии категорий (понятий) онтологии, а *профиль пользователя* – как некоторая структура, заданная на множестве интересов-категорий онтологии, возможно, структурированных иерархически. Поэтому *построение онтологии* и *выделение в ней профиля пользователя* в настоящее время являются двумя тесно связанными задачами разработки рекомендующих систем. Обычно эта связь обусловлена тем, что и для построения онтологии предметной области рекомендаций, и для построения профиля пользователя используются одни и те же источники данных.

Если онтология предметной области построена, то поиск персонального профиля пользователя сводится к поиску множества его интересов, выраженных в терминах понятий этой онтологии или их примеров. Это может выполняться различными алгоритмическими средствами. Выбор зависит и от объема и состава исходной информации об интересах пользователя, и от метода, который далее предполагается использовать для принятия решений, и от других факторов.

Методы построения профиля пользователя варьируются от простого опроса самого пользователя, до полностью автоматического (машинного) обучения профиля путем обработки всей доступной информации, полученной из разных источников, в которой, так или иначе, проявляются «следы» поведения пользователя, отражающие его интересы, предпочтения, намерения, контекст принятия решения и прочее.

Далее приводится обзор результатов в области построения персонифицированного профиля пользователя в контексте онтологии предметной области рекомендаций для рекомендующих систем третьего поколения.

3 Онтологии и персонификация профиля пользователя: начальные модели

Идея построения персонального профиля пользователя с использованием категорий или понятий естественного языка была первоначально предложена для рекомендующих систем в области поиска документов в Интернет [10–13]. Эти методы принято называть методами на основе понятий (англ. *concept-based methods*). По существу, эти работы еще не использовали в полной мере онтологические модели профиля пользователя, однако они в последующем стимулировали их развитие. Общая идея этой группы методов состоит в том, чтобы найти в текстах документов, выдаваемых поисковой машиной в ответ на запросы пользователя, слова, которые могут рассматриваться как понятия. При этом принимается гипотеза, что эти слова–понятия отражают, с одной стороны, семантику запроса пользователя, а с другой – персональные интересы пользователя. Хотя эти подходы и напоминают традиционный подход на основе ключевых слов, однако, они не являются таковыми в полной мере. Первый шаг в этих подходах состоит в поиске часто встречающихся терминов с использованием традиционной меры, применяемой при поиске ключевых слов с помощью методов обнаружения часто встречающихся паттернов. Для того чтобы быть более уверенными в том, что термины, выделенные таким способом, не являются просто ключевыми словами, далее тем или иным способом выполняется их сравнение с понятиями некоторой базы данных понятий, выбранной разработчиком. Например, работа [12] формирует разделы интересов пользователя по множеству веб-документов его истории поиска и использует базу данных понятий *Open Directory Project (ODP)* [14]. Далее профиль пользователя представляется в виде множества найденных категорий с *весами*, пропорциональными частоте их появления в документах. Важно отметить, что далее делается шаг обобщения: каждая категория сопровождается списком ключевых слов, встречающихся в документах, которые могут рассматриваться в каче-

стве *примеров* этого понятия. Аналогичный подход используется в работе [10]. Преобразование найденных терминов в онтологию выполняется с помощью специального механизма классификации, построенного авторами. В работе [13] дополнительно к документам, которые ранее просматривались пользователем, используются также письма его электронной почты. В них находятся часто встречающиеся термины, и при этом так же, как и в ранее названных работах, полагается, что они отражают интересы пользователя в содержании веб-документов, которые он запрашивал. В качестве базы данных понятий в этой работе, как и в работе [12], используется структура категорий *ODP* [14].

Авторы работы [11] указывают в качестве общего недостатка работ [12–13] *статический* характер профиля пользователя, который строится с использованием практически не обновляемой базы данных *ODP* [14]. Новым в работе [11] является *персонификация каждого запроса* пользователя. И это очень важно для поиска веб-документов, так же, как и в других задачах поиска *персонифицированных* ответов на запросы. Для достижения этой цели авторы анализируют преамбулу (служебную информацию) каждого документа, возвращаемого на запрос пользователя поисковой машиной, называемую *Web-snippet*. Информация преамбулы включает в себя название документа, аннотацию его содержания, а также URL-адрес веб-документа, возвращаемого поисковой машиной в ответ на запрос. Далее авторы полагают, что те ключевые слова, которые часто встречаются в служебной информации возвращаемых документов, являются искомыми понятиями, которые отражают интересы пользователя. Для отбора таких понятий авторы, как и большинство авторов других работ, используют пороговое значение меры, которая обычно используется в алгоритмах фильтрации (отбора) часто встречающихся паттернов. Эта мера вычисляется как произведение числа слов в понятии и частоты встречаемости понятия (число преамбул, в которых встречается это понятие, деленное на общее число возвращенных документов или проанализированных преамбул). Дополнительным новым элементом работы [11] является использование как *позитивных*, так и *негативных* интересов. По каждому запросу авторы сначала извлекают множество понятий на основе служебной информации, затем выполняется кластеризация найденных понятий по некоторой мере сходства и построение на этой основе персонифицированного профиля пользователя. Авторы рассматривают несколько таких стратегий, предложенных в различных их работах.

Достоинство проанализированной группы работ состоит в том, что они отказываются от простого использования ключевых слов, объединяя их в понятия. Они привлекают дополнительную информацию (например, базу данных *ODP* или аналогичную по содержанию другую базу данных понятий) для поиска понятий, представляющих персонифицированный профиль пользователя в форме множества найденных понятий с приписанными им весами. Другое достоинство этих работ – это использование сжатой информации о веб-документе, содержащейся в его преамбуле.

Однако эти работы делают лишь первый шаг от традиционных стратегий, которые используют в качестве атрибутов ключевые слова, к стратегиям на основе онтологий. Они используют только частотные характеристики для взвешивания интересов в профиле пользователя, но не анализируют зависимости между ними и не устанавливают отношений на множестве извлеченных терминов, которые рассматриваются далее как понятия. Бинарная постановка задачи, в которой не используется понятие рейтинга документа с точки зрения пользователя, обычно обеспечивающего обратную связь от него, не позволяет корректировать значимость различных интересов в профиле пользователя, и это сильно ограничивает возможности описанной группы методов.

Другие работы периода 2004–2008 гг., аналогичные по используемой идее⁴, предлагающие «почти» онтологические модели персонифицированного профиля пользователя на основе понятий, как правило, посвящены разработке конкретных приложений, а потому сильно ориентированы на их особенности. Примерами таких работ являются [15–17] и ряд других.

В [17] используется классификация веб-страниц. Основное внимание в ней уделяется персонификации профиля пользователя, представленного терминами *выбранной заранее* онтологии, в данном случае, в терминах *иерархии ODP* [14]. В качестве исходной информации для моделирования профиля пользователя используются веб-страницы (документы), классифицированные вручную по типу интереса для пользователя таким образом, чтобы каждый построенный класс отражал некоторый его конкретный интерес. Множество документов выборки, поставленной в соответствие одному классу, авторы называют «супердокументом» класса, который далее используется в качестве обучающей выборки этого класса. Каждому классу ставится в соответствие то, что авторы называют понятием (по существу – это некоторый интерес пользователя). Процесс обучения состоит в том, что каждому понятию-классу далее ставится в соответствие вектор терминов, сформированных с помощью ключевых слов и словаря понятий *ODP* [14] с весами, которые вычисляются по «супердокументу» класса. В итоге каждое понятие-класс представляется вектором терминов из словаря *ODP* с весами.

Классификация веб-страниц проводится на основе вычисления меры сходства вектора весов терминов, созданного для веб-страницы, с аналогичным вектором каждого понятия-класса, причем в качестве меры сходства выбрано значение косинуса угла между этими векторами. Для поиска лучших совпадений используется также метод *K-ближайших соседей*.

Работа [16] предлагает онтологическую модель профиля пользователя в онлайн-системе рекомендаций научно-исследовательских работ. Эти работы классифицируются с использованием онтологических классов и механизма коллаборативной фильтрации (англ. *collaborative filtering*). Профиль интересов пользователя представляется в понятиях онтологии научной области исследовательских работ, разработанной в рамках инициативы АКТ [18]. Связи, присутствующие в этой онтологии, используются для того, чтобы выводить другие понятия-темы, которые могут представлять интерес для пользователя, в том числе и те, которые еще не были им просмотрены.

В [15] предлагается модель профиля пользователя, охватывающая несколько предметных областей, представленных единой моделью онтологии. Эта работа имеет целью преодолеть ограниченность работ, в которых профиль пользователя представлен ключевыми словами, не согласованными между собой семантически, т.к. последнее может приводить к противоречивым моделям. Кроме того, авторы встраивают в модель онтологии контекст, сопровождающий процесс выработки рекомендаций. В работе предлагается рекомендующая система *COReS (Context-aware Ontology-Based Recommender System)*, использующая онтологию. Эта система способна работать с несколькими предметными областями и может привлекать контекст для получения более точных персонифицированных рекомендаций. «Еще одно новшество *COReS* – разбиение профиля пользователя на части в соответствии с различными областями рекомендаций, которое формируется в контексте решаемой задачи», – отмечают авторы. К сожалению, в работе рассматриваются аспекты мульти-доменного профиля пользователя, но не кросс-доменного профиля. Последнее было бы значительно интереснее, поскольку является гораздо более востребованным на практике.

Далее рассматриваются работы такого же направления, которые ориентированы уже на полноценное использование онтологий предметной области рекомендаций и ее категорий для представления персональных пользовательских интересов. Эти работы активно исполь-

⁴ В этот период появились также работы, которые использовали уже полномасштабные онтологические модели для представления профиля пользователя.

зуют также различные существующие общедоступные онтологии и базы знаний для того, чтобы *автоматизировать* процессов построения онтологических профилей пользователей.

4 Онтологии и персонификация профиля пользователя: современные модели

Модели, которые в полной мере используют возможности онтологии как структуры для представления знаний о предметной области рекомендаций для представления профиля пользователя в терминах ее понятий, активно развиваются в течение последних примерно 6 лет. Важно отметить, что ключевым фактором во всех этих исследованиях является состав и свойства исходной информации, которая доступна для построения онтологии.

Другой общей чертой этих работ является поиск путей *автоматизации* процессов, связанных с разработкой персонального профиля пользователя в контексте предметной онтологии. Для этих целей используются различные словари понятий, стандартные онтологии повторного использования, викификация⁵ веб данных [19], развиваемые в настоящее время параллельно со всем хорошо известным вебом документов, и другие источники знаний, стандартные программные средства и программные платформы.

В работе [20] представлена частная технология автоматизированного построения формальной модели персонального профиля пользователя в терминах онтологии, характеризующей его интересы и намерения. Она ориентирована, главным образом, на пользователей, работающих в социальной сети *Twitter*, например, для последующей рекомендации друзей в этой сети или для рекомендации им товаров и услуг с помощью более целенаправленной рекламы.

В качестве входной информации в разработанной модели используются URL-адреса, извлеченные из постов пользователя в *Twitter* (из «твитов»). Верхний уровень онтологии, используемой авторами, который отвечает категоризации веб-сайтов, формируется вручную с использованием понятий, извлеченных из дополнительных источников. Примерами таких источников являются базы коллективных знаний *OpenDNS* [21], таксономии рекламных объявлений и онтологии *DBpedia* [22]. *OpenDNS* накапливает концептуализацию *тэгов* и реализуется как особый облачный сервис, который позволяет членам сообщества пополнять множество пар *тэг – категория* веб-сайта и относить веб-сайты к некоторым категориям, например, *Музыка*, *Кинофильм* и т.п. Категории, к которым может относиться веб-сайт, должны быть одобрены сообществом, прежде чем они попадут в эту базу данных. *DBpedia* извлекает категоризацию из дампа категорий Википедии и преобразует их представление в *RDF*-формат. *OpenDNS* и *DBpedia* покрывают целый ряд предметных областей, т.е. они позволяют строить мульти-доменные онтологии интересов пользователей.

Для построения модели профиля пользователя авторы представляют онтологию на языке *OWL* или *RDF/XML*. Для хранения твитов, содержащих URL, и сформированных профилей пользователя в формате *RDF* авторы используют *NoSQL*-базы данных *Cassandra* [23] и *Virtuoso* [24], соответственно. Заметим, что система управления базами данных *Cassandra* представляет собой распределённую *NoSQL*-систему [25], рассчитанную на создание масштабируемых и надёжных хранилищ огромных массивов данных, представленных в виде хэша. *Virtuoso Universal Server* является системой с открытым исходным кодом, которая сочетает в себе возможности реляционной базы данных, объектно-реляционной базы данных и вирту-

⁵ Так принято называть технологию автоматизации построения онтологий, в которой используется дамп категорий Википедии в качестве словаря понятий с возможностью получения статей, представляющих семантику этого понятия. Технология предложена И. Виттенем (I. Witten) из Университета Вайкато, Новая Зеландия, в 2007 году.

альной базы данных. Она хранит данные в формате *RDF* и *XML*. Она может выполнять функции сервера веб-приложений и файлового сервера.

Следует отметить, что использование *NoSQL*-баз данных позволяет улучшить масштабируемость приложений и обеспечивает возможность работы с большими данными, в частности, поступающими в реальном времени, что сегодня является обязательным требованием к системам, работающим с социальными сетями. Заметим, что использование *NoSQL*-баз данных отражает современную тенденцию рассматривать рекомендующие системы как системы, функционирующие на основе анализа больших данных.

Рассмотрим процедуру формирования профиля пользователя, предложенную в работе [20]. На первом этапе URL-адреса извлекаются из твитов и помещаются в базу данных *Cassandra*. Авторы полагают, что *все* URL, размещенные пользователями в твитах, представляют его интересы. С этим нельзя согласиться безоговорочно, поскольку пользователи могут делиться не только тем, что их интересует, но и чем-то, что вызывает у них негативные эмоции, или же размещать в твите какую-либо информацию по просьбе своих друзей или в рекламных целях. При этом понятие «негативного» интереса или интересов с присвоенными значениями рейтингов авторами в работе вообще не рассматривается. Далее, анализ сопутствующих текстов, а также текста, содержащегося в твите, не выполняется. Вообще говоря, это может приводить к потере существенной информации.

Твиты, помещенные в базу данных *Cassandra*, постоянно анализируются в пакетном режиме и обрабатываются с целью извлечения интересов и намерений пользователя (см. далее по тексту, как это выполняется), которые затем добавляются в его онтологический профиль, хранимый в базе данных *Virtuoso* в виде *RDF*-триплетов. Для поиска дополнительных интересов в существующей онтологии профиля пользователя используется механизм рассуждений *Pellet* [26].

Алгоритм извлечения интересов и намерений пользователя состоит в следующем. Сначала все URL-адреса пользователя разделяются на *интересные* и *неинтересные*. К «неинтересным» авторы относят URL-адреса поисковых машин (*Google*, *Yahoo* и т.п.), сервисы сокращения ссылок (*Goo*, *Su* и т.п.), URL социальных сетей (*Facebook*, *Twitter* и т.п.). Далее анализируются только «интересные» URL-адреса. Отметим, что такую фильтрацию (очистку, редукцию исходных данных) нельзя отнести к бесспорным преимуществам алгоритма, поскольку таким способом можно удалить большой объем полезной информации о пользователе.

Далее, для каждого URL-адреса из баз коллективных знаний *OpenDNS* и *DBpedia* извлекаются категории и понятия, которые добавляются к онтологии профиля пользователя, вместе с отношениями *hasIntent* или *hasInterest*. Дополнительно извлекаются категории, связанные отношением *sameAs* с категориями URL, для которых в онтологии пользователя создается отношение *categorySameAs*.

Фрагмент онтологии, разработанной в [20], в части, касающейся представления профиля пользователя, изображен на рисунке 1. В нем понятие *Персона* – класс, который идентифицирует пользователя; URL – класс, включающий в качестве примеров URL-адреса, размещенные в твитах; *Интерес* – класс, который представляет понятия, связанные с понятием URL; *Намерение* – класс, который представляет понятия, связанные с понятием URL, выражающие покупательские намерения пользователя; *Неизвестно* – класс, содержащий URL-адреса, которые не существуют в базах данных коллективных знаний (*OpenDNS*, *DBpedia*); *НеизвестнаяКатегория* – класс категорий, которые не существуют в БД коллективных знаний (*OpenDNS*, *DBpedia*).

Авторы определяют несколько подклассов классов *Интерес* и *Намерение*, которые соответствуют различным предметным областям, например *Спорт*, *Здоровье*, *Религия* и т.п.

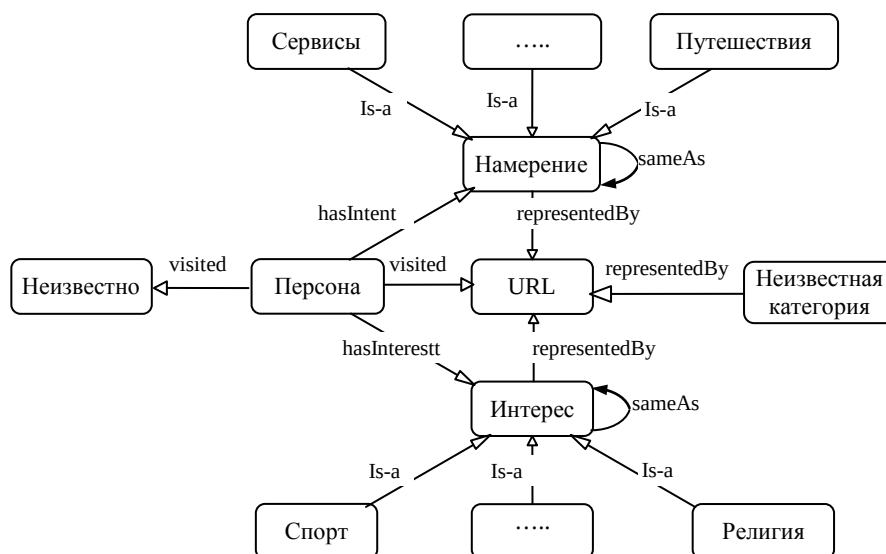


Рисунок 1 — Онтология, разработанная в [20] для представления профиля пользователя (для её графического представления использовано словесное описание, приведённое в указанной работе)

Основные отношения предложенной онтологии - это *hasInterest* и *hasIntent*. Отметим, что авторы разделяют все предпочтения пользователя на *Интересы* и *Намерения*. Намерения представляют сущности, которые существуют определённый период времени, в то время как *Интересы* полагаются постоянными для *Персоны*.

Отметим еще одну особенность данной работы. *OpenDNS* позволяет выполнять категоризацию только для доменов верхнего уровня и предоставляет информацию только об общей тематике сайта. Иначе говоря, если, например, пользователь разместил в твите ссылку на страницу своей любимой команды на сайте, посвященном футболу, то запрос к *OpenDNS* по данной ссылке, вернёт категории всего сайта, например, *Спорт*, *Футбол* и т.п. Но пользователь совсем необязательно интересуется футболом и спортом в целом. Возможно, его интересуется только его любимая команда или какой-то из её игроков.

Таким образом, методика, предложенная в [20], позволят извлечь только достаточно общие интересы пользователя с невысокой, по нашему мнению, точностью. Тем не менее, подход может быть использован для первичного извлечения областей интересов пользователя, которые затем могут быть уточнены, например, путем анализа содержания страниц, размещённых по URL-адресу из твита. Описанная модель профиля пользователя также не предусматривает учёт контекста, в котором пользователь имеет тот или иной интерес. В планы авторов входит дальнейшее развитие технологии, например, анализ текстов, сопутствующих URL-адресу из твита, для получения лучших результатов.

Несмотря на недостатки, перечисленные выше, данная работа хорошо иллюстрирует, как можно *обогащать информацию о пользователе*, извлеченную из публичных источников, с помощью онтологий и фольксономий, разработанных сообществом, например, *OpenDNS* и *DBpedia*. Данная работа является еще одним подтверждением общего мнения о том, что использование таких дополнительных источников знаний является неотъемлемой частью рекомендующих систем третьего поколения, а также может быть использована для автоматизации построения онтологий.

Работа [27] предлагает словарь для описания профиля пользователя, представленный на языке *RDF*, что дает возможность его применения сразу в нескольких классах приложений, ориентированных на веб-поиск. Идея авторов состоит в том, чтобы использовать такой профиль пользователя в качестве ограничения совместно с лексикой запроса клиента («уточне-

ние запроса на основе интересов» в объяснении авторов). Авторы мотивируют такой подход тем, что обычно запрос бывает слишком расплывчатым, а потому без дополнительного уточнения он может приводить к генерации большого количества ненужных документов. В работе предлагается расширенное определение понятия *интерес пользователя*, которое включает в себя не только само понятие *интерес* («тема, по которой агент хочет что-либо узнать, чему-либо обучиться или быть в нее вовлечен», в соответствии с [27]), но и добавляет к нему некоторые атрибуты, задающие контекст, например, время. В реальности это расширение понятия *интерес* предполагает создание стандартного словаря понятий, представляющих интересы, связанные с фиксированным набором атрибутов, а также с некоторыми контекстными атрибутами. Но, во-первых, в настоящее время не представляется возможным разработать словарь, охватывающий все виды интересов, во-вторых, работа игнорирует такую важную компоненту спецификации профиля пользователя, как структура интересов пользователей. В сущности, результаты этой работы применимы, главным образом, к веб-поиску.

Статья [28] мотивирует использование онтологии необходимостью восполнить нехватку семантической информации при построении персонализированного профиля пользователя, который может динамически меняться с течением времени. Как известно, рекомендации, вырабатываемые на основе коллаборативной фильтрации, основаны на *семантическом* сходстве целевого пользователя с другими пользователями и на мнениях последних (они представляются значениями рейтингов) относительно рекомендуемого экземпляра товара/услуги. Основываясь на онтологии интересов пользователя, работа предлагает модель профиля пользователя и алгоритм, выполняющий ее автоматическое обновление по мере получения новых данных. Иерархическая структуризация интересов формулируется только в качестве цели будущих исследований. Поэтому результаты этой работы не могут напрямую использоваться в рекомендующих системах третьего поколения.

Резюмируя содержание работ, рассмотренных в данном разделе, а также других, которые здесь не упоминаются, можно сделать следующее заключение. Эти работы, хотя и подготовили необходимый теоретический и технологический базис для широкого использования онтологических моделей в качестве единой структуры для представления различных компонент знаний рекомендующих систем, тем не менее, нуждаются в дополнительных исследованиях. Особенно убедительным этот тезис представляется применительно к специфическим видам информации, которую необходимо использовать в рекомендующих системах третьего поколения. Некоторые из них рассматриваются в последующих разделах.

5 Категоризация тэгов для построения персонифицированного профиля пользователя

Тэгами называют семантически осмысленные метки, которые любой пользователь может присваивать некоторым Интернет-ресурсам, например, документам, видео и аудио файлам, информации, представленной в социальных сетях и т.п. Иногда они играют роль, подобную ключевым словам, иногда они выражают эмоциональное отношение пользователя к соответствующему ресурсу, иногда используются для улучшения поиска. В любом случае тэги содержат весьма полезную информацию для идентификации свойств Интернет-ресурсов и отношения к ним пользователя. Роль тэгов многообразна, однако в данной работе они рассматриваются с позиций их полезности для извлечения информации об интересах пользователя.

Работы, посвященные использованию тэгов для построения онтологии предметной области рекомендаций и профиля пользователя, появились недавно. К настоящему времени эта проблема все еще разработана недостаточно, хотя она исследуется уже около 8 лет. Рассмотрим некоторые работы, которые посвящены решению указанной проблемы.

В работе [29] предлагается методология моделирования онтологического профиля пользователя с использованием тэгов Интернет-ресурсов и сервисов, к которым пользователь проявляет интерес. Использование информации о тэгах позволяет вовлекать в построение онтологического профиля пользователя дополнительные знания. Заметим, что информация о тэгах часто носит персонифицированный характер и потому важна для построения персонального профиля пользователя.

В качестве примера в статье [29] рассматриваются сервисы *Flickr* и *Delicious*. Для сведения, *Flickr* представляет собой сервис, предназначенный для хранения и дальнейшего использования цифровых фотографий и видеороликов. *Delicious* [30] – это веб-сайт, предоставляющий зарегистрированным пользователям бесплатную услугу по хранению и публикации закладок на страницы Интернет. Поскольку информация об интересах пользователя распределена по очень многим источникам в Интернет, то дополнительной мотивацией работы [29] является отработка технологии построения онтологии по информации из различных источников. В частности, авторы говорят о технологии консолидации тэгов, распределенных по многим источникам, в рамках онтологии.

Информация о профиле пользователя, сформированная на основе множества тэгов, удобна для реализации механизма сравнения (англ. *matching*) ее с семантической аннотацией рекомендуемого продукта/услуги, представленной в терминах *той же онтологии*, построенной для того же множества тэгов. Такое семантическое сходство рассматривается в работе как основа персонификации профиля пользователя и, соответственно, персонификации рекомендаций.

Авторы предлагают *отображать* теги на понятия онтологии профиля пользователя с помощью словарей *WordNet* и категорий Википедии. Напомним, что *WordNet* – это электронный тезаурус/семантическая сеть для английского языка, разработанный в университете Принстон (США) и выпущенный вместе с сопутствующим программным обеспечением под свободной лицензией [31]. Отображение тэгов на понятия профиля пользователя реализуется процедурой из трех шагов: (а) фильтрация тэгов, (б) получение семантической информации о тэгах из сети Интернет и (в) категоризация полученных понятий согласно классам понятий существующей онтологии.

Опишем несколько подробнее процесс извлечения и обработки тэгов пользователя, а также процесс их отображения на понятия онтологии с использованием семантической информации, полученной из сети Интернет.

На первом шаге, на котором выполняется фильтрация тэгов, сначала выбирается некоторый общий английский словарь терминов, например, *WordNet* [31-32], Википедия и/или Google и выполняется предобработка тэгов с использованием их морфологических и семантических трансформаций с последующим связыванием результирующего множества тэгов с терминами выбранного словаря. Лексический фильтр отбрасывает слишком короткие тэги (длиной в один символ), слишком длинные (более 25 символов), а также артикли, предлоги и т.п. Выполняется также преобразование специальных символов (например, символов *à, á, â* в символ *a*). Редко встречающиеся теги также удаляются с помощью пороговой фильтрации, хотя последнее и не является однозначно полезным. Оставшееся множество тэгов передается в *WordNet*, после чего выполняется их сравнение (англ. *matching*) с терминами словаря (по точному их совпадению). Далее «успешные тэги» (те, для которых нашлась пара в выбранном словаре) добавляются в результирующий набор тэгов. Для уточнения правильности написания тэгов используется также известный механизм Google «возможно, вы имели в виду». С помощью алгоритма, разработанного авторами [29], составные существительные типа *unitedkingdomsouthampton, san_francisco* и т.п. разбиваются на части. Полученные теги снова проверяются с помощью словаря *WordNet*. Теги, которые не были идентифицированы с по-

мощью WordNet (имена собственные, сокращения или сленг), далее анализируются на предмет поиска «пары» с помощью дампа категорий Википедии. Часто используемые аббревиатуры заменяются общепринятыми терминами, например, *нус* заменяется на *New York City*. Используется также морфологический анализ, который позволяет удалить морфологически одинаковые теги, например, *blog*, *blogs*, *blogging*. Затем, с помощью алгоритма, снова использующего WordNet, удаляются синонимы.

На втором шаге из Википедии извлекается семантическая информация по отфильтрованным тегам. Эта информация включает в себя наименование понятия, соответствующего тегу, принятое в Википедии, и его категорию.

Наконец, на третьем шаге выполняется отображение понятий из Википедии, которые соответствуют тегам, на понятия разработанной ранее онтологии. Тем самым выполняется расширение ранее построенной онтологии. Для этого выполняется морфологическое сравнение наименований понятий из Википедии и наименований его категорий с классами разработанной ранее онтологии. Понятие, сформированное на основе тэгов, добавляется в онтологию как экземпляр класса с наиболее похожим именем (категории, извлеченные из Википедии, добавляются к экземпляру как RDFS-метки).

Для формирования профиля пользователя авторы [29] используют представление предпочтений пользователя в следующей форме [33–34]⁶:

$$U_m = [u_{m,1}, \dots, u_{m,i}, \dots, u_{m,K}]^T, \text{ где } u_{m,i} \in [0,1], c_m \in (c_1, \dots, c_M).$$

Бинарные компоненты $u_{m,i}$ вектора U_m выражают меру интереса пользователя $c_m \in C$ к понятию O_i онтологии предметной области O или отдельных экземпляров этого понятия, т.е. $O_i \in O$; K – общее число понятий в онтологии. Рекомендуемые продукты/услуги (*items*) $s_n \in S$ описываются аналогичным вектором:

$$V_n = [v_{n,1}, v_{n,2}, \dots, v_{n,i}, \dots, v_{n,K}]^T, \text{ где } v_{n,i} \in [0,1], s_n \in (s_1, \dots, s_N),$$

в котором бинарные компоненты $v_{n,i}$ вектора V_n выражают вес понятия $O_i \in O$ для товара/услуги $s_n \in (s_1, \dots, s_N)$, K – общее число понятий в онтологии.

Далее, в этом подходе вместо матрицы рейтингов $R = \{r_{i,j}\}_{i=1, j=1}^{M,N}$, введенной в разделе 1, используется матрица $D = \{d_{m,n}\}_{i=1, j=1}^{M,N}$, в которой элемент $d_{m,n} \in D$ равен косинусу угла между векторами $U_m = [u_{m,1}, \dots, u_{m,i}, \dots, u_{m,K}]^T$ и $V_n = [v_{n,1}, v_{n,2}, \dots, v_{n,i}, \dots, v_{n,K}]^T$. Эта величина используется в качестве меры сходства вектора интересов U_m пользователя $c_m \in C$ и вектора V_n , описывающего свойства товара/услуги $s_n \in S$. Напомним, что значение косинуса между парой векторов равно их нормированному скалярному произведению:

$$d_{m,n} = (U_m, V_n) / (\|U_m\| \times \|V_n\|).$$

При выработке рекомендации для пользователя $c_m \in C$ предпочтение отдается товару/услуге $s_n \in S$ с наименьшим значением меры $d_{m,n}$.

Очевидный недостаток описанного подхода к построению персонифицированного профиля пользователя с использованием информации о тэгах состоит в том, что размерность бинарных векторов $U_m \in U$ и $V_n \in V$ всегда равна числу понятий в онтологии предметной области, а оно может исчисляться тысячами, и их число может увеличиваться в процессе работы системы. Матрица «пользователь–продукт» будет очень большой размерности. Работа

⁶ Здесь и далее используются обозначения, частично введенные ранее в разделе 1.

с такой моделью потребует огромных вычислительных ресурсов. Однако, поскольку здесь вычисления ведутся с бинарными переменными, то не возникнет эффекта накопления ошибок, свойственного процедурам обработки больших данных с вещественными атрибутами.

Среди преимуществ этого подхода авторы выделяют инвариантность технологии его реализации к предметной области и возможность ее применения для любых продуктов и услуг.

Описанный подход и реализующая его технологии проверены с помощью системы рекомендации новостей *News@hand*, разработанной авторами реферируемой статьи. В этой системе поддерживается также автоматическое получение новостных статей из сети, автоматическое семантическое аннотирование этих статей с использованием средств обработки естественных языков [35] и средства *Lucene* [36].

В целом, работа достаточно убедительно демонстрирует интересную и практически реализуемую перспективу комбинированного использования *фольксономии* (*folksonomy*)⁷, как иначе принято называть словари и базы данных тэгов и знаний, которые могут быть получены из различных источников с использованием иерархической категоризации понятий (онтологической информации), содержащейся в словарях, в частности, в дампе категорий Википедии.

Интересный подход использования тэгов для построения онтологий предложен в работе [37]. Он, хотя и не имеет прямого отношения к построению онтологического профиля пользователя в рекомендующих системах, тем не менее, предлагает новый подход к построению онтологий с использованием тэгов, который представляется достаточно продуктивным и для рекомендующих систем, в частности, для построения онтологического профиля пользователя на основе тэгов. С методологической точки зрения, все основные операции этого алгоритма направлены на максимальное повторное использование имеющегося опыта разработок, доступного авторам.

На первом шаге построения онтологии предложенный алгоритм генерирует некоторую базовую (черновую) онтологию, которая содержит в себе множество тэгов данных, каждому из которых поставлен в соответствие термин словаря (понятие). Они структурированы некоторым множеством отношений. Для автоматической генерации базовой онтологии используется распределенная структура фольксономии, представленная в ресурсе *OpenDNS* [21]. Этот ресурс, созданный волонтерами с помощью накопления и объединения собственных фольксономий (результатов по аннотации тэгов), представляет собой достаточно простую (в первоисточнике англ. - *lightweight*) концептуальную структуру, которая связывает пользовательские термины с веб-ресурсами тэгов. Эта процедура позволяет существенно ускорить концептуализацию и классификацию тэгов данных и, тем самым, ускорить построение онтологии. Понятия и отношения базовой онтологии, построенной таким способом, далее корректируются экспертом по предметным знаниям. В этом процессе какие-то понятия и отношения из базовой онтологии удаляются, какие-то могут добавляться.

Второй шаг описываемого подхода также опирается на имеющийся опыт построения онтологий и пытается максимально его использовать. На этом шаге авторы обращаются к облачному сервису *веба данных* (англ. *Linked Data*) [19] для того, чтобы использовать внешние связи к классам понятий этого ресурса и к отношениям. Эта работа выполняется предметными экспертами. Затем выполняется совместный просмотр аннотаций и отношений веб сайтов *веба данных*, которые были рекомендованы экспертами, и сайтов фольксономии, которые были использованы при построении базовой онтологии.

⁷ Фольксономии обычно содержат термины, которые отсутствуют в стандартных глоссариях.

Исследования в области методов концептуализации тэгов и построения онтологий на этой основе ведутся в настоящее время достаточно активно (см., например, [38]). Расширенный обзор по состоянию исследований в этой области можно найти в работе [39].

6 Построение профиля пользователя в контексте социальных сетей: пример

Широкий и важный класс задач выработки рекомендаций охватывают приложения, функционирующие в контексте социальных сетей. Одним из множества имеющихся примеров такого класса приложений является рекомендующая система поиска экспертов для ответа на вопросы, рассмотренная в работе [40], когда эксперты образуют некоторую сеть. В этом приложении эксперты выступают в роли «пользователей», которым адресуется («рекомендуется») некоторый запрос (продукт, услуга – в терминологии традиционной постановки задачи), и задача состоит в том, чтобы адресовать этот запрос экспертам, которые наиболее компетентны по тематике запроса. Компетентность эксперта здесь играет роль, аналогичную роли персонального профиля пользователя. Другой пример – задача подбора специалистов на некоторые должности при заданных требованиях к профессиональным знаниям для каждой из должностей. Главной задачей здесь, по-прежнему, является задача построения *персонального профиля* пользователя, которым является *эксперт*, в рассматриваемом приложении. Особенность этой задачи состоит в том, что при поиске рекомендаций используется «сетевой контекст», задаваемый топологией и числовыми характеристиками социальной сети экспертов. Этот сетевой контекст задает меры сходства различных пользователей (экспертов – в задаче [40]) и связи между ними, а значит, позволяет использовать методы коллаборативной фильтрации при формировании рекомендаций.

Рассмотрим пример сетевого контекста для приложения из [40]. В социальной сети экспертов каждый из них имеет некоторую область экспертизы, описываемую множеством ключевых слов. Сходство областей экспертизы любой пары экспертов (узлов социальной сети, в общем случае), которое задается формально некоторой моделью, определяет то, что названо выше «сетевым контекстом». Заметим, что постановка задачи, рассматриваемая в работе [40], обобщается на широкий класс приложений, решающих задачи формирования рекомендаций в социальных сетях, например, поиск друзей в сети, формирование групп по некоторому множеству интересов и прочие.

Опишем кратко, каким образом сетевой контекст (сходство на множестве пар экспертов) описывается и вычисляется формально в задаче, рассматриваемой в [40]. Эта модель представляет определенный интерес, поскольку является одним из примеров меры *семантического сходства пользователей*, столь важной для применения методов коллаборативной фильтрации.

Сеть экспертов (в общем случае – социальная сеть) описывается графом, в котором каждому эксперту сети ставится в соответствие узел графа, а в качестве атрибутов узла рассматриваются ключевые слова, задающие область экспертизы (область знаний) эксперта, представленного узлом. Дуги между вершинами этого графа задают отношение цитирования одного эксперта другим, так что это отношение не является симметричным. Эти данные используются для вычисления матрицы сходства на множестве узлов графа. Авторы используют совместно несколько мер, которые затем некоторым образом комбинируются в итоговой мере сходства. Одна из этих мер называется *структурной* мерой сходства. Она задается либо числом общих узлов, связанных описанным выше отношением цитирования с каждым элементом пары (в общем случае – числом общих соседей рассматриваемой пары узлов графа), либо числом других авторов, которые цитируют оба элемента пары. В общем случае вид меры сходства зависит от приложения.

Что касается модели персонального профиля пользователя, то работа [40] использует для этих целей модель, которая для каждого пользователя является экземпляром некоторого класса предметной онтологии. Другими словами, область экспертизы каждого эксперта представляется множеством терминов. Для этих терминов с использованием английского словаря терминов *WorldNet* отыскиваются наиболее близкие понятия предметной онтологии, причем близость оценивается с помощью весов релевантности (к сожалению, авторы не сообщают, каким именно образом эти веса вычисляются). Профиль пользователя определяется тем классом понятий предметной онтологии, который наиболее близок к представлению профиля пользователя.

По-видимому, на момент написания статьи авторы [40] еще не закончили разработку модели онтологического профиля пользователя, поскольку в работе представлена только модель концептуального уровня. В целом, предложенный подход к построению онтологического профиля пользователя проработан достаточно слабо, поэтому о новых решениях в этой части авторы не сообщают.

7 Онтологические модели профиля пользователя для множества источников данных

Необходимость формирования профиля пользователя по данным, распределенным по множеству источников, возникает практически в каждом случае, если не считать те случаи, которые задаются стандартными наборами данных, предназначенными для проверки и сравнения различных моделей построения персонального профиля пользователя (англ. *benchmarks*). По этой причине необходимо уметь работать с распределенными источниками данных, в которых проявляются «следы» пользователя. Обычно для решения таких задач привлекается опыт исследователей в области слияния данных из различных источников (англ. *data fusion*). Этот опыт достаточно большой, он дает много примеров использования онтологий, и поэтому здесь об аналогичных задачах применительно к построению онтологического профиля пользователя говорится только с той целью, чтобы обозначить существование такой проблемы в области рекомендующих систем.

Одна из таких моделей предлагается в работе [41]. В ней предполагается, что для построения профиля пользователя привлекается информация о его работе в социальных сетях, например, в *Twitter*, *Facebook*, *LinkedIn*, а также используются данные с его домашней страницы.

Авторы рассматривают разные источники данных как результаты, измеряемые разными сенсорами, а саму задачу объединения информации – как задачу слияния сенсорной информации (англ. *multi-sensor information fusion*). Рассматриваются три стратегии слияния информации, которые традиционно приняты в теории слияния данных и информации, а именно [42]:

- 1) *Слияние информации на уровне данных*, когда данные из разных источников собираются в общей базе и обрабатываются централизованно. Обычно это возможно при небольших размерностях данных о пользователе. Здесь, однако, затруднительно обеспечить конфиденциальность данных, которые в разных источниках могут иметь разных владельцев.
- 2) *Слияние информации на уровне признаков, атрибутов данных* (англ. *feature level fusion*), когда в каждом источнике данных формируется потенциальное множество интересов, они объединяются вместе в общей онтологии, и далее производится их фильтрация по важности и персонификация. Этот вариант слияния данных несколько лучше предыдущего, однако, страдает многими аналогичными недостатками.

- 3) *Слияние информации на уровне решений.* В этом случае информация об интересах и предпочтениях пользователя, представленная в терминах онтологий, формируется в каждом источнике независимо друг от друга. Затем понятия и отношения локальных онтологий объединяются с помощью метаонтологии, для которой понятия локальных онтологий являются подклассами ее классов. Этот подход, в общем случае, представляется наиболее перспективным. Он позволяет успешно преодолевать различные проблемы, связанные и с большим масштабом объединенных данных, и с конфиденциальностью. Кроме того, в каждом источнике данных может использоваться своя стратегия разработки онтологии интересов пользователя. Авторы работы [41] принимают именно эту стратегию.

Интересы пользователей описываются в ней в терминах понятий и отношений онтологии. Онтология используется для установления различных взаимосвязей между интересами пользователя, представленными в нескольких источниках данных с различными степенями детализации. Онтология верхнего уровня, объединяющая несколько источников данных, используется также для вывода неявных интересов пользователей путем рассуждений с использованием иерархии интересов в онтологии. В целом, в работе кратко описываются потенциальные направления и перспективы богатой семантически-ориентированной модели профиля пользователя, но не более того, так как основные идеи демонстрируются в работе на слишком упрощенных примерах. Пока не ясно, насколько адекватно предлагаемые идеи могут быть использованы в приложениях реальных масштабов.

Заключение

В рекомендующих системах третьего поколения онтологии рассматриваются в качестве базовой семантической модели и структуры для представления персонального профиля пользователя. В ней модель знаний о профиле пользователя представляется в терминах классов понятий, которые описывают онтологию предметной области рекомендаций. Сформулируем основные выводы о роли онтологий в рекомендующих системах третьего поколения, в частности, о представлении персонального профиля пользователя, а также о некоторых особенностях такого представления.

- 1) Онтология позволяет представлять интересы пользователя в терминах семантически интерпретируемых категорий естественного языка, т.е. в терминах, понятных пользователю, в которых он о своих интересах обычно говорит. Например, интересы формулируются в таких терминах, как интерес к фильмам комедийного жанра или интерес пользователя к фильмам с участием конкретных актеров, интерес к новостям на конкретную тему и т.п. Семантически понятная интерпретация интересов пользователя позволяет получать обратную связь от пользователя как при построении онтологии и профиля пользователя, так и при оценке качества работы системы в целом. В онтологической модели профиль пользователя представляется как структурированное подмножество его интересов, которые представлены конкретными понятиями/категориями предметной области, или/и экземплярами некоторых классов понятий. Такая семантически ясная модель профиля пользователя может дать полезную информацию для объяснения рекомендаций. Особенно полезной эта информация является в том случае, когда интересы пользователя построены на основе поиска причинных связей интересов и оценок, которые он дает продуктам/услугам.
- 2) При построении профиля пользователя разработчики должны стремиться использовать всю доступную информацию, в которой, так или иначе, проявляются «следы» поведения пользователя, отражающие его позитивные или негативные интересы. В особенности полезной является информация, в которой представлены субъективные оценки пользовате-

ля, его отношение к тем или иным продуктам/услугам. Одним из важных источников такой семантически-интерпретируемой информации является множество тэгов. Ими пользователи метят данные. Часто пользователи отражают в тэгах информацию, которую невозможно или трудно описать в терминах свойств товаров/услуг. Поэтому иногда для оценки мнения пользователя тэги могут быть полезнее, чем любая другая информация. Они чаще являются *субъективными* оценками и потому дают важную информацию для персонализации рекомендаций.

К настоящему времени проблема использования тэгов для формирования знаний о профиле пользователя пока исследована недостаточно, хотя она, очевидно, и является проблемой первостепенной важности. Сейчас эта область исследований и разработок остается все еще на уровне отдельных удачных решений, но не на уровне хорошо отработанной технологии. Она требует серьезных дополнительных исследований, которые должны быть направлены, прежде всего, на создание *автоматизированных* технологий обработки тэгов для построения онтологического профиля пользователя. Некоторые примеры таких технологий, а также методы и средства их автоматизации были рассмотрены в данном обзоре.

- 3) Большой интерес представляют исследования, которые для построения онтологии и профиля пользователя используют не только ранее накопленную (статическую) информацию, но и динамически отслеживают новые интересы пользователя и трансформируют его профиль на основе новых данных. Очевидно, что методы обучения профиля пользователя должны быть инкрементными, т.е. они должны уточнять модель профиля пользователя по мере поступления новой информации. Во многих случаях это уточнение сводится к расширению множества примеров уже существующих классов понятий и модификации правил выработки рекомендаций.
- 4) В рекомендующих системах третьего поколения необходимо решать проблему кросс-доменных рекомендаций. Онтологическая модель профиля пользователя позволяет существенно упростить решение этой задачи, если его представить единой структурой на множестве понятий-интересов пользователя, которые относятся к различным предметным областям. Профиль пользователя при этом не должен быть представлен просто множеством локальных подмоделей, каждая из которых относится к своей предметной области. Их объединение с помощью понятий метауровня и использование технологий «выравнивания онтологий» (англ. *ontology alignment*) и другие приемы современной инженерии онтологий позволят связать знания из разных областей в единую модель.
- 5) Во многих работах онтология строится так, чтобы любому ее понятию можно было бы поставить в соответствие множество конкретных примеров из базы данных. Такая структура является предметом исследований в области анализа формальных понятий (англ. *Formal Concept Analysis*) [43]. Эта опция применения онтологий представляется весьма полезной: она позволяет проводить динамическую переоценку важности тех или иных интересов для пользователя по мере получения новых данных. Если рекомендующая система функционирует на продолжительном интервале времени (измеряемом годами), то информация о примерах тех или иных понятий онтологии и об их динамике позволит отслеживать «возрастную» динамику интересов пользователя, включая динамику его интересов в различных контекстах. Другие достоинства такой структуры представления обучающих данных кратко описаны в [44].
- 6) В данном обзоре не рассмотрены вопросы использования онтологий для представления зависимостей профиля пользователя от контекста, хотя именно эта компонента профиля пользователя во многом определяет современные тенденции развития теории и практики рекомендующих систем. Контекстно-зависимые рекомендации и роль онтологий в таких

системах – это очень обширная и важная тема, которая достойна отдельного обзора. Учёт контекста в рекомендующих системах исследуется уже около 12–15 лет. В настоящее время достигнутые в ней результаты позволяют создавать достаточно эффективные рекомендующие системы. Тем не менее, некоторые важные аспекты этой проблемы исследованы недостаточно. Например, пока нет однозначного мнения о том, как отделить контекст от самой предметной области рекомендаций и нужно ли это делать вообще, если принимать во внимание тенденцию использования онтологической модели контекста, интегрированной в предметную онтологию.

Среди ключевых проблем контекстно-зависимых рекомендующих систем авторы работы [8] упоминают необходимость сравнительного изучения разных моделей встраивания контекста в технологию поиска рекомендаций. Пока неясно, какая из предложенных моделей, а именно предварительная фильтрация, пост-фильтрация и встраивание контекста в модель знаний предпочтительнее и в каких случаях. По-видимому, возможны также и альтернативные методы учета контекста. Не менее важным является вопрос о полезности их разумной комбинации. Требуется дополнительное исследование задачи категоризации контекстов и возможности автоматизации построения онтологии контекстов. Необходимо более глубоко изучить корректность использования процедур обобщения контекстов, а также построить различные, например, топологические метрики оценки близости контекстов. Это поможет справиться с проблемой дефицита данных, которая характерна для задач обучения профиля пользователя в контекстно-зависимых системах.

Следует обратить внимание на то, что все проблемы вычислительной эффективности и качества решений, которые возникают в области интеллектуальной обработки *больших данных*, во многом характерны и для приложений в области рекомендующих систем третьего поколения. В обоих случаях необходимо решать задачу обработки данных большого объема и размерности. Как и в задачах построения рекомендующих систем, в задачах из области больших данных остро стоит проблема автоматизации процессов построения онтологий: в них число базовых понятий может исчисляться тысячами и более. Это говорит о том, что по сути своей задачи в области рекомендующих систем третьего поколения и задачи в области обработки больших данных во многом схожи. Поэтому исследования в этих актуальных областях могут и должны обогащать друг друга.

Благодарности

Данная работа выполнена при частичной поддержке Исследовательского Центра Самсунг, г. Москва, а также проекта 1.12 Программы фундаментальных исследований отделения нано- и информационных технологий Российской академии наук «Интеллектуальные информационные технологии, системный анализ и автоматизация».

Список источников

- [1] **Tuzhilin, A.** Keynote presentation at International Conference on Data Mining (ICDM 2013) / A. Tuzhilin. - Dallas, Texas, December, 2012.
- [2] **Cao, L.** Actionable knowledge discovery and delivery / L. Cao // In: WIREs Data Mining and Knowledge Discovery, Volume 2, March/April 2012, John Wiley & Sons, Inc., 2012. – P. 149–163.
- [3] **Argyri, C.** Actionable Knowledge / C. Argyri // The Oxford Handbook of Organization Theory. Edited by Knudsen C. and Tsoukas H. - Oxford University Press, 2005.
- [4] Recommender Systems Handbook / Ricci F., Rokach L. and Shapira B (Eds.). - Springer, 2011. – 842 P.
- [5] **Jannach, D.** Tutorial: Recommender Systems / D. Jannach, G. Friedrich // International Joint Conference on Artificial Intelligence.(Beijing, August 4, 2013). Available at <http://ijcai-11.iiia.csic.es/files/proceedings/Tutorial%20IJCAI%202011%20Gesamt.pdf> (Актуально на 18.07.2014).

- [6] **Segaran, T.** Programming Collective Intelligence / T. Segaran // O'Reilly, 2006 (Русский перевод: Тоби Сегаран. Программируем коллективный разум. Издательство Символ +, 2008. – 368 с.
- [7] **Adomavicius, G.** Incorporating contextual information in recommender systems using a multidimensional approach / G. Adomavicius, R. Sankaranarayanan, S. Sen, A. Tuzhilin // ACM Transactions on Information Systems (TOIS). – 2005 - 23(1). - P. 103–145.
- [8] **Adomavicius, G.** Context-Aware Recommender Systems / G. Adomavicius, B. Mobasher, F. Ricci, A. Tuzhilin // AI Magazine, FALL 2011. - P. 67–80.
- [9] **Adomavicius, G.** Context-Aware Recommender Systems / G. Adomavicius, A. Tuzhilin // In Ricci F., Rokach L, Shapira B, Kantor P (Eds.). Recommender Systems Handbook, Springer, 2011. - P. 217–256.
- [10] **Gauch, S.** Ontology-Based Personalized Search and Browsing / S. Gauch, J. Chaffee, A. Pretschner // ACM Web Intelligence and Agent System. - 2003. - Vol. 1. - No. 3/4. - P. 219–234.
- [11] **Leung, K.W.T.** Deriving Concept-Based User Profiles from Search Engine Logs / K.W.T. Leung, D.L. Lee // IEEE Transaction on Data and Knowledge Engineering. - 2010. - Vol. 22. - No. 7. - P. 969–982.
- [12] **Liu, F.** Personalized Web Search by Mapping User Queries to Categories / F. Liu, C. Yu, W. Meng // In Proc. of Intern. Conf. on Information and Knowledge Management (CIKM), 2002.
- [13] **Xu, Y.** Privacy-Enhancing Personalized Web Search / Y. Xu, K. Wang, B. Zhang, Z. Chen // Proceedings of World Wide Web (WWW) Conference, 2007. - P. 591–600.
- [14] Open Directory Project, <http://www.dmoz.org/> (Актуально на 18.07.2014).
- [15] **Costa, A.C.** Cores: Context-aware, ontology-based recommender system for service recommendation / A.C. Costa, R.S.S. Guizzardi, J.G.P. Filho // In Proc. 19-th Intern. Conf. on Advanced Information Systems Engineering (CAISE07). 2007.
- [16] **Middleton, S.E.** Ontological user profiling in recommender systems / S.E. Middleton, N.R. Shadbolt, D.C. de Roure // ACM Transaction on Information Systems. - 2004. - 22(1). - P. 54–88.
- [17] **Trajkova, J.** Improving Ontology-Based User Profile / J. Trajkova, S. Gauch // RIAO, 2004. - P. 380–390.
- [18] **O'Hara, K.** The AKT Manifesto / K. O'Hara, N. Shadbolt, S. Buckingham. - <http://citeseer.uark.edu:8080/citeseerx/showciting?cid=3895624> (Актуально на 18.07.2014).
- [19] **Bizer, C.** Linked Data – The Story So Far / C. Bizer, T. Heath, T. Berners-Lee // International Journal on Semantic Web and Information Systems (IJSWIS), 2009.
- [20] **Peña, P.** Collective Knowledge Ontology User Profiling for Twitter / P. Peña, R. del Hoyo, J. Vea-Murguía, C. González, S. Mayo // 2013 IEEE/WIC/ACM International Conferences on Web Intelligence (WI) and Intelligent Agent Technology (IAT), 2013.
- [21] OpenDNS cloud websites tagging: <http://community.opendns.com/domaintagging/> (Актуально на 18.07.2014).
- [22] DBpedia—a crowd-sourced community effort to extract structured information from Wikipedia: <http://dbpedia.org> (Актуально на 18.07.2014).
- [23] The Apache Cassandra database: <http://cassandra.apache.org/> (Актуально на 18.07.2014).
- [24] Virtuoso, a grade multi-model data server: <http://virtuoso.openlinksw.com/> (Актуально на 18.07.2014).
- [25] NoSQL data bases: <http://ru.wikipedia.org/wiki/NoSQL> (Актуально на 18.07.2014).
- [26] Pellet reasoning server: <http://clarkparsia.com/pellet/> (Актуально на 18.07.2014).
- [27] **Zeng, Y.** User Interests: Definition, Vocabulary, and Utilization in Unifying Search and Reasoning / Y. Zeng, Y. Wang, Z.S. Huang, D. Damjanovic, Zh. Ning, C. Wang // In An A. et al. (Eds.): Active Media Technology 2010, Lecture Notes in Computer Science, vol. 6335, Springer, 2010. - P. 98–107.
- [28] **Su, Z.G.** Research on Personalized Recommendation Algorithm Based on Ontological User Interest Model / Z.G. Su, J. Yan, H.F. Ling, H.S. Chen // J. of Computational Information Systems. – 2012. - Vol. 8. - No 1. - P. 169–181. Available also at <http://www.jofcis.com/> (Актуально на 10.07.2014).
- [29] **Cantador, I.** Enriching Ontological User Profiles with Tagging History for Multi-Domain Recommendations / I. Cantador, M. Szomszor, H. Alani, M. Fernández, P. Castells // In Proc of 1st Intern. Workshop on Collective Semantics: Collective Intelligence & the Semantic Web (CISWeb 2008), Tenerife, Spain. 2008.
- [30] Delicious: Social Bookmark manager: <http://ru.wikipedia.org/wiki/Delicious> (Актуально на 18.07.2014).
- [31] WordNet: <http://ru.wikipedia.org/wiki/WordNet> (Актуально на 18.07.2014).
- [32] **Miller, G.A.** WordNet: A Lexical Database for English / G.A. Miller // Communications of the Association for Computing Machinery. – 1995. - 38(11). - P. 39–41.
- [33] **Castells, P.** An Adaptation of the Vector-Space Model for Ontology-based Information Retrieval / P. Castells, M. Fernández, D. Valle // IEEE Transactions on Knowledge and Data Engineering. – 2007. - 19 (2). - P. 261–272.
- [34] **Vallet, D.** Personalized Content Retrieval in Context Using Ontological Knowledge / D. Vallet, P. Castells, M. Fernández, P. Mylonas, Y. Avrithis // IEEE Trans. on Circuits and Systems for Video Technology. – 2007. - 17(3). - P. 336–346.

- [35] **Alfonseca, E.** The Wraetlic NLP Suite / E. Alfonseca, A. Moreno-Sandoval, J.M. Guirao, M. Ruiz-Casado // In Proc. of the 5th Intern. Conf. on Language Resources and Evaluation. 2006. - P. 2277–2280. Available at <http://www.lrec-conf.org/proceedings/lrec2006/> (Актуально на 18.07.2014).
 - [36] Lucene: An Open Source Information Retrieval Library: <http://lucene.apache.org/> (Актуально на 18.07.2014).
 - [37] **García-Silva, A.** Social Tags and Linked Data for Ontology Development: A Case Study in the Financial Domain / A. García-Silva, L.J. García-Castro, A. García // In Proceedings of 4-th International Conference on Web Intelligence, Mining and Semantics (June 2–4, 2014, Thessaloniki, Greece).
 - [38] **Jaschke, R.** Discovering shared conceptualizations in folksonomies / R. Jaschke, A. Hotho, C. Schmitz, B. Ganter, G. Stumme // Web Semantics Science Services and Agents on the World Wide Web. - 2008. - 6(1). - P. 38-53.
 - [39] **Garcia-Silva, A.** Review of the state of the art: discovering and associating semantics to tags in folksonomies / A. Garcia-Silva, O. Corcho, H. Alani, A. Gomez-Perez // The Knowledge Engineering Review. - 2012. - 27(2). - P. 57-85.
 - [40] **Kadima, H.** Toward ontology-based personalization of a Recommender System in social network / H. Kadima, M. Malek // International Journal of Computer Information Systems and Industrial Management Applications. - 2013. - Vol. 5. - P. 499-508.
 - [41] **Ma, Y.F.** User Interests Modeling Based on Multi-source Personal Information Fusion and Semantic Reasoning / Y.F. Ma, Y. Zeng, R. Xu, Zh Ning // In Zhong N., Callaghan V., Ghorbani A., Hu B. (Eds.): Active Media Technology-2011, Lecture Notes in Computer Science, vol. 6890, Springer, 2011. - P. 195–205.
 - [42] **Varshney, P.K.** Multisensor data fusion / P.K. Varshney // Electronics & Communication Engineering J. – 1997. - 9(6). - P. 245–253.
 - [43] **Ganter, B.** Formal concept analysis: foundations and applications / B. Ganter, R. Wille. - Springer, 1999.
 - [44] **Gorodetsky, V.** Agent-based Customer Profile Learning in 3G Recommending Systems / V. Gorodetsky, V. Samoylov, O. Tushkanova // In: Proc. of the 9-th Intern. Workshop “Agent and Data Mining Interaction” (ADMI - 2014) associated with International Conference “Autonomous Agents and Multi-agent Systems” (AAMAS - 2014) (Paris, May 5–9 2014). To be also published in Post Proceedings of the Workshop as a volume of Lecture Notes in Computer Science, Springer, 2014.
-

ONTOLOGY-BASED USER PROFILE PERSONIFICATION IN 3G RECOMMENDER SYSTEMS

V.I. Gorodetsky¹, O.N. Tushkanova²

St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, St. Petersburg, Russia

¹*gor@iias.spb.su*, ²*tushkanova@iias.spb.su*

Abstract

3G recommender system (3G RecSys) is a novel and, in basic aspects, future paradigm of human decision support/prediction systems operating based on human-like semantic categories in knowledge representation, knowledge discovery and knowledge usage. According to the experts' opinion, it is expected that 3G RecSys will focus on semantically transparent personal user profile structuring his/her multidimensional personal interests and preferences. It will bring novel perspectives from data mining and knowledge discovery, while emphasizing user's decision explanation and detection of context-aware causality core determining this or that user's choice. As an effect of these capabilities, 3G RecSys should be capable, in what concerns with recommendation proposed by it, to answer “why?” questions using explanatory user interface. At present days, ontology is recognized as the most natural well-developed modeling framework and knowledge representation paradigm aimed at formal specification of human-like semantic categories specifically tuned for human-like knowledge-based decision-making. Although ontology-based user profile model is being developed from recent times, to present days, it has become the mature approach to user profile modeling. The paper presents a critical survey on current state-of-the-art in ontology-based user profiling for 3G RecSys. In this survey, the special attention is paid to specific properties of source information to be represented in ontology model of user profile and corresponding technologies proposed.

Key words: *recommender systems, ontologies, user profile, personalization, context-aware recommendations, folksonomies, network context, automation of ontology development.*

References

- [1] **Tuzhilin, A.** Keynote presentation at International Conference on Data Mining (ICDM 2013) / A. Tuzhilin. - Dallas, Texas, December, 2012.
- [2] **Cao, L.** Actionable knowledge discovery and delivery / L. Cao // In: WIREs Data Mining and Knowledge Discovery, Volume 2, March/April 2012, John Wiley & Sons, Inc., 2012. – P. 149–163.
- [3] **Argyri, C.** Actionable Knowledge / C. Argyri // The Oxford Handbook of Organization Theory. Edited by Knudsen C. and Tsoukas H. - Oxford University Press, 2005.
- [4] **Recommender Systems Handbook** / Ricci F., Rokach L. and Shapira B (Eds.). - Springer, 2011. – 842 P.
- [5] **Jannach, D.** Tutorial: Recommender Systems / D. Jannach, G. Friedrich // International Joint Conference on Artificial Intelligence.(Beijing, August 4, 2013). Available at <http://ijcai-11.iiia.csic.es/files/proceedings/Tutorial%20IJCAI%202011%20Gesamt.pdf>
- [6] **Segaran, T.** Programming Collective Intelligence / T. Segaran // O'Reilly, 2006.
- [7] **Adomavicius, G.** Incorporating contextual information in recommender systems using a multidimensional approach / G. Adomavicius, R. Sankaranarayanan, S. Sen, A. Tuzhilin // ACM Transactions on Information Systems (TOIS). – 2005 - 23(1). - P. 103–145.
- [8] **Adomavicius, G.** Context-Aware Recommender Systems / G. Adomavicius, B. Mobasher, F. Ricci, A. Tuzhilin // AI Magazine, FALL 2011. - P. 67–80.
- [9] **Adomavicius, G.** Context-Aware Recommender Systems / G. Adomavicius, A. Tuzhilin // In Ricci F., Rokach L., Shapira B., Kantor P (Eds.). Recommender Systems Handbook, Springer, 2011. - P. 217–256.
- [10] **Gauch, S.** Ontology-Based Personalized Search and Browsing / S. Gauch, J. Chaffee, A. Pretschner // ACM Web Intelligence and Agent System. - 2003. - Vol. 1. - No. 3/4. - P. 219–234.
- [11] **Leung, K.W.T.** Deriving Concept-Based User Profiles from Search Engine Logs / K.W.T. Leung, D.L. Lee // IEEE Transaction on Data and Knowledge Engineering. - 2010. - Vol. 22. - No. 7. - P. 969–982.
- [12] **Liu, F.** Personalized Web Search by Mapping User Queries to Categories / F. Liu, C. Yu, W. Meng // In Proc. of Intern. Conf. on Information and Knowledge Management (CIKM), 2002.
- [13] **Xu, Y.** Privacy-Enhancing Personalized Web Search / Y. Xu, K. Wang, B. Zhang, Z. Chen // Proceedings of World Wide Web (WWW) Conference, 2007. - P. 591–600.
- [14] Open Directory Project, <http://www.dmoz.org/>
- [15] **Costa, A.C.** Cores: Context-aware, ontology-based recommender system for service recommendation / A.C. Costa, R.S.S. Guizzardi, J.G.P. Filho // In Proc. 19-th Intern. Conf. on Advanced Information Systems Engineering (CAISE07). 2007.
- [16] **Middleton, S.E.** Ontological user profiling in recommender systems / S.E. Middleton, N.R. Shadbolt, D.C. de Roure // ACM Transaction on Information Systems. - 2004. - 22(1). - P. 54–88.
- [17] **Trajkova, J.** Improving Ontology-Based User Profile / J. Trajkova, S. Gauch // RIAO, 2004. - P. 380–390.
- [18] **O'Hara, K.** The AKT Manifesto / K. O'Hara, N. Shadbolt, S. Buckingham. - <http://citeseer.uark.edu:8080/citeseerx/showciting?cid=3895624>
- [19] **Bizer, C.** Linked Data – The Story So Far / C. Bizer, T. Heath, T. Berners-Lee // International Journal on Semantic Web and Information Systems (IJSWIS), 2009.
- [20] **Peña, P.** Collective Knowledge Ontology User Profiling for Twitter / P. Peña, R. del Hoyo, J. Veja-Murguía, C. González, S. Mayo // 2013 IEEE/WIC/ACM International Conferences on Web Intelligence (WI) and Intelligent Agent Technology (IAT), 2013.
- [21] OpenDNS cloud websites tagging: <http://community.opendns.com/domaintagging/>
- [22] DBpedia—a crowd-sourced community effort to extract structured information from Wikipedia: <http://dbpedia.org>
- [23] The Apache Cassandra database: <http://cassandra.apache.org/>
- [24] Virtuoso, a grade multi-model data server: <http://virtuoso.openlinksw.com/>
- [25] NoSQL data bases: <http://ru.wikipedia.org/wiki/NoSQL>
- [26] Pellet reasoning server: <http://clarkparsia.com/pellet/>
- [27] **Zeng, Y.** User Interests: Definition, Vocabulary, and Utilization in Unifying Search and Reasoning / Y. Zeng, Y. Wang, Z.S. Huang, D. Damjanovic, Zh. Ning, C. Wang // In An A. et al. (Eds.): Active Media Technology 2010, Lecture Notes in Computer Science, vol. 6335, Springer, 2010. - P. 98–107.
- [28] **Su, Z.G.** Research on Personalized Recommendation Algorithm Based on Ontological User Interest Model / Z.G. Su, J. Yan, H.F. Ling, H.S. Chen // J. of Computational Information Systems. – 2012. - Vol. 8. - No 1. - P. 169–181. Available also at <http://www.Jofcis.com/>

- [29] **Cantador, I.** Enriching Ontological User Profiles with Tagging History for Multi-Domain Recommendations / I. Cantador, M. Szomszor, H. Alani, M. Fernández, P. Castells // In Proc of 1st Intern. Workshop on Collective Semantics: Collective Intelligence & the Semantic Web (CISWeb 2008), Tenerife, Spain. 2008.
- [30] Delicious: Social Bookmark manager: <http://ru.wikipedia.org/wiki/Delicious>
- [31] WordNet: <http://ru.wikipedia.org/wiki/WordNet>
- [32] **Miller, G.A.** WordNet: A Lexical Database for English / G.A. Miller // Communications of the Association for Computing Machinery. – 1995. - 38(11). - P. 39-41.
- [33] **Castells, P.** An Adaptation of the Vector-Space Model for Ontology-based Information Retrieval / P. Castells, M. Fernández, D. Valle // IEEE Transactions on Knowledge and Data Engineering. – 2007. - 19 (2). - P. 261-272.
- [34] **Vallet, D.** Personalized Content Retrieval in Context Using Ontological Knowledge / D. Vallet, P. Castells, M. Fernández, P. Mylonas, Y. Avrithis // IEEE Trans. on Circuits and Systems for Video Technology. – 2007. - 17(3). - P. 336-346.
- [35] **Alfonseca, E.** The Wraetlic NLP Suite / E. Alfonseca, A. Moreno-Sandoval, J.M. Guirao, M. Ruiz-Casado // In Proc. of the 5th Intern. Conf. on Language Resources and Evaluation. 2006. - P. 2277–2280. Available at <http://www.lrec-conf.org/proceedings/lrec2006/>
- [36] Lucene: An Open Source Information Retrieval Library: <http://lucene.apache.org/>
- [37] **García-Silva, A.** Social Tags and Linked Data for Ontology Development: A Case Study in the Financial Domain / A. García-Silva, L.J. García-Castro, A. García // In Proceedings of 4-th International Conference on Web Intelligence, Mining and Semantics (June 2–4, 2014, Thessaloniki, Greece).
- [38] **Jaschke, R.** Discovering shared conceptualizations in folksonomies / R. Jaschke, A. Hotho, C. Schmitz, B. Ganter, G. Stumme // Web Semantics Science Services and Agents on the World Wide Web. - 2008. - 6(1). - P. 38-53.
- [39] **Garcia-Silva, A.** Review of the state of the art: discovering and associating semantics to tags in folksonomies / A. Garcia-Silva, O. Corcho, H. Alani, A. Gomez-Perez // The Knowledge Engineering Review. - 2012. - 27(2). - P 57-85.
- [40] **Kadima, H.** Toward ontology-based personalization of a Recommender System in social network / H. Kadima, M. Malek // International Journal of Computer Information Systems and Industrial Management Applications. - 2013. - Vol. 5. - P. 499-508.
- [41] **Ma, Y.F.** User Interests Modeling Based on Multi-source Personal Information Fusion and Semantic Reasoning / Y.F. Ma, Y. Zeng, R. Xu, Zh Ning // In Zhong N., Callaghan V., Ghorbani A., Hu B. (Eds.): Active Media Technology-2011, Lecture Notes in Computer Science, vol. 6890, Springer, 2011. - P. 195–205.
- [42] **Varshney, P.K.** Multisensor data fusion / P.K. Varshney // Electronics & Communication Engineering J. – 1997. - 9(6). - P. 245–253.
- [43] **Ganter, B.** Formal concept analysis: foundations and applications / B. Ganter, R. Wille. - Springer, 1999.
- [44] **Gorodetsky, V.** Agent-based Customer Profile Learning in 3G Recommending Systems / V. Gorodetsky, V. Samoylov, O. Tushkanova // In: Proc. of the 9-th Intern. Workshop “Agent and Data Mining Interaction” (ADMI - 2014) associated with International Conference “Autonomous Agents and Multi-agent Systems” (AAMAS - 2014) (Paris, May 5–9 2014). To be also published in Post Proceedings of the Workshop as a volume of Lecture Notes in Computer Science, Springer, 2014.

Сведения об авторах



Городецкий Владимир Иванович, 1937 г. рождения. Окончил Ленинградскую военно-воздушную инженерную академию им. А.Ф. Можайского (1960) и математико-механический факультет Ленинградского госуниверситета (1970), д.т.н. (1973), профессор (1991). Гл. научный сотрудник лаборатории интеллектуальных систем Санкт-Петербургского института информатики и автоматизации РАН. Член Российской ассоциации искусственного интеллекта, IEEE Computer Society, International Society of Information Fusion (ISIF), International Foundation for Autonomous Agents and Multi-agent Systems (IFAAMAS). Автор более 300 публикаций и нескольких монографий. Область научных интересов: искусственный интеллект, в частности, технология многоагентных систем и инструментальные средства, прикладные многоагентные системы, распределенное обучение, извлечение знаний из баз данных, анализ и объединение данных различных источников, P2P сети принятия решений и P2P методы извлечения знаний из данных, обработка больших данных, планирование и составление расписаний, алгоритмы улучшения изображений, получаемых с помощью мобильных устройств, рекомендующие системы. (<http://space.iias.spb.su/ai/gorodetsky>)

Vladimir Ivanovich Gorodetskiy, Professor of Computer Science, Chief Scientist of Intelligent Systems Lab. of St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, received MS degree in mechanics from the Military Air Force Engineering Academy in St. Petersburg (1960) and MS degree in mathematics from Mathematical and Mechanical Department of the St. Petersburg State University (1970); Ph.D and Doctor of Technical Sciences (1973) Science. Published more than 300 journal and conference papers and several books. Current scientific interests: Intelligent Data Analysis, Information Fusion, P2P Data Mining and Machine Learning, Multi-Agent Systems Technology and Software Tools, Agent-based Applications (Air Traffic Control, Intelligent Logistics, etc.), Recommender systems, Mobile Image Enhancement. Web site: <http://space.iias.spb.su/ai/gorodetsky>



Тушканова Ольга Николаевна, 1988 г. рождения. Получила степень магистра техники и технологии в области «Системный анализ и управление» в Южном федеральном университете, г. Ростов-на-Дону (2011). Аспирант, мл. научный сотрудник Санкт-Петербургского института информатики и автоматизации РАН. Область научных интересов: интеллектуальный анализ данных и извлечение знаний, многоагентные системы, рекомендующие системы, облачные технологии, онтологии.

Olga Nikolaevna Tushkanova (b. 1988) received MS degree in engineering and technology from Southern Federal University, Rostov-on-Don (2011). Currently is a graduate student and a junior researcher at the St. Petersburg Institute for Informatics and Automation of Russian Academy of Sciences. Current scientific interests: data mining, multi-agent systems, recommender systems, cloud computing, ontologies, knowledge extraction technologies.

УДК 004.81

ПРОЕКТИРОВАНИЕ ИНДИКАТОРОВ МОНИТОРИНГА В СФЕРЕ НАУКИ: ТЕОРЕТИЧЕСКИЕ ОСНОВАНИЯ И МОДЕЛИ*

И.М. Зацман¹, П.С. Бунтман²

¹Институт проблем информатики РАН, г. Москва, Россия
IZatsman@yandex.ru

²Независимый учёный, г. Москва, Россия
pbuntman@gmail.com

Аннотация

Рассматривается семиотический подход к описанию итераций проектирования новых индикаторов мониторинга, включая верифицируемые его варианты, и генерации новых экспертных знаний в процессе проектирования. Для описания итераций проектирования и динамики процессов генерации используется пространство Фреге. Результаты каждой итерации проектирования индикаторов фиксируются в базе экспертных знаний, названной проективным словарем, который является компонентом системы мониторинга и оценивания программно-целевой деятельности. С помощью этого словаря могут быть реализованы две основные функции системы мониторинга. При формировании индикаторов он позволяет экспертам описывать каждую итерацию процесса их проектирования, представлять результаты формирования различных вариантов индикаторов в графической форме, сравнивать и оценивать эти варианты. При использовании сформированных и утверждённых индикаторов словарь обеспечивает возможность верификации их значений с учётом динамики пополнения информационных ресурсов системы.

Ключевые слова: база экспертных знаний, генерация знаний, представление знаний, верифицируемый мониторинг, оценивание программной деятельности в сфере науки, семиотический подход.

Введение

В настоящее время в Институте проблем информатики Российской академии наук (ИПИ РАН) проводятся исследования процессов генерации экспертных знаний в разных предметных областях (ПрО), в частности, при проектировании новых индикаторов для мониторинга в сфере науки и при пополнении и детализации двуязычных контрастивных грамматик [1, 2]. В отличие от проектируемых физических артефактов, для формируемых экспертами индикаторов как цифровых объектов, часто допускающих многозначную интерпретацию, трудно предложить объективные критерии их работоспособности, результативности и эффективности. Поэтому для проектируемых индикаторов вместо объективных критериев используются, как правило, субъективные оценки самих экспертов-проектировщиков и/или внешних экспертов, не участвующих в проектировании. При этом требуется согласование личностных оценок в группе экспертов.

В теоретических разделах статьи представлен семиотический подход к исследованию процессов генерации и согласования экспертных знаний при проектировании новых индикаторов. В прикладных ее разделах рассматривается база экспертных знаний для поддержки процесса проектирования индикаторов мониторинга и оценивания программно-целевой деятельности в сфере науки, названная проективным словарем [3, 4].

* Публикация подготовлена в ИПИ РАН в рамках поддержанного РГНФ научного проекта № 12-02-00407.

Этот словарь был создан для формирования и применения новых индикаторов в четырех вариантах методологии мониторинга и оценивания:

- 1) стационарный вариант методологии без функции верификации вычисленных значений индикаторов, как правило, усредняющий значения на заданном интервале времени;
- 2) нестационарный вариант методологии без верификации;
- 3) стационарный вариант методологии с верификацией;
- 4) нестационарный вариант методологии с верификацией.

Категоризация вариантов методологии определяется двумя следующими признаками:

- стационарное оценивание программно-целевой деятельности или нестационарное (при нестационарном оценивании отображается динамика изменения вычисленных значений индикаторов в дискретных точках заданного интервала времени);
- верифицируемое оценивание программно-целевой деятельности или оценивание без верификации значений индикаторов.

В первом и третьем вариантах методологии проективный словарь требуется только для формирования новых индикаторов, а во втором и четвертом вариантах он может использоваться для формирования новых и применения утверждённых индикаторов.

Первый вариант методологии является одним из наиболее проработанных с теоретической и прикладной точек зрения. С его помощью вычисляются, например, индикаторы взаимосвязей областей знаний и критических технологий для проектов РФФИ, руководители которых выявили и зафиксировали эти взаимосвязи в отчётных формах по своим проектам на моменты времени их завершения.

Второй вариант также хорошо проработан и реализован во многих организациях, например, в Институте научной информации компании Томсон-Рейтер (система Web of Science), издательстве Elsevier (система Scopus), ООО «НЭБ» (система Science Index).

Каждый из последних двух вариантов методологии с верификацией является актуальной и во многом ещё нерешённой задачей, если учитывать несколько аспектов понятия верифицируемости мониторинга и оценивания. Чтобы проиллюстрировать многоаспектность этого понятия, отметим лишь два основных.

Первый аспект представлен принципом «сохранения истории» обработки той исходной информации, которая использовалась при вычислении значений индикаторов. Суть этого принципа заключается в том, что на каждом очередном этапе процесса обработки исходной информации, на котором целью является вычисление значений некоторого индикатора, фиксируются и запоминаются:

- результаты обработки информационных ресурсов, включая вычисленные значения самих индикаторов и промежуточные данные вычислений;
- все те исходные информационные ресурсы, которые использовались на этапах обработки с целью вычисления значений индикаторов, и связи полученных значений именно с этими информационными ресурсами;
- используемые версии программ вычисления значений индикаторов и связи полученных значений именно с этими программами.

Второй аспект представлен принципом «верификации по дополнительным источникам информации». Его суть заключается в том, что на стадии планирования программно-целевой деятельности выясняются возможности использования дополнительных и независимых источников информации, которые могут быть использованы для верификации значений индикаторов, вычисленных на основе результатов обработки информационных ресурсов, собираемых и обрабатываемых в процессе программно-целевой деятельности. Однако этот принцип реализуем лишь в тех случаях, когда есть эти дополнительные и независимые источники информации.

Проективный словарь был создан в рамках работ по Программе фундаментальных научных исследований Российской академии наук (ПФНИ РАН). ПФНИ РАН представляет собой инструмент программно-целевого управления в сфере фундаментальных научных исследований. Первоначально проективный словарь был разработан в ИПИ РАН в интересах создания новых программно-ориентированных индикаторов как компонент экспериментальной системы мониторинга и оценивания программно-целевой деятельности. При формировании индикаторов верифицируемого мониторинга он позволяет экспертам описывать каждую итерацию их проектировании, представлять результаты формирования различных вариантов индикаторов в графической форме, сравнивать и оценивать эти варианты [1, 3, 4].

Опыт его применения показал, что этот словарь может выполнять еще одну функцию в системе мониторинга: обеспечивать верификацию значений индикаторов в процессе оценивания программно-целевой деятельности для нестационарного варианта методологии мониторинга.

В процессе разработки проективного словаря были использованы две семиотические модели для описания итераций проектирования индикаторов и процессов генерации личностных и коллективных (согласованных) экспертных знаний в процессе проектирования:

- модель, описывающая состояние процесса проектирования индикаторов в некоторый момент времени, называемая моделью фиксированного состояния [5];
- модель для идентификации динамики процесса проектирования индикаторов, называемая нестационарной моделью [6].

Разработанный проективный словарь дает возможность экспертам фиксировать в нем, во-первых, их личностные знания в процессе интроспекции ментальных результатов проектирования, во-вторых, коллективные знания в процессе согласования личностных знаний между экспертами. С помощью словаря, в результате итерационного описания личностных и коллективных экспертных знаний, в компьютерной форме фиксируется динамика их формирования.

Процессы генерации новых концептов как ментальных единиц экспертных знаний об индикаторах по сравнению с процессами представления уже сформированных и общедоступных знаний не связаны с текстами как источниками экспертных знаний, так как концепты новых индикаторов еще не нашли свое отражение в существующих текстах. Источниками новых концептов являются объекты интерпретации, каждый из которых есть состояние денотата, являющегося совокупностью версии компьютерной программы вычисления значений нового индикатора, использованных при вычислении данных и самих вычисленных значений индикаторов.

Следовательно, до начала первой итерации формирования новых концептов экспертных знаний должен быть разработан предварительный вариант программы и отобраны некоторые исходные данные. Это условие говорит о том, что у экспертов уже должны быть некоторые предварительные знания о целевом назначении проектируемого индикатора, которые требуется модифицировать и/или детализировать в процессе проектирования, а затем согласовать в группе экспертов. Использование денотатов, являющихся артефактами цифровой природы, как источников экспертных знаний об индикаторах является основным отличием рассматриваемых процессов генерации новых знаний от процессов представления уже сформированных и устоявшихся знаний. Ниже приведены краткие сведения о процессе проектирования новых индикаторов мониторинга и оценивания:

- эксперты изучают в последовательные моменты времени некоторые объекты интерпретации, которые есть состояния денотата, являющегося совокупностью

программы вычисления значений индикатора, использованных при вычислении данных и самих вычисленных значений индикаторов;

- эксперты выполняют персональную семантическую интерпретацию этих объектов (результаты их интерпретации могут отличаться от эксперта к эксперту) и интроспекцию результатов пополнения и/или изменения своих персональных (личностных) знаний;
- эксперты фиксируют результаты интерпретации в проективном словаре (персональная семантическая интерпретация этих объектов описывается на каждой итерации в виде записей-дескрипторов словаря);
- эксперты сравнивают свои записи в словаре, описывающие их персональные семантические представления об одном и том же объекте интерпретации, после чего они пытаются сформулировать согласованное семантическое представление (в случае, если у экспертов есть разные варианты интерпретации, и они смогли их согласовать).

Отметим, что на каждой итерации эксперты интерпретируют состояние денотата индикатора в некоторый момент времени, анализируя программу вычисления его значений, использованные при вычислении данные и сами вычисленные значения. При этом денотат может изменяться во времени, так как между итерациями проектирования экспертами-проектировщиками могут вноситься изменения в программу и/или данные, что, как правило, приводит к изменению вычисленных значений проектируемых индикаторов.

Для описания семиотического подхода к моделированию процесса проектирования и формирования экспертных знаний необходимо предварительно определить систему терминов. В статье даётся описание предлагаемого подхода, который включает систему терминов, две семиотических модели, семиотическое пространство Фреге и концепцию проективного словаря. Для демонстрации реализуемости семиотического подхода приводится описание результатов эксперимента по моделированию итераций формирования экспертных знаний в процессе проектирования новых индикаторов.

1 Система терминов

Задача создания чётких дефиниций используемых терминов является актуальной и для описания моделей формирования экспертных знаний, и для информатики в целом. Причина в том, что в информатике (да и в других отраслях науки) каждое из слов «знания», «данные» и «информация» имеет несколько разных значений [7-9].

Главной целью этого раздела является краткое описание основных терминов, которые будут обозначать основные понятия в моделях процессов формирования экспертных знаний. Эти термины были определены в работах [5, 10]; для построения этих определений использовались результаты работ [11, 12].

Ниже представлены основные термины для описания моделей процессов формирования экспертных знаний, относящиеся к трём средам ПрО информатики как информационно-компьютерной науки [13-15], названия которых указаны в скобках после определения каждого термина (см. также рисунок 1):

- *знание* – результаты познавательной и креативной деятельности человека, носителем которых может быть только человек и в которых могут быть выделены отдельные «кванты» знаний (ментальная среда);
- *концепты* – «кванты» знаний, которые могут быть выражены в рамках некоторого естественного языка, а в общем случае, в рамках знаковой системы; все концепты делятся на личностные, коллективные (здесь – согласованные группой экспертов), и конвенциональные (ментальная среда);

- *знаковая информация* – результаты процесса представления концептов человеком-генератором этих результатов в отчуждённых формах в рамках некоторой семиотической знаковой системы, которые являются сенсорно воспринимаемыми другими участниками коммуникаций и содержательно интерпретируются ими в рамках этой знаковой системы (социально-коммуникационная среда);
- *коды* – компьютерные эквиваленты двоичных цифр (или их последовательностей), которые могут представлять собой намагниченность или ее отсутствие, наличие электрического тока или его отсутствие, способность к отражению света или ее отсутствие [12] (цифровая среда, электрическая и/или оптическая).

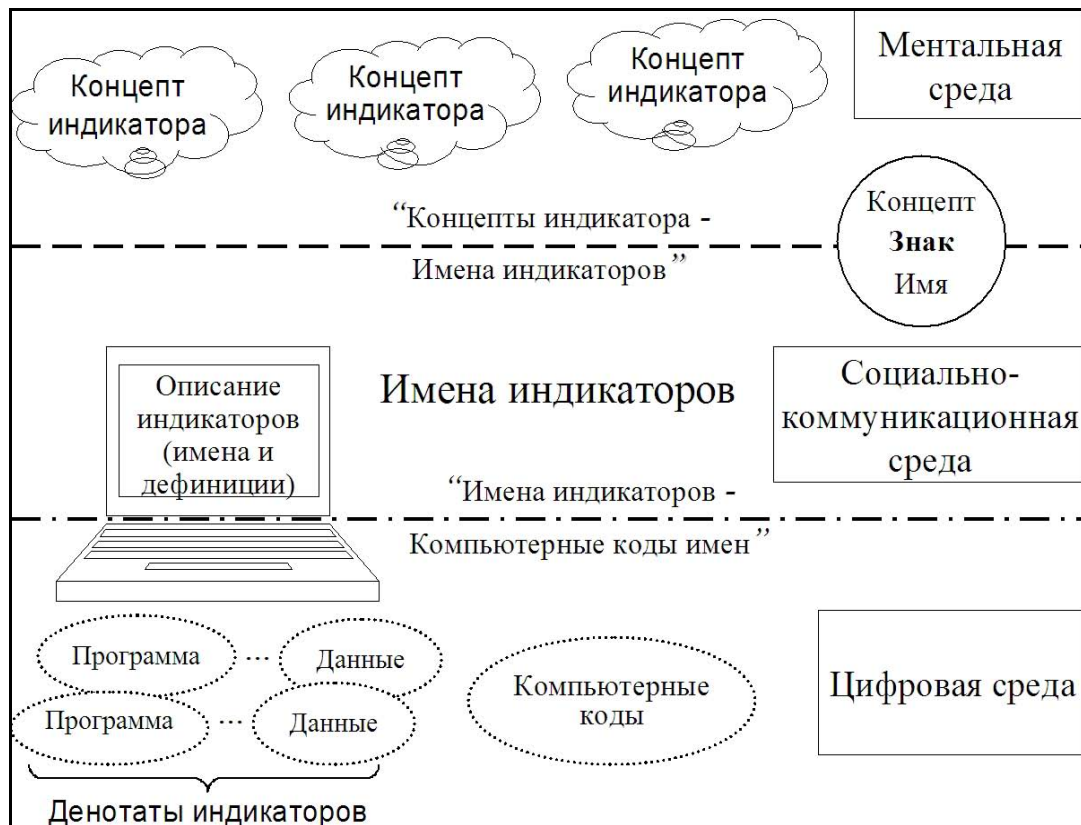


Рисунок 1 - Три среды предметной области информатики как информационно-компьютерной науки

В статье определены три категории кодов и ряд терминов, которые относятся к границам между вышеуказанными средами (ментальной, социально-коммуникационной и цифровой). Дефиниции строятся таким образом, чтобы из них всегда была понятна та среда или граница, к которой принадлежат обозначаемые объекты или сущности. Например, термины «знание» и «концепты» по определению относятся только к ментальной среде, «знаковая информация» - к социально-коммуникационной среде, а «коды» - к цифровой среде. Термин «знак» в семиотическом понимании [11], который будет использован в дальнейшем при описании моделей, относится как двуединая по своей природе сущность к границе между ментальной и социально-коммуникационной средами.

Обычно при обозначении концептов знаний используются имена или слова, применяя некоторый естественный язык или семиотическую знаковую систему. В парадигме трёх упомянутых сред язык как система вербальных знаков относится к границе между ментальной и социально-коммуникационной средами. На рисунке 1 показан один знак, объединяющий концепт с его именем, т. е. он есть двуединая по своей природе сущность.

Вторая граница находится между социально-коммуникационной и цифровой средами. Эта граница отделяет имена и любые другие слова от соответствующих им компьютерных кодов. При описании информатики как информационно-компьютерной науки не удалось найти устоявшегося термина, обозначающего двуединую сущность, объединяющую имя или любое другое слово как носитель смысла с компьютерным кодом (так же как знак объединяет концепт и его имя). Скорее всего, у этого понятия ещё нет устоявшегося обозначения¹. Поэтому в работе [16] оно было обозначено новым термином «формокод». В терминах трёх упоминавшихся сред формокоды принадлежат границе между социально-коммуникационной и цифровой средами. На рисунке 2 формокод изображён именно на этой границе.

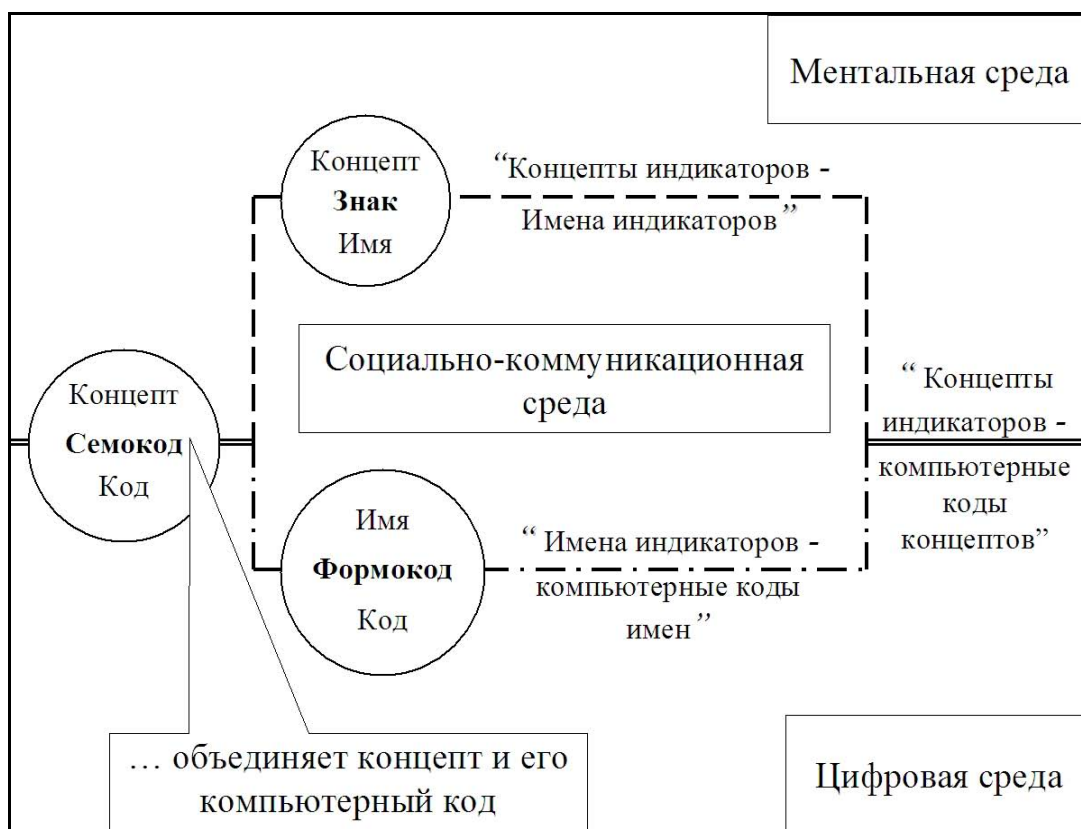


Рисунок 2 - Три среды предметной области информатики и три границы между ними

Существует третья граница между ментальной и цифровой средами. Эта граница отделяет концепты как ментальные единицы знаний человека от их компьютерного представления. Скорее всего, и эта двуединая сущность, объединяющая концепт с его компьютерным кодом, также не имеет устоявшегося обозначения. Поэтому в работе [16] это понятие было обозначено термином «семокод». Граница между этими средами изображена на рисунке 2 в виде двойной линии.

Таким образом, определена система терминов, которые будут использоваться далее для описания моделей формирования экспертных знаний в процессе проектирования новых индикаторов группой экспертов.

¹ Понятие таблицы кодирования здесь не может быть использовано, т. к. в этой таблице кодируемые символы, в общем случае, не являются носителями смысла, а играют только смысловоразличительную роль.

2 Две модели формирования экспертных знаний

Идея создания модели фиксированного состояния и нестационарной модели появилась в ходе изучения спиральной модели процесса генерации знаний, предложенной Икуджиро Нонака [17-19], и креативного пространства, описанного Вержбицким и Накамори [20, 21]. Эти семиотические модели (фиксированного состояния и нестационарная) основаны на семиотическом треугольнике Фреге, который в данном частном случае состоит из 3 вершин: концепта, имени и денотата индикатора.

В терминах семиотики для каждого состояния проектируемого индикатора в некоторый момент времени эти три вершины соответственно являются: значением знака (здесь концептом индикатора), формой знака (именем индикатора) и денотатом знака (здесь совокупностью версии компьютерной программы вычисления значений индикатора, использованных программой данных и самих вычисленных значений индикатора в этот момент времени). Будем использовать три среды ПрО информатики (рисунки 1 и 2) для описания семиотических моделей. На рисунке 3, изображающем семиотическую модель фиксированного состояния, цифровая среда условно разделена на две части. Верхняя ее часть включает денотаты индикаторов. Нижняя часть содержит компьютерные коды трёх разных категорий: для денотатов (объектные коды), концептов (семантические коды) и их имен (информационные коды). По определению, три компьютерных кода трёх разных категорий образуют цифровой семиотический треугольник [5, 6].

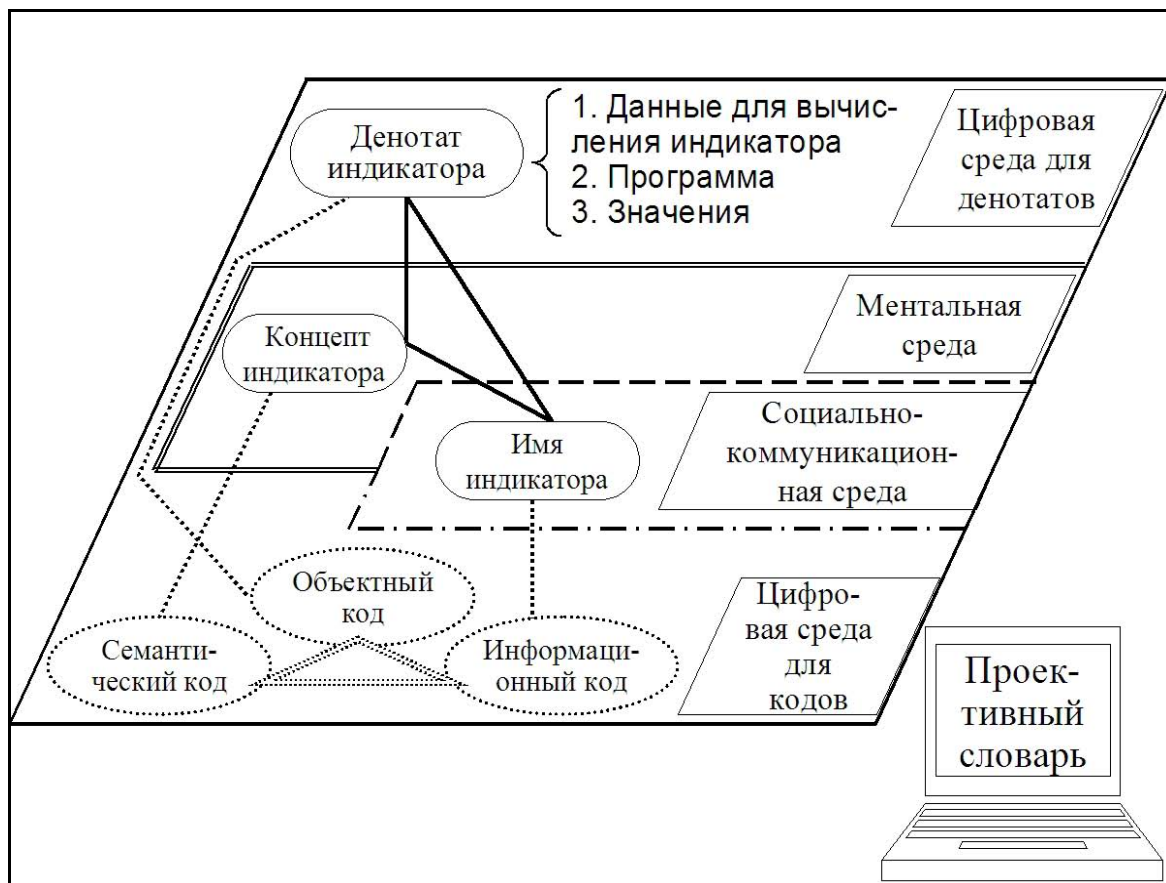


Рисунок 3 - Семиотическая модель фиксированного состояния процесса формирования экспертных знаний при проектировании индикаторов

Построение этого треугольника является начальным этапом описания модели фиксированного состояния и нестационарной модели. По определению из работы [5] первая семиотическая модель фиксирует состояние процесса формирования экспертных знаний в момент времени, соответствующий окончанию некоторой итерации процесса проектирования, и состоит из (см. рисунок 3):

- трёх сред ПрО информатики: ментальной, социально-коммуникационной и цифровой;
- треугольника Фреге, включающего денотат, концепт и имя формируемого индикатора в этот момент времени;
- цифрового семиотического треугольника, включающего компьютерные коды денотата, концепта и имени формируемого индикатора в этот же момент времени.

Согласно модели фиксированного состояния, три вершины треугольника Фреге кодируются тремя компьютерными кодами разных категорий, которые записываются в проективный словарь в конце каждой итерации:

- семантический код для концепта индикатора;
- информационный код для его имени, если оно создано экспертом (в противном случае этот код равен нулю);
- объектный код для его денотата.

Таким образом, после завершения каждой итерации, в проективном словаре по некоторому заданному алгоритму генерируются три компьютерных кода. Главная идея компьютерного кодирования с одновременным использованием кодов трёх категорий заключается в том, чтобы каждый проектируемый индикатор анализировался и описывался экспертом на каждой итерации с трёх точек зрения:

- как концепт нового индикатора;
- как итерационно уточняемое имя индикатора;
- как состояние денотата индикатора, которое может меняться между итерациями.

Если на некоторой итерации в проектировании участвовал один эксперт, то в проективном словаре создается только одна запись с результатами личностной семантической интерпретации и только три кода, а если несколько экспертов, то для каждого эксперта – одна запись и три кода. До начала согласования мнений экспертов о некотором индикаторе в конце итераций концепты, имена и денотаты кодируются как личностные с указанием идентификаторов соответствующих экспертов, их создавших.

Кроме итераций проектирования, существует отдельный вид итераций для согласования личностных концептов, имён и денотатов проектируемых индикаторов. На этих итерациях эксперты пытаются согласовать между собой свои личностные интерпретации и сформировать коллективные концепты и имена. Если это удастся им сделать, то в проективном словаре создается ещё одна запись с результатами коллективной семантической интерпретации и ещё три кода с указанием идентификаторов всех экспертов, которые приняли участие в процессе согласования и выработали единую позицию. Проектирование и согласование могут быть совмещены в рамках одной комплексной итерации.

Важно отметить, что перед любой итерацией проектирования индикатора или комплексной итерацией, совмещающей проектирование и согласование, денотат может меняться экспертом. Каждый эксперт-проектировщик может менять свой вариант программы и данных до начала итерации, а затем вычислить значениями индикатора и интерпретировать их вместе с программой и данными. Поэтому эксперты могут зафиксировать в проективном словаре несколько вариантов программы вычисления значений индикатора, используемых при вычислении данных и полученных значений индикаторов, а также соответствующие им результаты семантической интерпретации в виде своих дефиниций индикатора, их

тематических рубрик в классификаторе индикаторов и их имен (которые факультативны в начале проектирования индикаторов).

Генерация компьютерных кодов трёх категорий даёт возможность зафиксировать в проективном словаре (а затем отслеживать) следующие ситуации, список которых не претендует на полноту:

- состояние денотата (программа и данные) не изменилось между итерациями, при этом эксперт всё равно изменил свою семантическую интерпретацию, выразив это в новой дефиниции индикатора и/или в тематических рубриках, но не изменил имя индикатора;
- состояние денотата индикатора не изменилось между итерациями, эксперт изменил имя индикатора, но не изменил семантическую интерпретацию этого состояния денотата;
- состояние денотата не изменилось между итерациями, эксперт изменил свою семантическую интерпретацию, выразив это в дефиниции индикатора и/или в тематических рубриках, и имя индикатора;
- изменились только исходные данные в состоянии денотата между итерациями, эксперт изменил свою семантическую интерпретацию, выразив это в дефиниции индикатора и/или в тематических рубриках;
- изменились только исходные данные между итерациями, эксперт изменил только имя индикатора;
- изменились только исходные данные между итерациями, при этом эксперт изменил свою семантическую интерпретацию, выразив это в дефиниции индикатора и/или в тематических рубриках, а также имя индикатора.

Регистрация подобных ситуаций помогает восстановить ретроспективно все итерации проектирования индикаторов и их результаты.

Вторая, нестационарная, модель предназначена для идентификации динамики процесса проектирования индикаторов. В процессе проектирования концепты индикаторов, их имена и денотаты могут изменяться экспертами в широком диапазоне с целью согласования личностных концептов и имён индикаторов. По определению [6], нестационарная модель процесса проектирования одного индикатора одним экспертом состоит из (см. рисунок 4):

- трёх сред ПрО информатики: ментальной, социально-коммуникационной и цифровой;
- треугольников Фреге, построенных в моменты времени окончания итераций t_i , $i = 1, 2, \dots$;
- цифровых семиотических треугольников, построенных в эти же моменты времени t_i .

Аналогично можно определить модели для случая одного индикатора, проектируемого коллективом экспертов, пример которого рассматривается далее в статье, и случая нескольких индикаторов, проектируемых коллективно.

Нестационарная модель базируется на модели фиксированного состояния. На рисунке 4 представлены две итерации проектирования нового индикатора, первая из которых изображена в левой части этого рисунка. До начала первой итерации формируется денотат индикатора, интерпретируемый экспертом. В результате интерпретации может быть получен концепт и дано имя индикатору. Эта итерация завершается в момент времени t_1 генерацией цифрового семиотического треугольника, состоящего из трёх компьютерных кодов: семантического, информационного и объектного.

Вторая итерация проектирования нового индикатора изображена в правой части рисунка 4. В этом примере между первой и второй итерациями денотат был изменён и затем заново интерпретирован экспертом. В процессе интерпретации концепт и имя индикатора были изменены тем же экспертом. Эта итерация завершается в момент времени t_2 генерацией второго цифрового семиотического треугольника.

Используя нестационарную модель, в работе [6] было определено *пространство Фреге* как 4-мерное множество точек $\{t_i$, объектный код (t_i), семантический код (t_i),

информационный код (t_i) при $i = 1, 2, \dots$, сгенерированных в процессе проектирования одного индикатора одним экспертом. Аналогично можно определить пространство Фреге для случая одного индикатора, проектируемого коллективом экспертов, и для случая нескольких индикаторов, проектируемых коллективно.

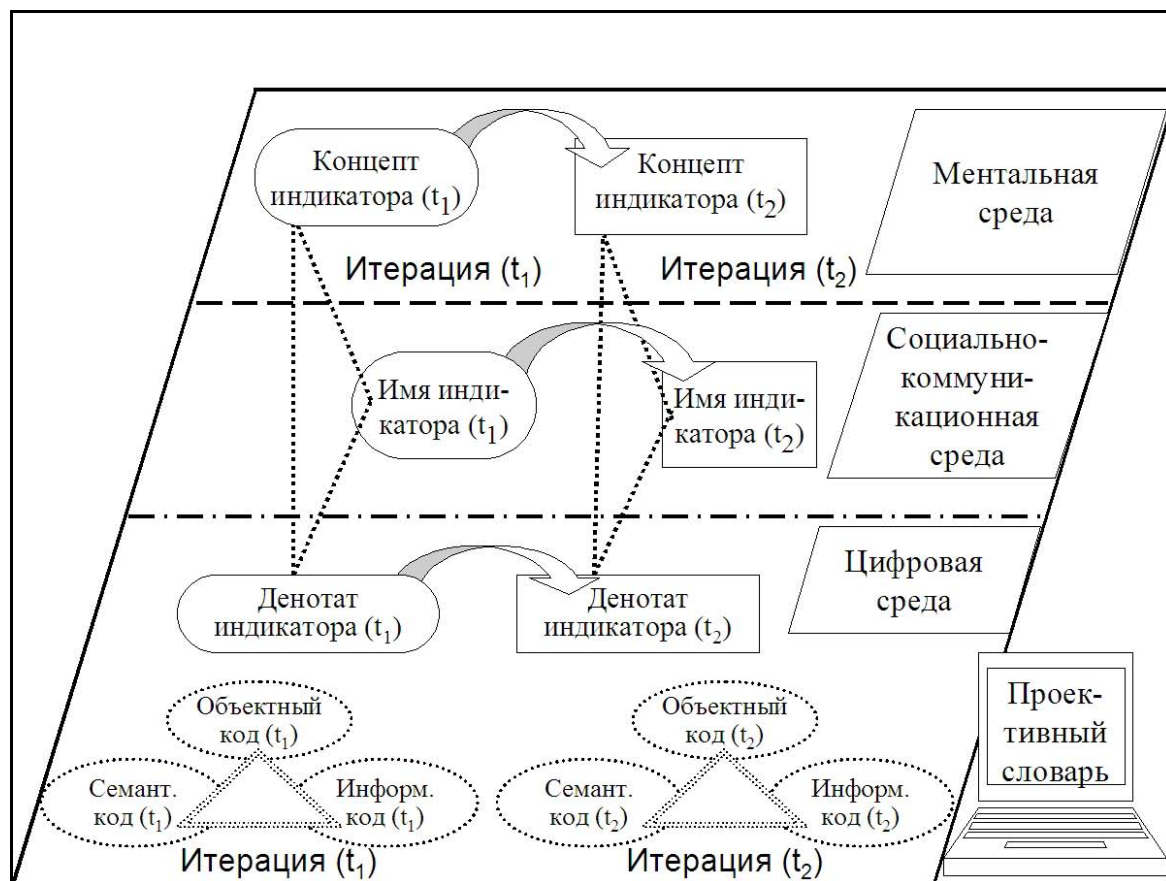


Рисунок 4 - Нестационарная семиотическая модель для двух итераций процесса проектирования одного индикатора одним экспертом

Пространство Фреге имеет три оси координат: семантическую, информационную и объектную, а так же четвертую – ось времени, содержащую дискретный набор точек окончания итераций проектирования, согласования или комплексных итераций (на рисунке 5 ось времени не показана). По определению пространство Фреге даёт возможность представить графически динамику процесса проектирования, используя последовательности значений семантических, информационных и объектных кодов в дискретные моменты времени их генерации.

Важно отметить, что вид дискретных траекторий точек будет определяться алгоритмами назначения семантических, информационных и объектных кодов. Можно ли выбирать эти алгоритмы таким образом, чтобы задать некоторую метрику в пространстве Фреге? В настоящее время этот вопрос является открытым. Пока этот вопрос находится в стадии изучения, термин «пространство Фреге» определён только как семиотическое понятие, но не математическое. Использование термина «семиотическое пространство Фреге» для 4-мерного множества точек подчеркивает это различие лексически.

Таким образом, нестационарная модель описывает все итерации процесса проектирования и согласования индикатора как дискретную траекторию в моменты времени (t_i , $i = 1, 2, \dots$), где t_i - это время завершения i -й итерации. При этом если проективный словарь

фиксирует эволюцию смыслового содержания (семантику) проектируемых индикаторов, то семиотическое пространство Фреге служит для количественного описания итераций проектирования.

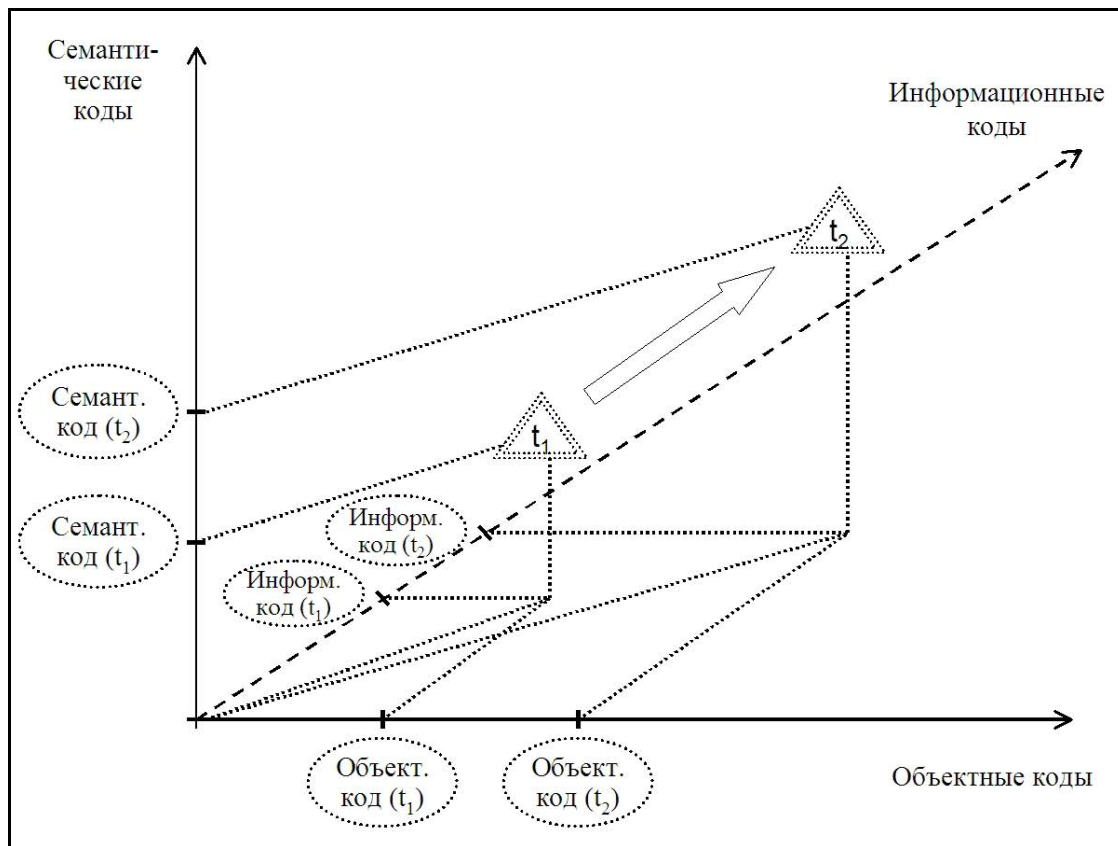


Рисунок 5 - Две точки в семиотическом пространстве Фреге, в котором ось времени в явном виде не показана, но переход от первого момента времени t_1 ко второму t_2 условно передан стрелкой

3 Проективный словарь

Пользователями проективного словаря могут быть только те эксперты-проектировщики и/или внешние эксперты, которые включены в группу по разработке новых индикаторов. Каждый индикатор разрабатывается на основе интерпретации его изменяемых денотатов, как правило, группой экспертов, хотя возможны случаи индивидуального проектирования. Поэтому словарь может поддерживать индивидуальный и коллективный режимы работы. Эксперты могут обращаться к существующим записям проективного словаря, называемых дескрипторами. Словарь является компонентом компьютерной системы мониторинга и оценивания программно-целевой деятельности и позволяет получать информацию о любом зафиксированном состоянии каждого из разрабатываемых индикаторов.

Помимо проективного словаря, который регулярно пополняется дескрипторами новых индикаторов (обозначены литерами U и V на рисунке 6), система включает семантический словарь [22], содержащий уже разработанные, согласованные и утвержденные индикаторы (обозначены литерами X и Y).

На рисунке 6 показаны семантический и проективный словари системы мониторинга и оценивания, поддерживающей только стационарные варианты методологии мониторинга и оценивания.

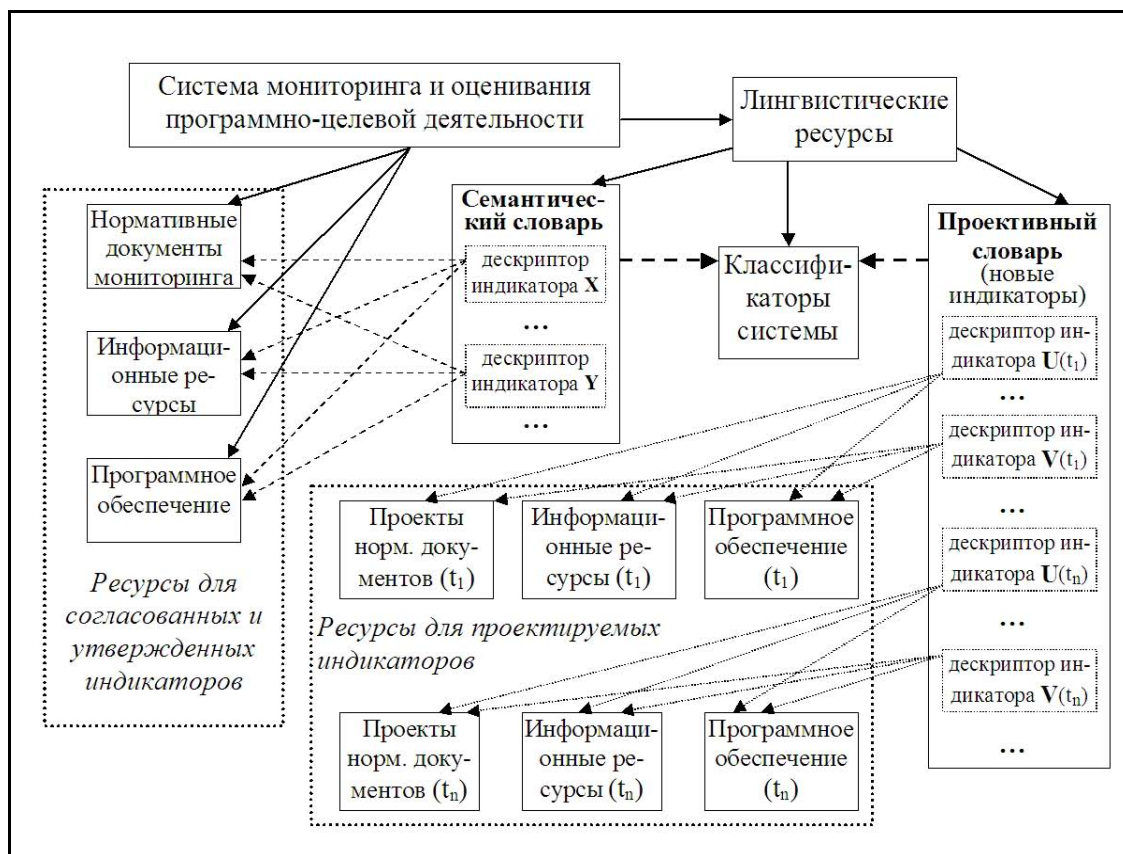


Рисунок 6 - Семантический и проективный словари системы мониторинга и оценивания

В проективном словаре фиксируются следующие события в процессе разработки:

- создание дескриптора для фиксированного состояния любого разрабатываемого индикатора (с указанием одного его автора при личностной интерпретации или нескольких авторов при согласованной интерпретации) с фиксацией момента времени окончания итерации;
- логическая модификация существующего дескриптора, включая изменение ссылок между дескрипторами словаря, что означает автоматическое создание нового дескриптора, так как физическая их модификация запрещена;
- изменение ссылок на компьютерные программы, которые вычисляют значения индикатора, и/или ссылок на его исходные данные с указанием времени окончания их модификации, что возможно только между итерациями.

По своей сути проективный словарь является темпоральной базой экспертных знаний. Каждый его дескриптор содержит момент времени окончания итерации, дефиницию индикатора, его рубрики и имя (если присвоены), три компьютерных кода, ссылки на другие дескрипторы, а так же ссылки на исходные данные, использованные при расчёте значений индикатора (см. ссылки в двух дескрипторах на информационные ресурсы в t_1 и t_n на рисунке 6), и на компьютерные программы, которые вычисляют значения индикатора (см. ссылки на программные ресурсы в t_1 и t_n).

Как было отмечено, помимо проективного словаря, регулярно пополняемого дескрипторами новых индикаторов, содержимое которых «привязано» к определённым моментам времени, система мониторинга и оценивания программно-целевой деятельности включает семантический словарь, который включает уже разработанные, согласованные и утверждённые индикаторы, т. е. неизменяемые.

Иначе говоря, проективный словарь содержит всю историю изменений денотатов, концептов и имён проектируемых индикаторов, а семантический словарь описывает только неизменяемые индикаторы, которые уже согласованы экспертами и утверждены лицами, принимающими решения. Каждый проектируемый индикатор потенциально может быть перенесён в семантический словарь. Для этого эксперты должны прийти к согласию по поводу его денотата, концепта и имени (что не всегда получается), после чего он должен быть утверждён вместе с нормативным документом, определяющим порядок его использования.

В рассмотренной архитектуре системы мониторинга и оценивания поддерживаются только стационарные варианты методологии мониторинга и оценивания. В этой архитектуре системы семантический словарь не является темпоральным, что не даёт возможность восстановить всю историю и динамику вычисления значений утверждённых индикаторов. Для поддержки нестационарных вариантов методологии мониторинга и оценивания необходима темпоральная версия семантического словаря, которая с функциональной точки зрения будет идентична проективному словарю, но регламент её использования будет отличаться кардинально.

Во-первых, при использовании утверждённых индикаторов любые изменения алгоритма вычисления значений индикаторов допускаются только в форме создания новой версии программы, которая должна пройти процедуру её утверждения лицами, принимающими решения, т. е. экспертные изменения алгоритма вычислений не разрешаются.

Во-вторых, информационные ресурсы, используемые программами вычисления значений утверждённых индикаторов, могут только пополняться, но не могут изменяться по их содержанию, так как это может привести к изменению смысла используемого индикатора (см. в следующем разделе описание четвёртой итерации проектирования индикатора, на которой эксперты принимают решение учитывать только те статьи, которые опубликованы в журналах из перечня ВАК, что меняет смысл проектируемого ими индикатора). Следовательно, только при перечисленных ограничениях проективный словарь может быть использован для поддержки нестационарных вариантов методологии мониторинга и оценивания программно-целевой деятельности. В этом случае, с функциональной точки зрения, он по своей сути будет выполнять функции темпоральной версии семантического словаря.

В заключение отметим, что для проективного словаря нет операций физического удаления или изменения дескрипторов. Любое логическое их выполнение приводит к созданию нового дескриптора с сохранением старого. При этом новые дескрипторы помечаются как «наследники» существующих. В проективном словаре такое «наследование» может быть множественным, потому что новый дескриптор может быть описан как «наследник» нескольких уже существующих с помощью ссылок между дескрипторами [1].

4 Пример проектирования индикатора

В ИПИ РАН был проведён эксперимент, в рамках которого были разработаны варианты индикатора, который характеризует распределение возраста авторов статей. Использовался массив публикаций с результатами проектов 2007-2008 годах подпрограммы «Фундаментальные исследования - медицине», являющейся частью ПФНИ РАН. Одной из задач подпрограммы является привлечение молодых исследователей к участию в её проектах. Эта подпрограмма включала 130 проектов в 2007 году и 146 - в 2008. Общее число участников составило 1334 в 2007 и 1504 в 2008 годах. Один из утверждённых индикаторов подпрограммы характеризует долю её участников в возрасте до 39 лет. Однако этот

индикатор не давал информации об их результативности, так как среди утверждённых индикаторов подпрограммы отсутствовали показатели распределения по возрастным группам авторов публикаций с результатами проектов подпрограммы. Поэтому, было предложено разработать новый индикатор, описывающий *возрастное распределение публикаций* (ВРП – это рубрика классификатора индикаторов, а не его имя) по 14 возрастным группам (20 – 24, 25 – 29 и так далее до 85 – 89 лет).

В текстовой форме опишем первые пять итераций, из которых для трёх представлены вычисленные значения индикаторов на рисунке 7 в графической форме (для одной комплексной итерации и двух итераций проектирования). В разработке нового индикатора принимали участие три эксперта-проектировщика, обозначенных литерами А, В, С и два внешних эксперта: D и E.



Рисунок 7 - Результаты вычислений двух вариантов индикатора на первой, третьей и четвертой итерациях

Первая итерация (комплексная). Эксперт А вычисляет значения первого варианта индикатора и формирует дескриптор, используя программу и данные его денотата, созданные до начала этой итерации при следующих условиях:

- 1) для вычисления значений этого варианта индикатора используются статьи участников проектов, напечатанные в любых журналах и сборниках (независимо от их рейтинга);
- 2) возрастная группа каждого соавтора статьи получает за эту статью 1 балл;
- 3) отсутствует нормализация относительно численности возрастных групп.

В это же время эксперты В и С формируют второй дескриптор совместно и согласованно, вычисляя значения второго варианта индикатора, используя программу и данные его

денотата, созданные до начала этой итерации при тех же первом и третьем условиях, но второе из трёх вышеперечисленных условий отличается, а именно: возрастная группа каждого соавтора получает $1/N$ балла за статью, у которой N соавторов.

Таким образом, на первой итерации строятся два дескриптора проективного словаря и генерируются два цифровых семиотических треугольника.

Вторая итерация (согласования). Эксперт C решает изменить свою точку зрения и принять точку зрения эксперта A . Иначе говоря, эксперт C на второй итерации отказывается в согласовании второму варианту индикатора, так как считает правильным добавлять всем соавторам по одному баллу, и согласовывает первый вариант индикатора. Первый вариант, спроектированный экспертом A становится согласованным (коллективным), а второй – авторским (личностным) эксперта B . Эти варианты индикатора идентичны вариантам первой итерации, но они отличаются уровнем согласованности. Иначе говоря, имеющиеся концепты, имена и денотатов на второй итерации не изменялись, а поменялись только уровни их согласованности экспертами. Поэтому на рисунке 7 эта итерация не показана.

Третья итерация (проектирования). Эксперты A , B и C принимают решения учесть численности возрастных групп при вычислении значений своих вариантов индикатора, что до начала итерации находит отражение в изменении соответствующих алгоритмов программ вычисления их значений с целью нормализации. Строятся ещё два дескриптора проективного словаря и генерируются два цифровых семиотических треугольника. При этом эксперты связывают новые дескрипторы с теми, которые изменили степень согласованности на второй итерации, отношениями наследования.

Четвертая итерация (проектирования). Эксперты A , B и C принимают решение учитывать только те статьи, которые опубликованы в журналах из перечня ВАК. Строятся ещё два дескриптора проективного словаря и генерируются два цифровых семиотических треугольника. На этой итерации эксперты связывают новые дескрипторы с теми, которые были построены на третьей итерации, отношениями наследования.

Пятая итерация (согласования). Варианты индикатора на этой итерации идентичны вариантам четвёртой итерации, но они отличаются уровнями согласованности, так как точки зрения экспертов изменяются следующим образом. Эксперт C отказывается от согласия на первый вариант эксперта A , он не даёт своего согласия на второй вариант эксперта B , но он может предложить новый вариант на одной из следующих итераций. При этом вариант эксперта B приобретает двух новых сторонников – внешних экспертов D и E , что приводит к изменению уровня его согласованности. На рисунке 7 эта итерация не показана, т.к. варианты индикатора на этой итерации идентичны вариантам предыдущей итерации.

В течение первых пяти рассмотренных итераций изменение уровней согласованности первого и второго вариантов представим в графической форме в виде двух дискретных функций $F_{\text{coor}}(t_i)$ на рисунке 8. Значения каждой из этих двух функций в момент времени окончания итерации t_i равно числу экспертов, согласовавших тот или иной вариант проектируемого индикатора в момент времени t_i .

Заключение

Предложенный семиотический подход к моделированию итераций формирования и согласования экспертных знаний в процессе проектирования новых индикаторов базируется только на субъективной интроспекции экспертов. Возникают закономерные вопросы. Почему не рассматривалась возможность дополнительного использования объективных нейроданных о функционировании мозга в процессе генерации и согласования новых знаний экспертами? Почему не использовались данные количественных измерений мозговой

деятельности экспертов (например, методом функциональной магнитно-резонансной томографии - Functional Magnetic Resonance Imaging, FMRI) для сопоставления с результатами процессов их личностного субъективного мышления?

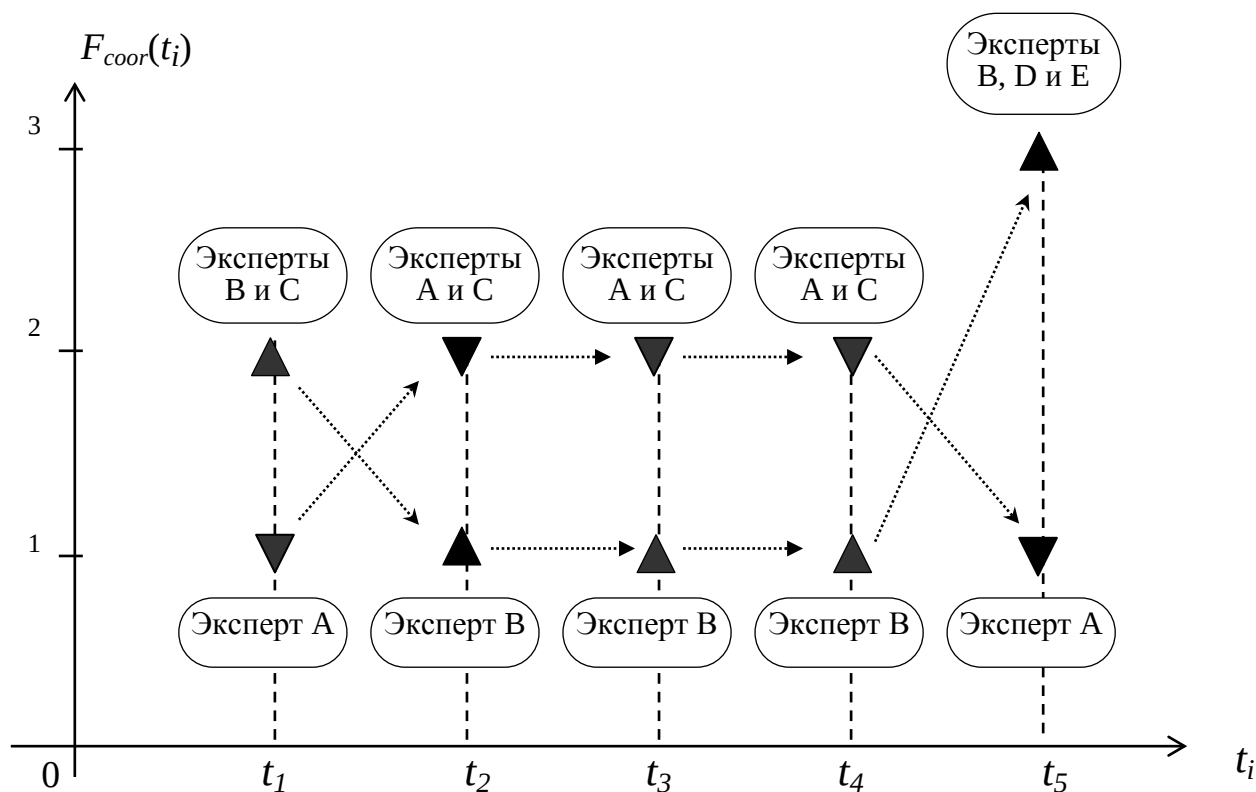


Рисунок 8 - Изменение уровней согласованности двух вариантов проектируемого индикатора в течение первых пяти итераций;

(треугольники с вершиной, направленной вниз, обозначают точки со значениями функции $F_{\text{coor}}(t_i)$ для первого варианта, а вершиной вверх – точки со значениями функции $F_{\text{coor}}(t_i)$ для второго варианта)

В настоящее время действительно есть возможность наблюдать одновременно и данные измерений, и результаты субъективного мышления, но из первых трудно выделить именно те нейроданные, которые соответствуют личностным концептам экспертов, чтобы провести такое сопоставление. Кроме того, есть гипотеза и подтверждающие её экспериментальные данные, что у экспертов существенная часть нейроинформации, соответствующей устоявшимся концептам экспертных знаний, носит структурный характер, скорее всего, на уровне связей между нейронами долговременной памяти, что не фиксируется FMRI и другими современными методами. Однако, стремительное развитие когнитивной нейронауки и её инструментальных средств позволяет надеяться, что в обозримом будущем станет возможным соотносить структурную нейроинформацию и количественные нейроданные измерений мозговой деятельности с устоявшимися и новыми концептами экспертных знаний с отражением динамики их формирования в процессе проектирования [23-26].

Отметим, что в будущем для описания соответствия нейроинформации и нейроданных с концептами с использованием методов информатики понадобится ещё одна среда, которая в работах [14, 15] была условно обозначена как нейросреда.

В нашем исследовании было достаточно трёх сред: ментальной, социально-коммуникационной и цифровой; трёх видов двуединых по своей природе сущностей: знака,

формокода и семокода; трёх категорий компьютерных кодов: для денотатов (объектный код), концептов (семантический код) и имени (информационный код) (см. рисунки 2 и 3).

В будущем ПрО информатики как информационно-компьютерной науки необходимо будет расширить за счёт нейросреды и определить ещё три вида двуединых по своей природе сущностей на границах следующих сред: ментальной и нейросреды; социально-коммуникационной и нейросреды; цифровой и нейросреды.

Расширение ПрО информатики даст возможность определить ещё одну категорию компьютерных кодов для представления нейроинформации в цифровой среде, которые условно назовём нейрокодами, и построить цифровой семиотический тетраэдр с кодами четырёх разных категорий в развитие цифрового семиотического треугольника.

Предложенная в работах [14, 15] парадигма расширяемой ПрО информатики, с одной стороны, даёт возможность более точно определить место информатики в современном процессе конвергенции отраслей науки [27], с другой стороны, поможет предвидеть появление новых поколений информационно-коммуникационных и конвергентных технологий для поддержки процессов проектирования, генерации и согласования новых экспертных знаний.

Список источников

- [1] **Зацман, И.М.** Моделирование процессов формирования экспертных знаний для мониторинга программно-целевой деятельности / И.М. Зацман, А.А. Дурново // Информатика и ее применения. – 2011. – 5(4). – С. 84–98.
- [2] **Zatsman, I.** Conceptual Framework for Development of Computer Technology Supporting Cross-Linguistic Knowledge Discovery. / I. Zatsman, N. Buntman, M. Kruzhkov, V. Nuriev, Anna A. Zaliziak // In the 15th European Conference on Knowledge Management Proceedings: Vol. 3. Reading: Academic Publishing International Limited, 2014. – P. 1063–1071.
- [3] **Zatsman, I.** Incompleteness problem for indicators system of research program / I. Zatsman, A. Durnovo // In Abstracts book of the 11th International Conference on Science and Technology Indicators. Leiden: Universiteit Leiden, 2010. – P. 309–311.
- [4] **Zatsman, I.** Proactive Dictionary of Evaluation System as a Tool for S&T Indicator Development / I. Zatsman, A. Durnovo // In É. Archambault, Y. Gingras, & V. Larivière (Eds.), Proceedings of 17th International Conference on Science and Technology Indicators: Vol. 2, 2012, Montréal: Science-Metrix and OST. – P. 905-906. http://2012.sticonference.org/Proceedings/vol2/Zatsman_Proactive_905.pdf.
- [5] **Зацман, И.М.** Семиотическая модель взаимосвязей концептов, информационных объектов и компьютерных кодов / И.М. Зацман // Информатика и ее применение. – 2009. – 3(2). – С. 65–81.
- [6] **Зацман, И.М.** Нестационарная семиотическая модель компьютерного кодирования концептов, информационных объектов и денотатов / И.М. Зацман // Информатика и ее применение. – 2009. – 3(4). – С. 87–101.
- [7] **Боргест, Н.М.** Ключевые термины онтологии проектирования: обзор, анализ, обобщения / Н.М. Боргест // Онтология проектирования. – 2013. – №3(9). – С. 9-31.
- [8] **Bryant, A.** Cognitive Informatics, Distributed Representation and Embodiment / A. Bryant // Brain and Mind. – 2003. – 4(2). – P. 215-228.
- [9] **Зацман, И.М.** Предпосылки и факторы конвергенции информационной и компьютерной наук / И.М. Зацман, О.С. Кожунова // Информатика и ее применения. – 2008. – 2(1). – С. 77–98.
- [10] **Зацман, И.М.** Построение системы терминов информационно-компьютерной науки: проблемно-ориентированный подход / И.М. Зацман // Теория и практика общественной научной информации: Сб. науч. тр. Вып. 21. М.: ИНИОН РАН, 2013. – С. 120-159.
- [11] **Eco, U.** A Theory of Semiotics / U. Eco. - Bloomington: Indiana University Press. 1976. – 59 p.
- [12] **McArthur, D.** Information, its forms and functions: The elements of semiology / D. McArthur - Lewinton: The Edwin Mellen Press, Ltd., 1997. – 228 p.
- [13] **Gorn, S.** Informatics (computer and information science): its ideology, methodology, and sociology / S. Gorn // In: The studies of information: Interdisciplinary messages / Ed. by F. Machlup and U. Mansfield. New York: Wiley, 1983. – P. 121-140.

- [14] **Зацман, И.М.** Информационно-компьютерная наука: предпосылки становления / И.М. Зацман // НТИ. Серия 1. – 2013. – (7). – С. 1-12.
- [15] **Зацман, И.М.** Информационно-компьютерная наука: технологические предпосылки становления / И.М. Зацман // Информационные технологии. – 2014. – (3). – С. 3-12.
- [16] **Зацман, И.М.** Концептуальный поиск и качество информации / И.М. Зацман. – М.: Наука, 2003. – 272 с.
- [17] **Nonaka, I.** The knowledge-creating company / I. Nonaka // Harvard Business Review. – 1991. – 69(6). – P. 96-104.
- [18] **Nonaka, I.** The Knowledge-Creating Company / I. Nonaka, H. Takeuchi. – New York: Oxford University Press. 1995. – 284 p.
- [19] Knowledge emergence / I. Nonaka, T. Nishiguchi, (Eds). – New York: Oxford University Press. 2001.
- [20] **Wierzbicki, A.** Basic dimensions of creative space / A. Wierzbicki, Y. Nakamori // In A. Wierzbicki & Y. Nakamori (Eds.). Creative Space: Models of Creative Processes for Knowledge Civilization Age. – Heidelberg: Springer. 2006. – P. 59-90.
- [21] **Wierzbicki, A.** Knowledge sciences: some new developments / A. Wierzbicki, Y. Nakamori // Zeitschrift für Betriebswirtschaft, 77(3), 2007 – pp. 271-295.
- [22] **Зацман, И.М.** Семантический словарь системы информационного мониторинга в сфере науки: задачи и функции / И.М. Зацман, О.С. Кожунова // Системы и средства информатики. – 2007. – 17(1). – С. 124-141.
- [23] **Секерина, И.А.** Метод вызванных потенциалов мозга в экспериментальной психолингвистике / И.А. Секерина // Вопросы языкознания. – 2006. – (3). – С. 22-45.
- [24] **de Charms, R.C.** Applications of real-time FMRI. / R.C. de Charms // Nature Reviews Neuroscience. – 2008. – 9(9). – P. 720–729.
- [25] **Baars, B.** Cognition, Brain, and Consciousness: Introduction to Cognitive Neuroscience / B. Baars, N. Gage. – Burlington, MA: Academic Press/Elsevier, 2010. – 677 p. – [http://www.hse.ru/data/2013/12/21/1338659679/Baars Gage 2010 Cognition, Brain and Consciousness \(2nd edition\).pdf](http://www.hse.ru/data/2013/12/21/1338659679/Baars%20Gage%20Cognition,%20Brain%20and%20Consciousness%20(2nd%20edition).pdf).
- [26] **Kumaran, D.** Tracking the emergence of conceptual knowledge during human decision-making / D. Kumaran, J.J. Summereld, D. Hassabis, E.A. Maguire // Neuron. – 2009. – 63(6). – P. 889–901.
- [27] NRC (National Research Council). Convergence: Facilitating Transdisciplinary Integration of Life Sciences, Physical Sciences, Engineering, and Beyond. Washington, DC: The National Academies Press, 2014. – 152 p. – http://www.nap.edu/catalog.php?record_id=18722.

DESIGNING INDICATORS FOR MONITORING IN SCIENCE: THEORETICAL FRAMEWORK AND MODELS

I. M. Zatsman¹, P.S. Buntman²

¹*Institute of Informatics Problems of the RAS, Moscow, Russia*
IZatsman@yandex.ru

²*Independent scholar, Moscow, Russia*
pbuntman@gmail.com

Abstract

The semiotic approach to the description of successive iterations of new indicators' designing for monitoring, including verified monitoring, and generating new expert knowledge in the course of designing is considered. For the description of iterations and dynamics of generation processes the Frege's space is used. Outcomes for successive iterations of new indicators' designing are fixed in an expert knowledge base, called by the proactive dictionary, which is a component of a monitoring system for evaluating goal-oriented activities. By means of this dictionary two main functions of the monitoring system can be implemented. It allows to describe each iteration, represent results of designing different variants for new indicators in a graphic form, compare and estimate these variants. When using the designed and approved indicators the dictionary provides possibility of verifying indicator values taking into account dynamics of adding information resources of the monitoring system.

Key words: expert knowledge base, knowledge creation, knowledge representation, verified monitoring, evaluating goal-oriented activities in science, semiotic approach.

References

- [1] **Zatsman, I.M.** Modelirovanie protsessov formirovaniya ehkspertnykh znanij dlya monitoringa programmno-tselevoj deyatel'nosti [Modeling of formation of expert knowledge for monitoring target-oriented activities] / I.M. Zatsman, A.A. Durnovo // *Informatika i ee primeneniya*, – 2011. – 5(4). – P. 84–98. (In Russian).
- [2] **Zatsman, I.** Conceptual Framework for Development of Computer Technology Supporting Cross-Linguistic Knowledge Discovery. / I. Zatsman, N. Buntman, M. Kruzhkov, V. Nuriev, Anna A. Zalizniak // In the 15th European Conference on Knowledge Management Proceedings: Vol. 3. Reading: Academic Publishing International Limited, 2014. – P. 1063–1071.
- [3] **Zatsman, I.** Incompleteness problem for indicators system of research program / I. Zatsman, A. Durnovo // In Abstracts book of the 11th International Conference on Science and Technology Indicators. Leiden: Universiteit Leiden, 2010. – P. 309–311.
- [4] **Zatsman, I.** Proactive Dictionary of Evaluation System as a Tool for S&T Indicator Development / I. Zatsman, A. Durnovo // In É. Archambault, Y. Gingras, & V. Larivière (Eds.), Proceedings of 17th International Conference on Science and Technology Indicators: Vol. 2, 2012, Montréal: Science-Metrix and OST. – P. 905-906. http://2012.sticonference.org/Proceedings/vol2/Zatsman_Proactive_905.pdf.
- [5] **Zatsman, I.M.** Semioticheskaya model' vzaimosvyazey kontseptov, informatsionnykh ob"ektov i komp'yuternykh kodov [Semiotic model the relationships of concepts, information objects and computer codes] / I.M. Zatsman // *Informatika i ee primeneniye*– 2009. – 3(2). – P. 65–81. (In Russian).
- [6] **Zatsman, I.M.** Nestatsionarnaya semioticheskaya model' komp'yuternogo kodirovaniya kontseptov, informatsionnykh ob"ektov i denotatov [Nonstationary semiotic model of computer coding concepts, information objects and denotations] / I.M. Zatsman // *Informatika i ee primeneniye*– 2009. – 3(4). – P. 87–101. (In Russian).
- [7] **Borgest, N.M.** Klyuchevye terminy ontologii proektirovaniya: obzor, analiz, obobshheniya [Key terms of the ontology of designing: review, analysis, synthesis] / N.M. Borgest // *Ontology of designing*. – 2013. – №3(9). – P. 9-31. (In Russian).
- [8] **Bryant, A.** Cognitive Informatics, Distributed Representation and Embodiment / A. Bryant // *Brain and Mind*. – 2003. – 4(2). – P. 215-228.
- [9] **Zatsman, I.M.** Predposylki i faktory konvergentsii informatsionnoj i komp'yuternoj nauk [Preconditions and factors of convergence of information and computer sciences] / I.M. Zatsman, O.S. Kozhunova // *Informatika i ee primeneniye*– 2008. – 2(1). – P. 77–98. (In Russian).
- [10] **Zatsman, I.M.** Postroenie sistemy terminov informatsionno-komp'yuternoj nauki: problemno-orientirovannyj podkhod. [Building a system of terms of information and computer science: a problem-oriented approach.] / I.M. Zatsman // *Teoriya i praktika obshhestvennoj nauchnoj informatsii: proceedings issue No. 21* Moscow: INION RAS, 2013. – P. 120-159. (In Russian).
- [11] **Eco, U.** A Theory of Semiotics / U. Eco. – Bloomington: Indiana University Press. 1976. – 59 p.
- [12] **McArthur, D.** Information, its forms and functions: The elements of semiology / D. McArthur - Lewinton: The Edwin Mellen Press, Ltd., 1997. – 228 p.
- [13] **Gorn, S.** Informatics (computer and information science): its ideology, methodology, and sociology / S. Gorn // In: The studies of information: Interdisciplinary messages / Ed. by F. Machlup and U. Mansfield. New York: Wiley, 1983. – P. 121-140.
- [14] **Zatsman, I.M.** Informatsionno-komp'yuternaya nauka: predposylki stanovleniya [Information and computer science: the preconditions of formation] / I.M. Zatsman // *NTI. Series 1.*– 2013. – (7). – P. 1-12. (In Russian).
- [15] **Zatsman, I.M.** Informatsionno-komp'yuternaya nauka: tehicheskie predposylki stanovleniya [Information and computer science: technical preconditions of formation] / I.M. Zatsman // *Informatsionnieologii*. – 2014. – (3). – P. 3–12. (In Russian).
- [16] **Zatsman, I.M.** Kontseptual'nyj poisk i kachestvo informatsii [Conceptual search and quality of information] / I.M. Zatsman. – Moscow: Science, 2003. – 272 p. (In Russian).
- [17] **Nonaka, I.** The knowledge-creating company / I. Nonaka // *Harvard Business Review*. – 1991. – 69(6). – P. 96-104.
- [18] **Nonaka, I.** The Knowledge-Creating Company / I. Nonaka, H. Takeuchi. – New York: Oxford University Press. 1995. – 284 p.
- [19] *Knowledge emergence* / I. Nonaka, T. Nishiguchi, (Eds.). – New York: Oxford University Press. 2001.
- [20] **Wierzbicki, A.** Basic dimensions of creative space / A. Wierzbicki, Y. Nakamori // In A. Wierzbicki & Y. Nakamori (Eds.). *Creative Space: Models of Creative Processes for Knowledge Civilization Age*. – Heidelberg: Springer. 2006. – P. 59-90.
- [21] **Wierzbicki, A.** Knowledge sciences: some new developments / A. Wierzbicki, Y. Nakamori // *Zeitschrift für Betriebswirtschaft*, 77(3), 2007 – pp. 271-295.

- [22] **Zatsman, I.M.** Semanticheskij slovar' sistemy informatsionnogo monitoringa v sfere nauki: zadachi i funktsii [Semantic dictionary information monitoring systems in the field of science: the tasks and functions] / I.M. Zatsman, O.S. Kozhunova // *Sistemy i sredstva informatiki*– 2007. - 17(1). - P. 124-141. (In Russian).
- [23] **Sekerina, I.A.** Metod vyzvannykh potentsialov mozga v ehksperimental'noj psikholingvistike [The method of evoked potentials in experimental psycholinguistics] / Problems of Linguistics] / I.A. Sekerina // *Voprosy yazykoznaniya*. – 2006. – (3). - P. 22-45. (In Russian).
- [24] **de Charms, R.C.** Applications of real-time fMRI. / R.C. de Charms // *Nature Reviews Neuroscience*. – 2008. - 9(9). – P. 720–729.
- [25] **Baars, B.** Cognition, Brain, and Consciousness: Introduction to Cognitive Neuroscience / B. Baars, N. Gage. - Burlington, MA: Academic Press/Elsevier, 2010. – 677 p. – [http://www.hse.ru/data/2013/12/21/1338659679/Baars Gage 2010 Cognition, Brain and Consciousness \(2nd edition\).pdf](http://www.hse.ru/data/2013/12/21/1338659679/Baars%20Cognition,%20Brain%20and%20Consciousness%20(2nd%20edition).pdf).
- [26] **Kumaran, D.** Tracking the emergence of conceptual knowledge during human decision-making / D. Kumaran, J.J. Summereld, D. Hassabis, E.A. Maguire // *Neuron*. – 2009. - 63(6). – P. 889–901.
- [27] NRC (National Research Council). *Convergence: Facilitating Transdisciplinary Integration of Life Sciences, Physical Sciences, Engineering, and Beyond*. Washington, DC: The National Academies Press, 2014. – 152 p. - http://www.nap.edu/catalog.php?record_id=18722.

Сведения об авторах



Зацман Игорь Моисеевич, 1952 г. рождения, защитил кандидатскую диссертацию в 1986 году. Заведующий отделом Института проблем информатики РАН. В 2011 году защитил докторскую диссертацию по специальности «Теоретические основы информатики»; в 2012 году присвоена степень доктора наук. Область научных интересов: информатика как информационно-компьютерная наука, когнитивная информатика, моделирование процессов генерации знаний и их представление в цифровой среде, верифицируемый мониторинг и оценивание программной деятельности в сфере науки.

Igor Moiseevich Zatsman (b. 1952) has the PhD in informatics (1986). Currently, he is the head of research department at the Institute of Informatics Problems of the Russian Academy of Sciences. He has the highest research diploma, obtained after the PhD (2012). Research interests are in the fields of informatics as computer and information science, cognitive informatics, modeling emerging meaning processes and their tracing by computer, verified monitoring and goal-oriented assessment in science.



Бунтман Павел Сергеевич, 1987 г. рождения, окончил биологический факультет Московского государственного университета. В Институте биоорганической химии РАН написал дипломную работу по методам молекулярной биологии. Работал переводчиком научно-популярной литературы. Область научных интересов: информатика как информационно-компьютерная наука и когнитивная информатика.

Pavel Sergeevich Buntman (b. 1987) graduated from the biological faculty of the Moscow State University. At the Institute of Bioorganic Chemistry he has written a thesis on molecular biology methods and after that he worked as a translator for publishing houses. Research interests are in the fields of informatics as computer and information science, cognitive informatics.

УДК 519.711.3

ПОВЫШЕНИЕ РЕЛЕВАНТНОСТИ ПРИ ПОИСКЕ ВЕБ-СЕРВИСОВ В МУЛЬТИОНТОЛОГИЧЕСКОЙ СРЕДЕ

Л.С. Глоба¹, М.В. Ковальский²

Национальный технический университет Украины «Киевский политехнический институт», Киев, Украина

¹lgloba@its.kpi.ua, ²mkovalskyi@luxoft.com

Аннотация

Кратко рассмотрены существующие подходы к решению задачи поиска веб-сервисов. Выявлены проблемы низкой релевантности, интероперабельности, существенного влияния человеческого фактора в процессе поиска веб-сервисов, которые решаются добавлением семантической информации к их описанию. Проанализированы существующие модели аннотированных онтологий веб-сервисов, а также предложена их модификация, позволяющая получать более релевантный ответ на запрос пользователя. Предложена концепция промежуточного абстрактного объекта - «шаблона сервиса», в который преобразовываются пользовательские запросы с одной стороны, а с другой — онтологические модели независимо от языка их описания. С целью повышения релевантности ответа предлагается также использовать входные и выходные параметры операций веб-сервисов, сравнивая их как синтаксически, так и семантически. Согласно полученным результатам, релевантность ответа на поисковый запрос увеличилась, а вероятность возникновения «ложного обнаружения» уменьшилась.

Ключевые слова: веб-сервисы, онтология, релевантность.

Введение

Веб-сервисы предоставляют единообразные программные интерфейсы, позволяющие их использовать конечными пользователями при необходимости получения информационных услуг в среде Интернет [1].

В текущий момент стоит задача нахождения такой модели построения распределенных систем на базе веб-сервисов, которая бы позволяла автоматизированно обнаруживать веб-сервисы согласно заданным критериям с наибольшей точностью (релевантностью) [2].

Релевантность поисковых запросов¹ является серьезной проблемой в набирающем всё большие объёмы данных Интернете. Например, реестр UDDI (Universal Description, Discovery and Integration) позволяет производить поиск веб-сервиса исключительно с помощью ключевых слов и предопределенных категорий [3]. Таким образом, окончательный выбор веб-сервиса производится пользователем.

В ряде работ предложено множество решений в этом направлении, в частности — семантические веб-сервисы, в основе которых лежит дополнение существующих WSDL-описаний (Web Services Description Language) посредством аннотирования онтологическими моделями, разрабатываемыми независимо от описаний веб-сервисов [4]. Однако они характеризуются низкой релевантностью в процессе поиска веб-сервисов, что осложняет их использование конечными пользователями при необходимости получения услуг в среде Интернет.

¹ **Релевантность** (лат. *relevo* — поднимать, облегчать) - семантическое соответствие поискового запроса и поискового образа документа. В более общем смысле, близко к понятию «адекватность», при этом оценивается не только степень соответствия, но и степень практической применимости результата поиска. *Прим.ред.*

1 Анализ существующих подходов

Де-факто стандартом обнаружения веб-сервисов является реестр UDDI, который предоставляет категоризированный поиск, а также поиск по ключевым словам. Множество исследователей считает, что такой подход не является масштабируемым, поскольку требует полной репликации всех опубликованных в реестре веб-сервисов [4].

Для решения этой проблемы были предложены варианты архитектуры «peer-to-peer» [5], которые включают в себя решения распределенного хранения каталога веб-сервисов с использованием полной или частичной репликации на множество серверов.

В [6] рассмотрена возможность применения таксономий процессов, описывающих не сущности, а действия, применимые к этим сущностям. Такой метод сопоставления использует семантические отношения между процессами для сравнения пользовательских запросов и моделей, описывающих веб-сервисы.

Работа [7] использует принцип подстановки Барбары Лисков² относительно концептов онтологической модели. Любой концепт может быть заменен своим дочерним концептом, если последний предоставляет больший или равный интерфейс, который имеет избыточность в некотором контексте. Таким образом, более детализированные онтологические модели можно адаптировать для сравнения с менее детализированными. Недостатком этого метода является то, что предусматривается аннотирование веб-сервисов онтологиями исключительно средствами DAML (DARPA Agent Markup Language), который не поддерживает описания входных и выходных данных. Анализ полученных описаний возможен только по описаниям и названиям веб-сервисов, что является естественным ограничением DAML. Поэтому значительно возрастает возможность ложного обнаружения.

Предложенный в [8] метод состоит в преобразовании веб-сервисов в вид RDF-графов (Resource Description Framework). Далее полученные графы сравниваются, при этом проводится декомпозиция узлов, представляющих сложные типы, к более простым типам.

Также был представлен способ сопоставления онтологических моделей средствами специально разработанного языка LARKS (Language for Advertisement and Request for Knowledge Sharing) [9]. Ядро системы, описанной в указанной работе, состоит из пяти так называемых «фильтров» - критериев поиска. Такая концепция «фильтров» основывается не на сравнении отдельных компонентов, а на аппроксимации онтологической модели с целью получения множества ключевых слов, которые определяют содержание модели. Далее это множество ключевых слов сопоставляется с пользовательским запросом. LARKS разработан с учётом особенностей мультионтологической среды, но при этом считается, что любая онтологическая модель, подлежащая сравнению, использует некоторый единый глоссарий или генерализированную онтологическую модель.

В [10] была дана классификация различий, которые возможны между онтологическими моделями в связи с использованием различных языков моделирования и намерений, с которыми некоторая модель создавалась. Это же исследование показывает, что применение статистических и лингвистических алгоритмов для сравнения свойств концептов представляется обоснованным.

Существуют также решения из области машинного обучения, например, LSD (Learning Source Descriptions) [10], которые основаны на вероятностном распределении экземпляров объектов внутри разнообразных онтологических моделей.

² Принцип подстановки Барбары Лисков (*Liskov Substitution Principle, LSP*) в объектно-ориентированном программировании является специфичным определением *подтипа*, предложенным Барбарой Лисков в 1987 году. Этот принцип служит критерием оценки качества принимаемых решений при построении иерархий наследования и фактически означает, что поведение наследуемых классов не должно противоречить поведению, заданному базовым классом. *Примеч. ред.*

На практике, несмотря на значительное количество исследований в данном направлении, всё же остаётся существенное влияние человеческого фактора в процессе поиска веб-сервисов, а также низкая релевантность результатов поиска, что вынуждает конечного пользователя выполнять ручной подбор необходимого ему сервиса.

Метод поиска аннотированных онтологическими моделями веб-сервисов, предложенный в [4], основан на оценке синтаксического подобия названий и описаний концептов, представляющих сами веб-сервисы. Концепты, описывающие искомые веб-сервисы, и ключевые слова из пользовательского запроса сравниваются с помощью синтаксического подобия, использующего различные алгоритмы сравнения строк (нормализация текста, стемминг³, алгоритмы с использованием N-грамм). Этот метод показывает хорошие результаты для моноонтологической среды, в которой используется единая онтология. Основываясь на результатах проведённого исследования утверждается, что релевантность ответа на пользовательский запрос значительно уменьшается в мультионтологической среде [4].

При этом для мультионтологической среды, использующей различные онтологические модели, разработанные разрозненно, метод, описанный в работе [4], не подходит. Анализ рассмотренных подходов позволяет сделать вывод о недостаточной эффективности вышеописанных методов. Однако, метод поиска аннотированных онтологическими моделями веб-сервисов, изложенный в [4], показывает наибольшую эффективность в случае поиска веб-сервисов в моноонтологической среде, поэтому его доработка с целью использования в мультионтологической среде является наиболее целесообразной.

Для использования метода поиска аннотированных онтологическими моделями веб-сервисов в мультионтологической среде предложены следующие модификации.

Во-первых, не использовать конкретный язык описания, в данном случае — DAML. Несмотря на то, что используемый алгоритм не зависит от конкретной реализации онтологической модели, он зависит от языка описания. С целью независимости от языка описания онтологических моделей предлагается подход с использованием программного промежуточного абстрактного объекта, который является отображением онтологической модели и называется «шаблоном сервиса». В такие объекты преобразовываются пользовательские запросы с одной стороны, а с другой — онтологические модели независимо от языка их описания.

Во-вторых, в работе [4] оценка подобия веб-сервисов производилась исключительно с помощью названий и описаний концептов, описывающих веб-сервис. С целью повышения релевантности ответа предлагается использовать также входные и выходные параметры операций, сравнивая их как синтаксически, так и семантически, например, учитывая кардинальность множества принимаемых параметром значений.

2 Метод оценки подобия веб-сервиса искомому

Подход к поиску веб-сервиса в мультионтологической среде сводится к трём последовательным шагам:

- 1) создание шаблона сервиса на основе пользовательского запроса;
- 2) сравнение шаблона сервиса с множеством веб-сервисов, которые были идентифицированы как вероятные кандидаты;
- 3) возврат веб-сервисов, удовлетворяющих минимально допустимой оценке подобия, пользователю в виде упорядоченного списка.

Семантический шаблон сервиса описывает пользовательский запрос. Он позволяет задавать набор требуемых пользователю операций, их свойств, входных и выходных данных. У

³ Stemming (англ.) - морфологический поиск, т.е. поиск слова во всех его морфологических формах. *Прим.ред.*

шаблона сервиса нет конкретной реализации, поскольку его следует рассматривать как промежуточную абстракцию – веб-сервис-посредник (proxy web service).

Более формально шаблон сервиса (ST, Service Template) можно определить следующим образом:

$$ST = \langle N_{ST}, D_{ST}, OP_{ST} \langle N_{OP}, D_{OP}, O_{OP}, I_{OP} \rangle \rangle,$$

где N_{ST} – название веб-сервиса, D_{ST} – текстовое описание веб-сервиса, OP_{ST} – множество операций веб-сервиса.

Каждая из операций веб-сервиса, в свою очередь, определяется с помощью N_{OP} – названия операции, D_{OP} – текстового описания операции, O_{OP} и I_{OP} – выходные и входные данные операции.

Шаблон сервиса сравнивается с множеством сервис-кандидатов (CS, Candidate Service) – шаблонов сервиса, полученных при анализе веб-сервисов из предопределённого множества, подлежащего перебору. Сравнение состоит в оценке синтаксического и семантического подобия:

$$\Theta(ST, CS) = [w_H \cdot H(ST, CS) + w_\Phi \cdot \Phi(ST, CS)] / [w_H + w_\Phi],$$

где $\Theta(ST, CS)$ – суммарная оценка, $H(ST, CS)$ – синтаксическое подобие, $\Phi(ST, CS)$ – семантическое (функциональное) подобие, w_i представляет весовой коэффициент соответствующей оценки подобия, предназначенный для более гибкого управления критериями сравнения.

Синтаксическое подобие вычисляется аналогично базовому методу [4] с использованием стемминга, т.е. процесса нахождения основы слова, представляющего описание или название, и дальнейшее вычисление расстояния Хемминга – числа позиций, в которых символы двух слов различны.

Алгоритм, используемый для нахождения основы слова, заключается в использовании таблицы правил для преобразования входного слова в нормализованную форму. Таблица правил определяет префиксы и суффиксы, которые необходимо удалять из слов или заменять на другие префиксы или суффиксы. Такой подход оправдан, поскольку показывает хорошее быстроедействие и, в то же время, достаточно прост. Например, алгоритм, использующий таблицу, определяющую правило полного соответствия слова и его основы, является менее производительным, хотя и обеспечивает большую эффективность для языков, имеющих сложные правила образования префиксов или суффиксов [11].

Семантическое подобие вычисляется с использованием множества критериев, которые учитывают оценку подобия двух веб-сервисов с использованием атрибутов и отношений концептов – подобие по входным и выходным данным операций веб-сервисов, концептуальное подобие и его составляющая – подобие по свойствам концептов.

Концептуальное подобие веб-сервисов $C(ST, CS)$ определяется как:

$$C(ST, CS) = [w_H \cdot H(ST, CS) + w_P \cdot P(ST, CS)] / [w_H + w_P],$$

где $C(ST, CS)$ – концептуальное подобие, $H(ST, CS)$ – синтаксическое подобие, $P(ST, CS)$ – подобие по свойствам концептов, w_i – весовые коэффициенты соответствующих оценок подобия.

Концептуальное подобие используется также при оценке подобия по входным и выходным данным $IO(OP_{ST}, OP_{CS})$:

$$IO(OP_{ST}, OP_{CS}) = (I(OP_{ST}, OP_{CS}) \cdot O(OP_{ST}, OP_{CS}))^{1/2},$$

где $IO(OP_{ST}, OP_{CS})$ – подобие по входным и выходным данным операций, $I(OP_{ST}, OP_{CS})$ – подобие по входным данным операций, $O(OP_{ST}, OP_{CS})$ – подобие по выходным данным операций, OP_{ST} – множество операций шаблона сервиса, OP_{CS} – множество операций сервиса-кандидата.

Оценка подобия по входным данным определяется как:

$$I(OP_{ST}, OP_{CS}) = \max[I(OP_{ST} - OPi_{ST}, OP_{CS} - OPi_{CS}) + C(OPi_{ST}, OPi_{CS})],$$

где $I(OP_{ST}, OP_{CS})$ – подобие по входным данным, OP_{ST} и OP_{CS} – множество операций шаблона сервиса и сервиса-кандидата соответственно, OPi_{ST} и OPi_{CS} – элементы множества операций шаблона сервиса и сервиса-кандидата, $C(OPi_{ST}, OPi_{CS})$ – концептуальное подобие элементов множества.

Оценка подобия по выходным данным операций производится аналогичным образом.

На рисунке 1 схематически представлена иерархия используемых критериев.

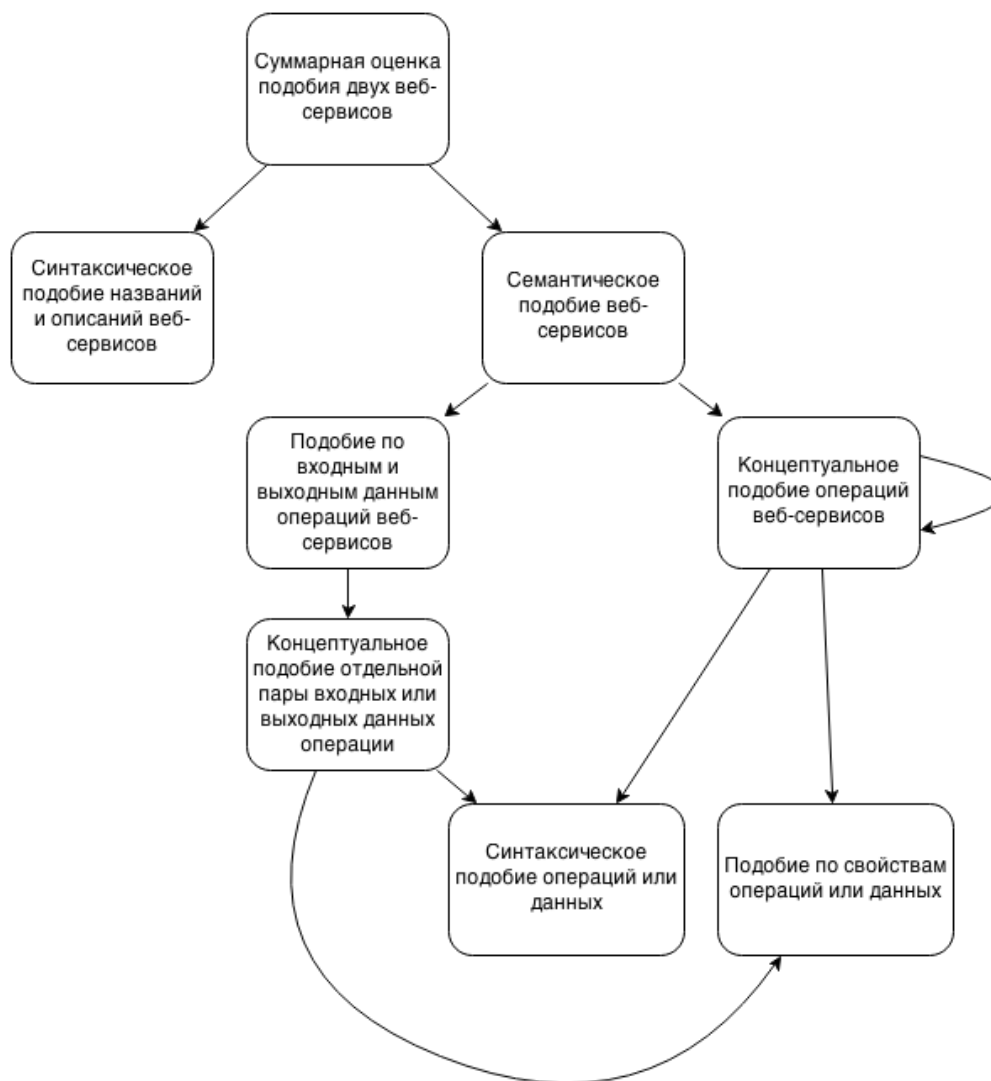


Рисунок 1 – Схематическое представление иерархии используемых критериев

Следует отметить, что концептуальное подобие операций веб-сервисов определено как задача динамического программирования, т.е. рекурсивная функция, поскольку включает в себя перебор всех возможных комбинаций операций с целью нахождения такого их отображения, которое даёт максимальную оценку подобия по этому критерию.

Наибольшее влияние на оценку семантического подобия имеет её составляющая – подобие по свойствам концептов, которыми описаны операции веб-сервиса.

Вычисление оценки подобия по свойствам концептов заключается в поочередном сравнении пар свойств двух выбранных для сравнения концептов. Во время сравнения этих свойств формируются их отображения, показывающие эквивалентность свойств между концептами. При этом следует обеспечить максимально возможную среднюю оценку подобия отображаемых свойств:

$$P(C_{ST}, C_{CS}) = \max (P(CP_{ST} - CPi_{ST}, CP_{CS} - CPi_{CS}) + p(CPi_{ST}, CPi_{CS})),$$

где $P(C_{ST}, C_{CS})$ – оценка подобия по свойствам, CP_{ST} и CP_{CS} – множество свойств концепта, описывающего операцию шаблона сервиса и сервиса-кандидата соответственно, CPi_{ST} и CPi_{CS} – отдельные свойства концепта, описывающего операцию шаблона сервиса и сервиса-кандидата соответственно. Подобие отдельных свойств $p(CPi_{ST}, CPi_{CS})$ описывается следующим образом:

$$p = k \cdot (H \cdot K \cdot \chi)^{1/3} - 0.05 \cdot \|\text{свойства, не имеющие отображения}\|,$$

где H – синтаксическое подобие, K – подобие по кардинальности, χ – подобие по ограничениям, применяемым к типу.

Множитель k является константой, значение которой составляет 1 в случае, если оба сравниваемых свойства являются инверсными функциональными свойствами, и 0,8 – в противном случае. Инверсными функциональными называют такие свойства, которые имеют значения, уникальные для каждого экземпляра объекта. Иными словами, это свойства концепта, которые определяют его идентичность в пределах предметной области.

Подобие по ограничениям вычисляется с учётом типа данных, представляющего свойство. Свойство концепта может быть представлено с помощью примитивного типа или другого концепта. Исходя из этого, возможны три случая:

- 1) Оба сравниваемых свойства представлены примитивным типом данных. Оценка подобия по ограничениям сводится к вычислению количества информации, теряемой при преобразовании из одного типа в другой. На практике используются эвристические оценки в пределах от 0 (полная потеря информации) до 1 (полное сохранение информации).
- 2) Оба сравниваемых свойства являются онтологическими концептами. Задача оценки подобия по свойствам двух концептов сводится к рекурсивному вычислению оценки и является задачей динамического программирования.
- 3) Одно из свойств представлено примитивным типом, а другое – концептом. В этом случае оценка равна нулю, поскольку типы несовместимы. Такая оценка может быть нежелательной, поскольку данный случай может возникнуть в связи с различным уровнем детализации двух сравниваемых концептов, тогда нужно использовать некоторые оценки по умолчанию или же сравнивать такие свойства с помощью синтаксического подобия.

Подобие по кардинальности также играет значительную роль при сравнении двух свойств. Оценка подобия по кардинальности предусматривает подсчёт кардинальности множества принимаемых свойством значений. Кардинальность множества принимаемых свойством значений может быть как конечной малой величиной (в случае с типом перечисления), так и условно конечной величиной (например, вещественные числа, количество возможных значений которых ограничено только конкретной реализацией окружения).

Оценка подобия по кардинальности также является эмпирической и равняется:

- 1) единице, если кардинальности множества принимаемых значений двух свойств равны;
- 2) единице, если оба свойства являются инверсными функциональными;
- 3) 0,9, если кардинальность свойства, представляющего шаблон сервиса, меньше;
- 4) 0,7, если кардинальность свойства, представляющего шаблон сервиса, больше.

В случае, если количество свойств концепта сервиса-кандидата не меньше количества свойств концепта шаблона сервиса, представляется возможным отображение «один к одно-

му» каждого свойства шаблона. В противном случае, когда отображение имеет вид «один ко многим», следует вводить штраф, уменьшающий оценку подобия по свойствам пропорционально количеству свойств, имеющих несколько отображений. С помощью моделирования значение штрафа было установлено на уровне 0,05.

В конечном итоге, в результате сравнения каждый сервис-кандидат получает нормированную в пределах $[0, 1]$ оценку подобия. Считаем, что оценка со значением 1 соответствует наилучшему совпадению в выборке. Затем веб-сервисы, удовлетворяющие минимально допустимой оценке, возвращаются пользователю в виде упорядоченного списка.

3 Полученные результаты

Метод был проверен на наборе веб-сервисов для оценки повышения релевантности ответа на пользовательский запрос. Веб-сервисы были получены из индекса компании xIgnite и аннотированы двумя разными онтологиями. Первая была получена в результате анализа предметной области фондовых бирж при помощи материалов NASDAQ, находящихся в открытом доступе в сети Интернет. Вторая онтология была создана на основе понятий, используемых компанией xIgnite при описании собственных веб-сервисов.

Такой выбор используемых данных обусловлен тем, что они в полной мере удовлетворяют необходимым требованиям.

Во-первых, рассматриваются четыре сценария, которые могут возникнуть при сравнении двух онтологических концептов в моноонтологической (в пределах одной онтологической модели), и четыре сценария относительно мультионтологической среды (в пределах нескольких онтологических моделей). Таковыми являются сравнение (в моноонтологической среде под эквивалентным концептом подразумевается идентичный):

- 1) эквивалентных концептов;
- 2) концепта и его подконцепта (дочернего концепта);
- 3) концепта и его надконцепта (родительского концепта);
- 4) несвязанных концептов.

Набор используемых для анализа веб-сервисов включает в себя 14 различных веб-сервисов, 13 из которых содержат по две операции и один – одну операцию.

Моделирование производится в два этапа. Сначала шаблон сервиса для пользовательского запроса описывается заранее определенным (базовым) концептом из онтологической модели. Затем, для каждого из восьми вышеописанных сценариев, сервис-кандидат описывается одним из концептов, удовлетворяющих условиям сценария; а потом производится оценка.

Второй этап заключается в аннотировании операций, принадлежащих веб-сервисам из набора, и их свойств; входящих и исходящих данных при помощи ранее выбранного базового концепта. Далее производится оценка по базовому и расширенному методам.

На рисунке 2 изображены результаты первого этапа моделирования. Сценарии 1-4 соответствуют моноонтологической среде, а сценарии 5-8 – мультионтологической.

Из полученных результатов следует, что суммарная оценка при использовании синтаксического и семантического подобия ниже оценки только по синтаксическому подобию для сценариев 3, 4, 7, 8. Следовательно, учёт семантического подобия снижает суммарную оценку для несвязанных концептов и концептов, которые неспособны в полной мере покрыть свойства базового концепта. Это, в свою очередь, понижает вероятность так называемого «ложного обнаружения».

Суммарная оценка для сценария 5 тоже уменьшилась, что также связано с неполным покрытием свойств базового концепта, что является следствием различий между уровнем детализации онтологических моделей.

Анализируя сценарии 1, 2, 6, можно сделать вывод о том, что суммарная оценка повысилась для концептов, удовлетворяющих свойствам базового концепта (в данном случае – идентичные и дочерние по отношению к базовому концепту). Таким образом, при поиске веб-сервиса предпочтение будет отдано веб-сервисам, которые входят в эту категорию.

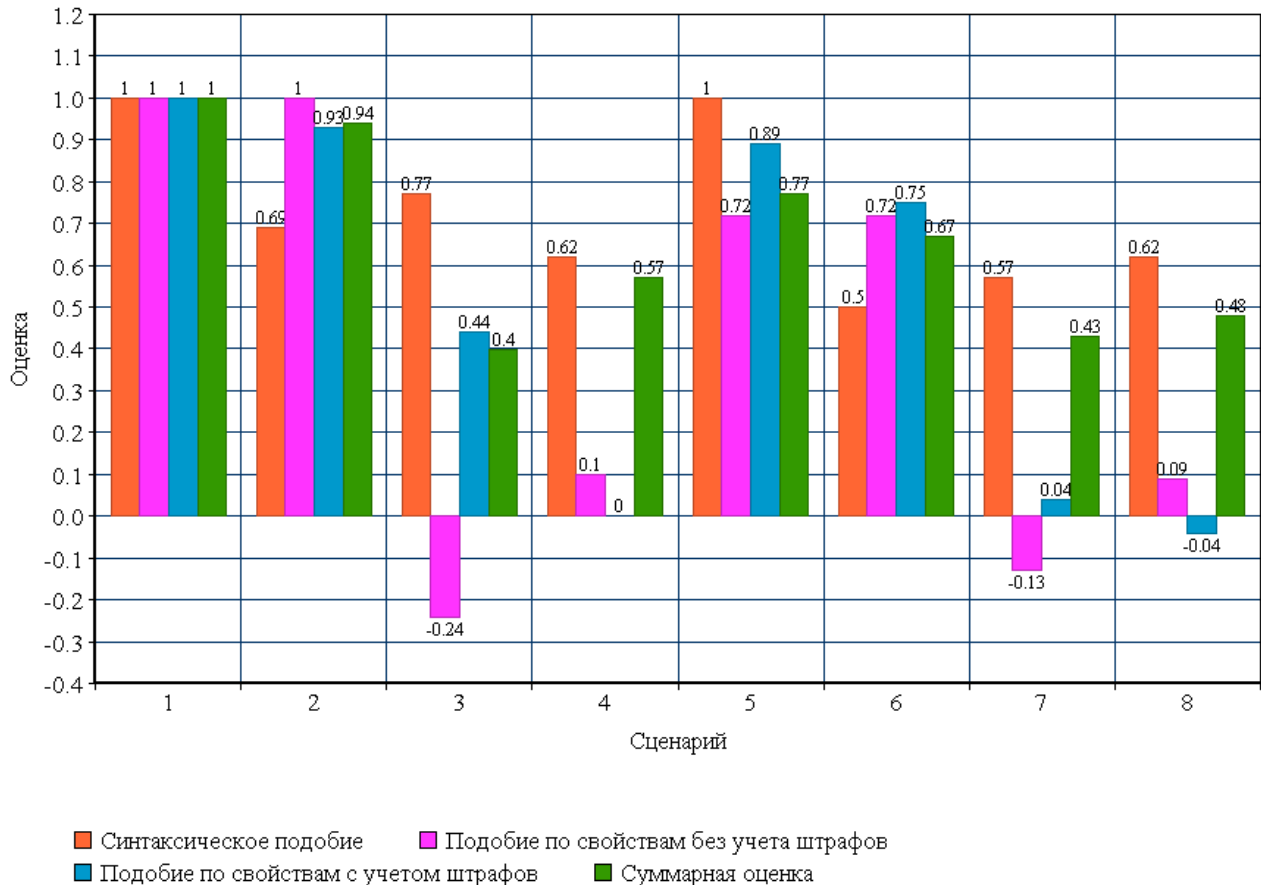


Рисунок 2 – Подобие двух онтологических концептов согласно видам оценок

На рисунке 3 показаны результаты второго этапа моделирования. Операции веб-сервисов 1-7 имеют большее количество свойств, чем может предоставить базовый концепт, при этом веб-сервис 7 предоставляет только одну операцию, в то время как шаблон-сервиса, описывающий пользовательский запрос, содержит две операции.

Согласно полученным результатам, оценка метода, учитывающего синтаксическое и семантическое подобие, оказалась ниже оценки только по синтаксическому подобию для веб-сервисов 1-7. Таким образом, веб-сервисы, которые не могут полностью покрыть требования шаблона сервиса, отсеиваются.

Задавая при поиске минимально допустимую оценку подобия значением 0,7, получаем, что при использовании предложенного метода получим пять веб-сервисов (9, 10, 12, 13 и 14), а при использовании метода, учитывающего только синтаксическое подобие, – семь (4, 6, 10-14).

Следовательно, оценка семантического подобия помогает уменьшить вероятность возникновения «ложного обнаружения» (в данном случае – веб-сервисов 4, 6, 11)⁴.

⁴ В приведенном примере авторы сочли возможным не отмечать повышение вероятности «ложного обнаружения», которая проявилась у веб-сервиса 9 (см. рисунок 3). *Прим.ред.*

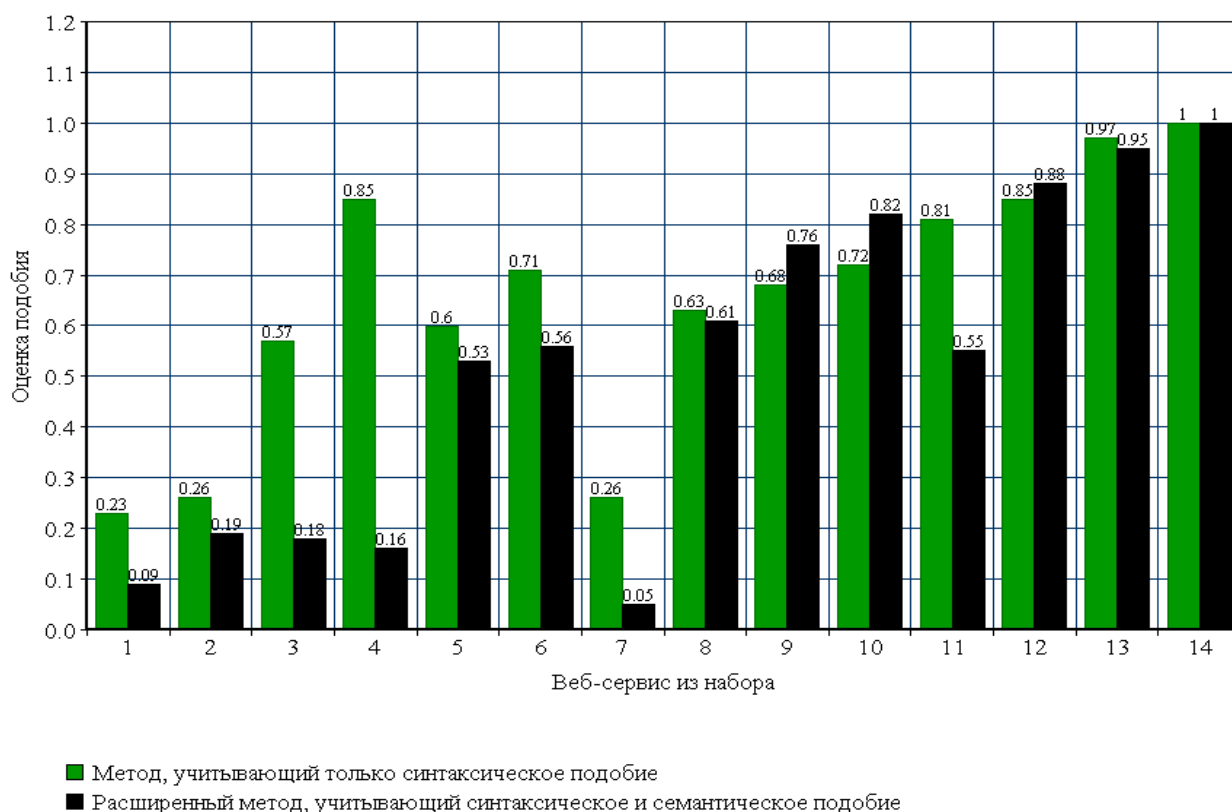


Рисунок 3 – Сравнение оценки подобия веб-сервисов из набора и пользовательского запроса

Заключение

Обнаружение веб-сервисов является исключительно семантической проблемой. Текущий стандарт веб-сервисов на основе UDDI и WSDL не использует семантическую информацию, поэтому релевантность обнаруженных пользователями веб-сервисов остается невысокой.

Предложено усовершенствование метода поиска веб-сервиса в мультионтологической среде, суть которого заключается в дополнительном использовании программного промежуточного абстрактного объекта – «шаблона сервиса», а также в использовании для сравнения сервисов дополнительно типов входных и выходных параметров операций веб-сервиса с целью увеличения релевантности ответа на запрос пользователя.

Дальнейшие исследования направлены на расширение предложенного метода за счёт поддержки описаний ошибок в терминах WSDL, которые могут возникнуть во время обработки запроса, а также требований к QoS (Quality of Service). Необходимо определить оптимальные оценки вместо эвристических для подобия по свойствам концепта с помощью экспертных систем и проверить предлагаемый метод на большем количестве реальных веб-сервисов и их онтологических описаний.

Также метод может быть расширен за счёт поддержки композиции веб-сервисов во время их обнаружения, что позволит формировать цепочку их вызовов. Это необходимо в ситуациях, при которых требуемый веб-сервис не был найден в реестре, но, комбинируя другие существующие, можно получить искомый. Предлагается проводить такую композицию с помощью онтологических описаний входных и выходных параметров операций каждого веб-сервиса в цепочке.

СПИСОК ИСТОЧНИКОВ

- [1] **Fensel, D.** Semantic web services / D. Fensel, F.M. Facca, E. Simperl, I. Toma. - Springer, 2011 - 357 p.
- [2] **Pedrinaci, C.** Semantic web services / C. Pedrinaci, J. Domingue, A.P. Sheth // In Handbook of semantic web technologies. - Springer Berlin Heidelberg, 2011 - pp. 977-1035.
- [3] **Toch, E.** Humans, semantic services and similarity: A user study of semantic Web services matching and composition / E. Toch, I. Reinhartz-Berger, D. Dori // Web Semantics: Science, Services and Agents on the World Wide Web, 9(1), 2011 - pp. 16-28.
- [4] **Sheth, A.** Changing Focus on Interoperability in Information Systems: From System, Syntax, Structure to Semantics / A. Sheth // In Interoperating Geographic Information Systems. - Kluwer Academic Publishers, 1998 - pp. 5-30.
- [5] **Schmidt, C.** A Peer-to-Peer Approach to Web Service Discovery / C. Schmidt, M. Parashar // World Wide Web, Internet and Web Information Systems. - Kluwer Academic Publishers, Vol. 7, No. 2, 2004 – pp. 211–229.
- [6] **Klein, M.** Searching for Services on the Semantic Web using Process Ontologies / M. Klein, A. Bernstein // First Semantic Web Working Symposium (SWWS-1), Stanford, CA USA, 2001.
- [7] **Gonzales-Castillo, J.** Description logics for matchmaking of services / J. Gonzales-Castillo, D. Trastour, C. Bartolini // Proc. of Workshop on Application of Description Logics, Vol. 44. ADL-2001.
- [8] **Trastour, D.** A Semantic Web Approach to Service Description for Matchmaking of Services / D. Trastour, C. Bartolini, J. Gonzalez-Castillo // Proc. 1st Semantic Web Working Symposium, CA, 2001 - 15 p.
- [9] **Sycara, K.** Dynamic Service Matchmaking among Agents in Open Information Environments / K. Sycara, J. Lu, M. Klusch, S. Widom // Journal ACM SIGMOD Record, Special Issue on Semantic Interoperability in Global Information Systems, 28(1). 1999 - pp. 47-53.
- [10] **Magnini, B.** Linguistic Based Matching of Local Ontologies / B. Magnini, L. Serafini, M. Speranza // In Working notes of MeaN-02 (Workshop held in conjunction with AAAI-2002), Edmonton, Alberta, Canada, July 28 - August 1, 2002 - pp. 42-50.
- [11] **Porter, M.F.** An Algorithm for Suffix Stripping / M.F. Porter // The journal Program, Vol. 14, No. 3, July 1980 - pp 130-137; Program: electronic library and information systems, Vol. 40 Iss: 3, pp. 211–218.

INCREASING WEB SERVICES DISCOVERY RELEVANCY IN THE MULTI-ONTOLOGICAL ENVIRONMENT

L.S. Globa¹, M.V. Kovalskyi²

National Technical University of Ukraine «Kiev Polytechnic Institute», Kiev, Ukraine

¹lgloba@its.kpi.ua, ²mkovalskyi@luxoft.com

Abstract

The existing approaches to the solution of web services discovery issues are reviewed briefly in this paper. The problems of low relevancy, interoperability, high amount of the human factor existing in the process of web services discovery are found and can be solved by addition of semantic data to web service description. The existing models of ontology annotated web services are analyzed and their modification is proposed to allow achievement of higher discovery relevancy. The concept of the intermediate abstract object called «service template», in which end user requests and ontological models are converted, is introduced. It is also proposed to use inputs and outputs of web services operations to match them by syntactic and semantic criteria. As shown in the paper, web service discovery response relevancy has increased, and probability of false matching has decreased.

Key words: web services, ontology, relevancy.

References

- [1] **Fensel, D.** Semantic web services / D. Fensel, F.M. Facca, E. Simperl, I. Toma. - Springer, 2011 - 357 p.
- [2] **Pedrinaci, C.** Semantic web services / C. Pedrinaci, J. Domingue, A.P. Sheth // In Handbook of semantic web technologies. - Springer Berlin Heidelberg, 2011 - pp. 977-1035.

- [3] **Toch, E.** Humans, semantic services and similarity: A user study of semantic Web services matching and composition / E. Toch, I. Reinhartz-Berger, D. Dori // Web Semantics: Science, Services and Agents on the World Wide Web, 9(1), 2011 - pp. 16-28.
 - [4] **Sheth, A.** Changing Focus on Interoperability in Information Systems: From System, Syntax, Structure to Semantics / A. Sheth // In Interoperating Geographic Information Systems. - Kluwer Academic Publishers, 1998 - pp. 5-30.
 - [5] **Schmidt, C.** A Peer-to-Peer Approach to Web Service Discovery / C. Schmidt, M. Parashar // World Wide Web, Internet and Web Information Systems. - Kluwer Academic Publishers, Vol. 7, No. 2, 2004 – pp. 211–229.
 - [6] **Klein, M.** Searching for Services on the Semantic Web using Process Ontologies / M. Klein, A. Bernstein // First Semantic Web Working Symposium (SWWS-1), Stanford, CA USA, 2001.
 - [7] **Gonzales-Castillo, J.** Description logics for matchmaking of services / J. Gonzales-Castillo, D. Trastour, C. Bartolini // Proc. of Workshop on Application of Description Logics, Vol. 44. ADL-2001.
 - [8] **Trastour, D.** A Semantic Web Approach to Service Description for Matchmaking of Services / D. Trastour, C. Bartolini, J. Gonzalez-Castillo // Proc. 1st Semantic Web Working Symposium, CA, 2001 - 15 p.
 - [9] **Sycara, K.** Dynamic Service Matchmaking among Agents in Open Information Environments / K. Sycara, J. Lu, M. Klusch, S. Widom // Journal ACM SIGMOD Record, Special Issue on Semantic Interoperability in Global Information Systems, 28(1). 1999 - pp. 47-53.
 - [10] **Magnini, B.** Linguistic Based Matching of Local Ontologies / B. Magnini, L. Serafini, M. Speranza // In Working notes of MeaN-02 (Workshop held in conjunction with AAAI-2002), Edmonton, Alberta, Canada, July 28 - August 1, 2002 - pp. 42-50.
 - [11] **Porter, M.F.** An Algorithm for Suffix Stripping / M.F. Porter // The journal Program, Vol. 14, No. 3, July 1980 - pp 130-137; Program: electronic library and information systems, Vol. 40 Iss: 3, pp. 211–218.
-

Сведения об авторах



Глоба Лариса Сергеевна, 1954 г. рождения. Окончила Харьковский авиационный институт в 1978 г., к.т.н. (1984), д.т.н. (1996), проф. с 2001 г., с 2012 г. – академик Международной академии информационных технологий (Минск, Белоруссия). Заведует кафедрой информационно-телекоммуникационных систем Национального технического университета Украины «Киевский политехнический институт». Сфера научных интересов: разработка интеллектуальных корпоративных систем производства и систем управления, распределенные системы, технологии Web-сервисов. В списке научных трудов около 400 работ, среди которых учебники и учебные пособия.

Larysa Sergiivna Globa (b. 1954) graduated from Kharkov Aviation Institute in 1978, Ph.D., from 1984, Doctor of Technical Sciences from 1996, Prof. from 2001, academician of the International Academy of Information Technology (Minsk, Belorussia) from 2012. The head of the chair of Information and Telecommunication Systems of National Technical University of Ukraine "Kiev Polytechnic Institute". Research interests: development of intelligent enterprise production and management systems, distributed systems, Web-services technology. She is co-author of about 400 scientific works, a significant number of works issued abroad.



Ковальский Михаил Витальевич, 1994 г. рождения. Окончил Национальный технический университет Украины «Киевский политехнический институт» в 2014 г. Сфера научных интересов – распределенные в сети Интернет системы и облачные вычисления.

Mikhail Vitalievich Kovalskyi (b.1994) graduated from the National Technical University of Ukraine "Kiev Polytechnic Institute" in 2014. Works in the area of the distributed Internet-systems and cloud computing.

УДК 004.85

ОНТОЛОГИЧЕСКИЕ МОДЕЛИ СИТУАЦИЙ В ЗАДАЧАХ КОМПЬЮТЕРНОГО КОНТРОЛЯ ЗНАНИЙ ИНОСТРАННОГО ЯЗЫКА

В.М. Дёмкин, А.В. Соснин, С.С. Сусманова

Национальный исследовательский университет «Высшая школа экономики» – Нижний Новгород, Россия
vdyomkin@hse.ru

Аннотация

В работе обсуждаются современные подходы к проектированию сложных интеллектуальных систем компьютерного контроля знаний иностранного языка для поддержки учебного процесса в высшем учебном заведении. Исследуются механизмы построения онтологических моделей ситуаций в задачах компьютерного контроля знаний английского языка в рамках парадигмы ограниченной рациональности. Рассматриваются онтологические модели ситуаций на примере корпуса заданий на проверку знания личных местоимений в английском языке. На пути к становлению новой парадигмы компьютерного тестирования сделан очередной шаг, связанный с «очеловечиванием» поведенческой природы холонов, «примеривших» на себя одну из «шляп мышления» Эдварда де Боно. Процессы согласования «точек зрения» и «взаимных уступок» участников определяются цветом их шляп мышления, наиболее полно описывающих мотивы и логику шаблонов поведения людей.

Ключевые слова: компьютерная форма контроля знаний, открытый и закрытый тест, система тестирования, онтологический подход, холон, актор.

Введение

Среди ключевых факторов, оказывающих влияние на совершенствование и повышение качества учебного процесса в образовательных учреждениях и определяющих направление развития контроля знаний в целом, особое место по праву принадлежит информационным технологиям. Активное внедрение в учебный процесс различных форм компьютерного тестирования стало неотъемлемой частью современного образования.

Необходимо отметить, что многие из них до сих пор имеют отношение к так называемым традиционным компьютерным формам тестирования, где проверяются только теоретические знания. Ключевой идеей такого тестирования является «узнавание», поэтому здесь широко используются закрытые и открытые тесты на основе экранных форм шаблонов. Процедура проверки заключается в посимвольном сравнении шаблона с правильными ответами и заполненного шаблона с ответами тестируемого субъекта. Поскольку наряду с «узнаванием» имеет место также и «угадывание», то такие формы компьютерного тестирования служат основанием для мифологизации как самой процедуры контроля знаний, так и результатов этого контроля [1].

Контроль такого рода «знаний» не способен выявить наличие у тестируемого субъекта ни творческого начала, ни практических навыков. Его только спрашивают, что он может узнать или угадать из указанного контекста, не предоставляя ему возможности оказывать влияние на сам процесс тестирования. Результаты контроля всего лишь фиксируют состояние процессов «узнавания» и «угадывания». Они не достоверны и иллюзорны, именно поэтому способствуют укреплению мифа о действенности такой формы контроля знаний.

Накопленный опыт сообщества разработчиков программного обеспечения показывает, что предпринимаемые с их стороны усилия по «борьбе» с «узнаванием» и «угадыванием» неоправданны по затратам и бесплодны по своей сути. Миф о «всесилии» компьютерных технологий в вопросе контроля знаний культивируется десятилетиями!

Становление новой парадигмы компьютерного тестирования должно опираться на стремление вернуть человеку активную роль в осуществлении компьютерного контроля его знаний. В противном случае все усилия по совершенствованию традиционных форм компьютерного тестирования окажутся напрасными, и, как следствие, подобные попытки по-прежнему будут способствовать укреплению мифа о действенности форм такого рода контроля знаний.

Одним из возможных шагов на пути перехода от старых форм к принципиально новым может оказаться попытка применения такого подхода к построению сложных искусственных систем на основе онтологий, где краеугольным камнем новизны будет являться парадигма ограниченной рациональности принятия решений [2].

1 Компьютерный контроль знаний – это реальность

Иностранные языки, как и языки программирования, являются одной из привлекательных сфер применения новых форм компьютерного тестирования знаний благодаря присутствию в этих учебных дисциплинах ярко выраженных элементов творческого начала. Например, чтобы «побороть» рутину традиционных форм компьютерного тестирования в процессе изучения языков программирования, наряду с полной проверкой необходимо осуществлять и частичную проверку программного кода [3, 4]. Лишь в этом случае можно надеяться на успех в достижении контроля практических навыков программирования, а значит и контроля знаний.

Задания творческого характера при изучении иностранных языков, как правило, направлены на формирование таких приемов умственной деятельности, как синтез, анализ, обобщение, конкретизация, аналогия. Они позволяют организовать разнообразные деятельность-ные ситуации, способствующие лучшему усвоению материала и, в целом, логического мышления.

Перед тем как перейти к обсуждению свойств будущей новой системы тестирования знаний иностранного языка, перечислим основополагающие принципы парадигмы ограниченной рациональности:

- 1) Ограниченная рациональность опирается на фундаментальную онтологию М. Хайдеггера, в которой человек не только есть, но соотносится с собой, миром и миром других людей, что позволяет говорить о множестве персональных онтологий.
- 2) Ограниченная рациональность предполагает «встроенность» человека в ситуацию.
- 3) Ограниченная рациональность предполагает нахождение согласованного решения на основе солидарности неоднородных акторов (лиц, не только познающих, но и преобразующих ситуацию).
- 4) Ограниченная рациональность опирается на взаимодействие по принципу «часть-целое» элементов открытых систем.
- 5) Формализованное описание ситуаций осуществляется с помощью онтологических моделей ситуаций, включающих в себя персональные и групповые онтологии, а также объектные модели, разрабатываемые на основе взаимодействия неоднородных акторов [5].

Итак, парадигма ограниченной рациональности опирается на отношение «часть-целое». Каждая часть рассматривается как целостность (или целое), являющееся частью другого большого целостного (или целого). Целостность, участвующую в становлении новых цело-

стностей, принято называть холоном. Принцип целостности предполагает взаимодействие целостных компонентов (холонов).

Ключевой фигурой парадигмы ограниченной рациональности является человек-актор, берущий на себя полномочия принимать решения и несущий ответственность за последствия принимаемых решений. В этом смысле человек – это главное «действующее лицо» холона.

Следуя основополагающим принципам парадигмы ограниченной рациональности, укажем, что новая система тестирования знаний:

- функционирует при участии тестируемых субъектов, являющихся «одушевленными акторами», и их ответов на задания теста, являющихся «неодушевленными акторами»;
- является открытой системой, взаимодействующей с окружением и информацией, хранящейся и порождаемой в «недрах» системы;
- характеризуется неопределенностью и изменчивостью возникающих в ней проблемных ситуаций, побуждающих акторов принимать решения;
- организуется в качестве динамически упорядоченных целостностей, в которых образуются и совершенствуются взаимосвязи между частями целого.

Итак, в системе взаимодействуют тестируемые субъекты и их ответы, все они обладают статусом холонов. Акторы как «действующие лица» получают полномочия по управлению холонами и несут ответственность за последствия принимаемых решений.

Онтологические модели ситуаций исполняют роль платформы, на базе которой происходят процессы согласования «точек зрения» и «взаимных уступок» акторов. Акторы изучают проблемную ситуацию, сложившуюся в системе, и принимают решения о путях выхода из нее, согласуя свои действия с другими акторами. В конечном итоге будет найдено одно из решений: «актор обладает знанием» или «актор не обладает знанием». Принятие решения будет получено в темпе развития ситуации, т.е. в реальном масштабе времени.

Становление новой парадигмы позволит разрушить миф о действенности современных форм компьютерного контроля знаний. Тестируемый субъект и его ответы как активные творческие элементы открытой системы будут создавать и разрешать ситуации, которые позволят выявить глубину знаний.

Но чтобы и новый подход не стал очередным мифом, необходимо предпринять ряд усилий, связанных с «очеловечиванием» поведенческой природы холонов.

Одним из приемлемых подходов на этом пути могло бы стать следование концепции британского психолога и писателя, эксперта в области творческого мышления Эдварда де Боно, представленной в виде шести «мыслеварительных шляп», наиболее полно описывающих мотивы и логику шаблонов поведения людей [6].

Согласно этой концепции, при оценке полученной информации (например, от выступающего с презентацией человека), слушатели так или иначе разделяются на шесть групп относительно шаблона их мышления, а носителю каждого из шаблонов условно присваивается «мыслеварительная шляпа» определённого цвета.

Такой подход оказался очень полезен на тренингах, посвященных развитию мышления, логики и умственных способностей: он позволяет продуктивно строить дискуссию в форме игры, не переходя при этом на личности и пресекая возможность скандалов. Участники тренируются поочередно надевать на себя шляпы мышления и пробовать себя в той или иной роли. Меняя всякий раз формируемый каждой из них образ своих мыслей, они расширяют своё сознание и вскоре перестают мыслить в единственной присущей им парадигме.

Каждый шаблон поведения, согласно Э. де Боно, условно назван именно «шляпой» в силу того социального значения, которое в XX веке означал носимый её владельцем головной убор: он определял сословие своего хозяина, его социальный статус и часто даже род занятий.

Итак, в соответствии с окраской поведения своего владельца обнаруживаются белые, красные, чёрные, жёлтые, зелёные и синие шляпы.

Белая шляпа: информация. Цвет девственно-белого, холодного снега; в бесстрастной манере изложенные цифры, факты и объективная информация; белый цвет беспристрастен и объективен; в белой шляпе «варятся» мысли, «замешанные» на цифрах и фактах.

Белая шляпа предназначена для извлечения всей необходимой ей точной числовой, фактической информации. Она исключает доминирование эмоций, ориентируясь и руководствуясь уже полученной информацией, определяет, какая ещё информация необходима для составления полной картины проблемной ситуации.

Красная шляпа: чувства и интуиция. Красный цвет символизирует гнев, ярость и внутреннее напряжение; чувства и эмоции, предчувствия и интуицию.

Красная шляпа высказывает свои интуитивные опасения и догадки относительно предмета разговора, как правило, основанные на собственном либо чужом опыте, не рассуждая о дальнейшей стратегии поиска выхода решения из гипотетически возможных трудностей.

Чёрная шляпа: критика. Точность, доходящая до въедливости; негативизм и искренняя убеждённость в том, что «никогда в жизни ничто не может складываться так, как надо»; чёрный цвет мрачный, зловещий, словом – недобрый; чёрная шляпа покрывает собой все дурное – то, что страшится людских глаз.

Чёрная шляпа предназначена для эффективного поиска изъянов в теории, высказывает субъективную критику, показывает на возможные риски и не позволяет совершать непродуманные поступки, перепроверяет всё множество раз.

Жёлтая шляпа: поддержка и позитив. Солнечный свет; энтузиазм и оптимизм; позитивный взгляд на вещи; созидание; поиск благоприятных возможностей; жёлтый цвет солнечный, жизнеутверждающий; жёлтая шляпа полна оптимизма; под ней живет надежда и позитивное мышление.

Жёлтая шляпа действует в поддержку уже имеющейся теории, указывает на имеющиеся в ней достоинства и преимущества.

Зелёная шляпа: изобретательность. Символ плодородия; растение, пробивающее путь к свету сквозь толщу земли; творческое начало; движение; полёт фантазии; создание ситуаций, пробуждающих сознание; зелёная шляпа символизирует творческое начало и расцвет новых идей.

Зелёная шляпа предоставляет возможность изобиловать новыми идеями, бесконечно совершенствовать уже имеющиеся модели: в общем, даёт полную свободу фантазии и творчеству.

Синяя шляпа: контроль. Целостный контроль; дисциплина; образ – дирижёр крупного оркестра; управление процессом «мыслеварения»; синий цвет холодный; это – цвет неба; синяя шляпа связана с организацией и управлением мыслительным процессом, а также с применением шляп других расцветок.

Синяя шляпа работает не с предметом дискуссии, а управляет самим её процессом и другими шляпами мышления для составления объективной оценки обо всех высказанных теориях и предположениях. На основе полученных сведений синяя шляпа должна подвести итоги дебатов и наметить новые цели.

На наш взгляд, некоторые из вышеуказанных типов мышления могут оказаться непригодными в рамках конкретных задач, и такие шляпы следует отбросить. Так, для некоторых форм контроля знаний иностранных языков таковой, к примеру, может стать зелёная шляпа мышления. Она могла бы оказаться полезной в рамках олимпиад, а вот в случае аккредитации привитых знаний, которая зачастую сводится к проверке владения правилами грамматики, будет лишней.

Итак, для новой системы компьютерного контроля знаний иностранного языка, а именно английского, ограничим свой выбор «белым», «красным», «чёрным», «жёлтым» и «синим» шаблонами мышления, которым придадим реальный облик в контексте модели тестирования. Цвета шляп мышления, «надетых» на холоны, будут определять окраску их поведения. Рисунок 1 отражает общую концепцию работы холонов в рамках модели тестирования. Здесь наряду с указанием характеристик шаблонов поведения раскрыто содержание главных и контекстных задач холонов.



Рисунок 1 – Общая схема работы холонов

2 Моделирование общих принципов работы холонов

«Синий» холон управляет событиями в ходе тестирования таким образом, чтобы все остальные действующие холоны имели прочную базу для анализа в соответствии с их парадигмой действия. На рисунке 2 представлена общая схема работы «синего» холона, включающая два основных этапа. «Синий» холон обеспечивает первоначальный запуск «белого» «красного» и «жёлтого» холонов. Он же обеспечивает и завершение их работы, сопоставляет оценки, полученные в результате работы «красного» и «белого» холонов (этап I), и определяет дальнейшее участие в процессе тестирования «чёрного» или «жёлтого» холонов (этап II).

«Красный» – это один из базовых холонов, который во многом определяет всю стратегию поведения модели тестирования и оказывает самое сильное воздействие на конечную оценку. На рисунке 3 представлена общая схема работы «красного» холона. Суждение о знании строится либо на основе специфических параметров структуры вопросов, либо на основе времени, затраченного на ответ. Так как работа «красного» холона построена на интуитивном анализе, то основой для предположений о знании будет время отклика на вопросы теста. «Красный» холон просчитывает динамику только по тем вопросам, на которые был дан верный ответ. Он сравнивает средние времена выполнения заданий в каждой категории, оценивает средние времена выполнения заданий на каждом уровне сложности, рассчитывает итоговую оценку на основе среднего времени выполнения заданий теста (этап I, основной блок тестов). Наличие ответов, подозрительных на «узнавание» или «угадывание», приводит к передаче управления «белому» холону для подтверждения их достоверности.

«Жёлтый» – это позитивно настроенный холон. На рисунках 4 и 5 представлены общие схемы работы «жёлтого» холона на обоих этапах хода тестирования. Роль «жёлтого» холона связана с оказанием поддержки в виде упрощённых дополнительных вопросов в случае

«внезапного» появления неправильных ответов (этап I, рисунок 4). Если и в дополнительных вопросах «внезапно» появляются неправильные ответы, управление будет передано «белому» холому. В свою очередь на этапе II (дополнительный блок тестов) «жёлтому» холому может быть передано управление от актора, уполномоченного управлять «белым» холомом, с тем, чтобы он взял на себя роль «красного» холома (этап II, рисунок 5). В этом случае «жёлтым» холомом задаются вопросы, уровень сложности которых ранее уже был присвоен «красным» холомом, время отклика на вопросы теста при этом не учитывается. «Белый» холол вместо выключенного на время «красного» холола фиксирует теперь динамику ответов.

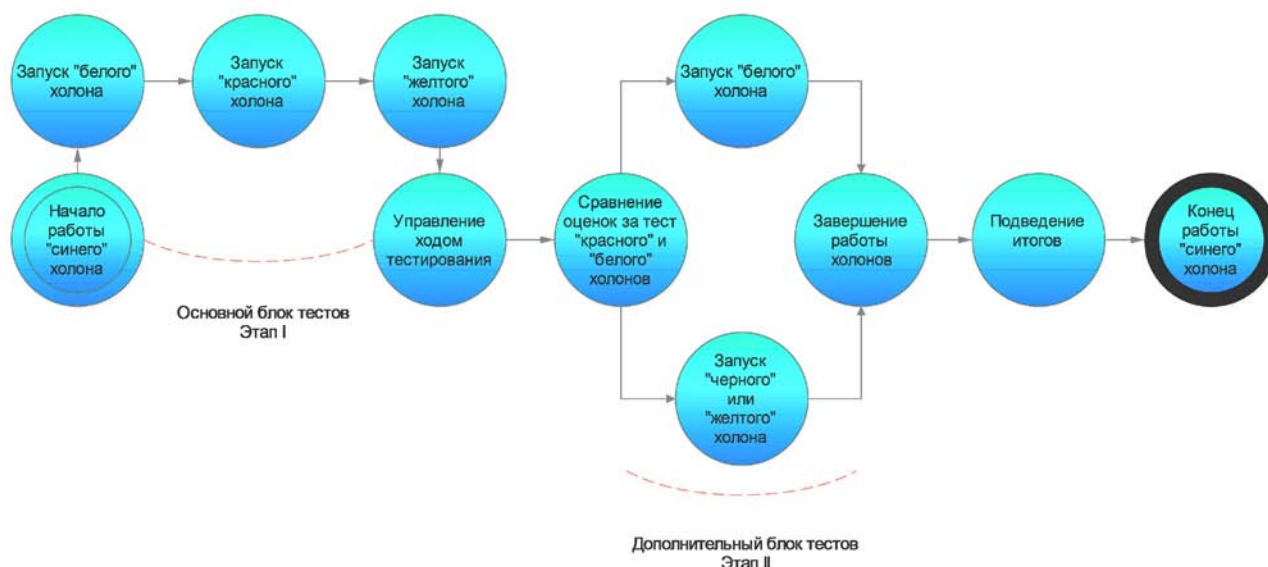


Рисунок 2 – Общая схема работы «синего» холола



Рисунок 3 – Общая схема работы «красного» холола



Рисунок 4 – Общая схема работы «жёлтого» холона на этапе I



Рисунок 5 – Общая схема работы «жёлтого» холона на этапе II

Роль «чёрного» холона определена таким образом, что на его вопросы непросто ответить. Здесь и открытые вопросы, и вопросы, в которых нет правильного ответа, и вопросы с большим количеством правильных ответов. Зачастую используется и психологический прием на отвлечение внимания. Проверка подлинности фактического уровня знаний – одна из главных задач актора, уполномоченного управлять «чёрным» холоном. Потому и разработке вопросов для «чёрного» холона уделяется самое пристальное внимание.

На рисунке 6 представлена общая схема работы «чёрного» холона. «Белый» холон предоставляет «чёрному» сведения обо всех недочётах, выявленных «красным» и «жёлтым» холонами во время проведения испытаний на каждом этапе хода тестирования. Преодоление препятствий, поставленных «чёрным» холоном, как правило, является неизбежным «злом» при подведении окончательных итогов тестирования.

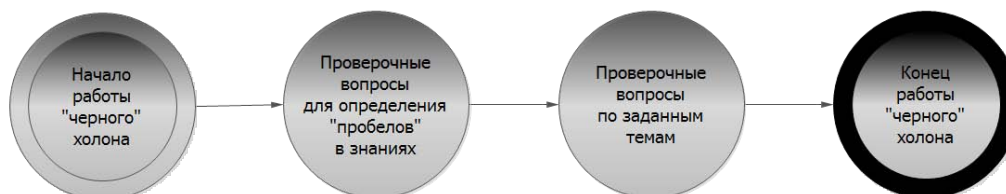


Рисунок 6 – Общая схема работы «чёрного» холона

«Белый» – это наиболее важный холон в модели тестирования. Помимо базовых расчётов он корректирует работу «красного» и «жёлтого» холонов. Поскольку единственной информацией является количество верных и неверных ответов на вопросы, то «белый» холон строит свою оценку исходя из общего количества вопросов и их уровня сложности. «Белый» холон работает на протяжении обоих этапов хода тестирования, выполняя следующие задачи:

- запись в собственную базу данных сведений о динамике прохождения заданий теста;
- проведение проверок на достоверность знаний и динамическую коррекцию собственной базы данных по итогам прохождения этих проверок;
- сбор статистики и коррекцию внутренних параметров «красного» и «жёлтого» холонов.

Общие схемы работы «белого» холона на обоих этапах хода тестирования представлены на рисунках 7 и 8.



Рисунок 7 – Общая схема работы «белого» холона на этапе I



Рисунок 8 – Общая схема работы «белого» холона на этапе II

3 Контроль знаний с позиций традиционного подхода

Перед тем как перейти к обсуждению механизмов контроля знаний иностранного языка, в качестве рабочего материала выберем личные местоимения. В таблице 1 представлен корпус заданий на проверку знания личных местоимений в английском языке. Личные местоимения изменяются по лицам и числам, а также имеют формы двух падежей: именительного и объектного. Полная парадигма приводится в таблице 2 (курсивом помечены редкие формы, употребление которых ограничено определёнными стилями или диалектами). Данный корпус является частью общего корпуса заданий по всем английским местоимениям.

Тестовые задания разделены на шесть уровней сложности – от простейшей подстановки местоимений по смыслу до стилистических нюансов и выбора архаичных форм в поэтическом дискурсе. Уровни условно соответствуют традиционно выделяемым уровням знания английского языка: Elementary, Pre-Intermediate, Intermediate, Upper-Intermediate, Advanced, Proficiency («начальный», «выше начального», «средний», «выше среднего», «продвинутый», «профессиональный»). Для каждого задания на каждом уровне даётся три проверочных задания, исключающих произвольный выбор ответа или случайную ошибку.

Проверочные задания организованы по степени возрастания сложности: первое задание по своей грамматической структуре практически дублирует исходное задание; во втором за-

дании допускается некоторое варьирование грамматических форм и конструкций; в третьем задании грамматический или лексический материал проверяется в его парадигматических связях, т.е. предусмотрена отсылка к смежным частям данного корпуса заданий или к другим корпусам. Однако на простейших уровнях корпуса (в нашем случае уровни 1 и 2) прямой выход в другие корпуса, как правило, не предусматривается.

В рассматриваемом корпусе для каждого задания предлагается три варианта ответа. На простейших уровнях выбор ответа осуществляется исключительно из проверяемого корпуса или даже подкорпуса грамматики (так, на уровне 1 – только из личных местоимений в именительном падеже, на уровнях 2 и 3 – из личных местоимений в именительном и объектном падежах, и только на уровне 4 в вариантах ответа появляются притяжательные местоимения).

В задания на самых сложных уровнях (в нашем случае на уровне 6 – «профессиональный») включены элементы текстовой интерпретации, что предполагает выход на уровни текстового анализа и в литературоведение как дисциплину, не отделимую от изучения языка и формирования языковой компетенции.

Стандартное время на выполнение заданий (прочтение вопроса, самого тестового задания (т.е. микроконтекста, например, предложения с пропуском) и вариантов ответа; время на раздумье; подстановка ответа) составляет в среднем 95 знаков без пробелов в минуту (данные получены на основе анализа ряда тестов из учебников «Language Leader» и «Intelligent Business» различных уровней сложности). Однако эти данные должны быть проверены эмпирически для каждого уровня сложности и каждого задания в корпусе.

Отметим, что представленный корпус заданий являет собой элементарную модель, которая скорее описывает принципы построения корпусов, нежели имеет прямую практическую ценность. Отметим также, что неудобство, связанное с разбиением таблицы на две составные части по столбцам, преодолимо вследствие контекстной обособленности каждой из них.

Таблица 1 – Корпус заданий (начало, столбцы I и II)

	I	II
	Уровень сложности	Задание
1.	Elementary Подстановка по смыслу местоимения в именительном падеже	John is late. Has ____ called? a) we b) he c) she [ответ: b)]
2.	Pre-Intermediate Подстановка местоимения в именительном падеже в функции подлежащего в зависимости от формы сказуемого	Nick is lucky. He earns much more than ____ do. a) I b) me c) she [ответ: a)]
3.	Intermediate Подстановка местоимения в объектном падеже	Mary works hard because she wants her boss to give ____ a pay rise. a) him b) her c) she [ответ: b)]
4.	Upper-Intermediate Подстановка местоимения в объектном падеже в оборотах «объектный падеж с инфинитивом» и «объектный падеж с причастием настоящего времени»	I saw ____ leave only a couple of minutes ago. They can't have gone far. a) they b) their c) them [ответ: c)]
5.	Advanced Стилистические различия в употреблении личных местоимений в именительном и объектном падежах	Выберите более разговорную форму: I hate jazz. - ____ too. a) I do b) Me [ответ: b)]
6.	Proficiency Архаичные формы личных местоимений в поэтическом дискурсе	Abandon hope all ____ who enter here. a) thou b) thee c) ye [ответ: c)]

Таблица 1 – Корпус заданий (продолжение, столбцы III и IV)

	III	IV
	Проверка задания	Отсылка к другим уровням или корпусам
1.	1) John and Henry are very rich. ____ are millionaires. a) We b) They c) He [ответ: b)] 2) Don't tease the dog. ____ may bite you. a) She b) I c) It [ответ: c)] 3) John gave us a great idea. ____ all admired it. a) We b) They c) He [ответ: a)]	Правильный ответ в проверочном задании (3) выбирается за счёт идентификации местоимения в объектном падеже «us» и подбора соответствующего местоимения в именительном падеже «we». → уровень 3
2.	1) I hate jazz. - ____ do, too. a) He b) Us c) We [ответ: c)] 2) I spend more money than ____ does. a) we b) her c) she [ответ: c)] 3) I hate cats. - ____ too. a) I b) Me c) They [ответ: b)]	Проверочное задание (3) основано на различии узуального употребления местоимения в личном и объектном падеже: «I do, too», но «Me too». → уровни 3 и 5
3.	1) Could you help ____ with this box? It's so heavy! a) me b) he c) I [ответ: a)] 2) Joe's late. It's not like ____ to be late. a) he b) him c) them [ответ: b)] 3) Susan is tall but I'm taller than ____. a) me b) she c) her [ответ: c)]	2) → корпусы: предлоги; синтаксис структур сравнения 3) → корпусы: прилагательные (степени сравнения); синтаксис структур сравнения
4.	1) Mary, dear, I want ____ to sing for the guests. a) me b) you c) your [ответ: b)] 2) I saw ____ crossing the street yesterday. Where was he going, I wonder? a) he b) her c) him [ответ: c)] 3) He suggested ____ leaving at once. a) our b) us c) we [ответ: a)]	1) → корпус: инфинитив (оборот «объектный падеж с инфинитивом») 2) → корпус: причастие (причастие настоящего времени); оборот «объектный падеж с причастием настоящего времени») 3) → корпусы: притяжательные местоимения; управление глаголов; герундий
5.	1) Выберите более разговорную форму: We spend more money than ____ . a) them b) they do [ответ: a)] 2) Выберите более литературную форму: I'm tired of this job. - ____ too. a) Me b) I am [ответ: b)] 3) Выберите более разговорную форму: Look! Isn't that Bill? – It can't be ____ . He is in New York now. a) him b) he [ответ: a)]	1) → корпусы: прилагательные (степени сравнения); синтаксис разговорной речи 2) → корпус: синтаксис литературной речи 3) → корпусы: синтаксис разговорной речи; модальные глаголы (can)

6.	1) ____ gods! What shall be the end to this? a) Ye b) Thou c) Thee [ответ: a)] 2) Though thy crest be shorn and shaven, ____, I said, art no craven, Ghastly grim and ancient raven, Wandering from the nightly shore. (<i>The Raven</i> by E.A. Poe) a) Thy b) Thou c) Thee [ответ: b)] 3) Our little systems have their day, They have their day and cease to be: They are but broken lights of ____ , And Thou, O Lord, art more than they (<i>In Memoriam</i> by A. Tennyson) a) Ye b) Thee c) Thy [ответ: b)]	→ корпуса: грамматика и синтаксис поэтической речи; словарный запас продвинутого уровня (архаизмы); интерпретация текста; англоязычная литература
----	---	---

Таблица 2 – Личные местоимения

Лицо	Именительный падеж	Объектный падеж
Единственное число		
1-е	I (я)	me (меня, мне)
2-е	thou (поэтич., архаич., диалектн. ты)	thee (поэтич., архаич., диалектн. тебя, тебе)
3-е	he (он) she (она) it (он, она, оно)	him (его, ему) her (ее, ей) it (его, ее, ему, ей)
Множественное число		
1-е	we (мы)	us (нас, нам)
2-е	you (вы); ye (поэтич., архаич. диалектн. вы)	you (вас, вам); ye (поэтич., архаич. диалектн. вас, вам)
3-е	they (они)	them (их, им)

4 Контроль знаний с позиций холонического подхода

Рассмотрим онтологические модели ситуаций на примере представленного корпуса заданий, «вживаясь» в роль каждого холона, «примерившего» на себя одну из шляп мышления.

«Синий» холон обеспечивает первоначальный запуск и завершение работы «белого», «красного» и «желтого» холонов, задает вопросы в порядке возрастания уровня сложности из столбца II таблицы 1.

«Жёлтый» холон по определению оказывает поддержку в виде дополнительных вопросов, подобных вопросам из столбца II таблицы 1, в случае «внезапного» появления неправильных ответов на вопросы, заданные «синим» холоном, и призывает «белый» холон на помощь в тех случаях, когда имеет место ошибка после череды правильных ответов. При этом «белым» холоном фиксируется «доверительное» количество верных ответов.

«Красный» холон определяет всю стратегию поведения модели тестирования, он занят анализом и расчётами, оказывает наибольшее воздействие на итоговую оценку. Он же передаёт управление «белому» холону для подтверждения достоверности ответов, подозрительных на «узнавание» или «угадывание».

«Белый» холон задаёт проверочные вопросы из столбца III таблицы 1, если в дополнительных вопросах, задаваемых «жёлтым» холоном, «внезапно» появляются ложные ответы. Он же задаёт один из вопросов из этого же столбца после передачи ему управления от «красного» холона в случае наличия ответов, подозрительных на «узнавание» или «угадывание», для подтверждения их достоверности. Вопрос при этом задаётся в открытой форме.

«Чёрный» холон принимает на себя управление от «синего» в том случае, если появилось «подозрение» на недобросовестность или везение. Его задача – уличить. Каждый из «правильных» или «неправильных» вариантов ответа хранит ссылку на соответствующий уровень данного или другого корпуса, где и должен осуществляться дополнительный этап проверки.

Заключение

Сформулируем основные результаты работы:

- 1) На пути к становлению новой парадигмы компьютерного тестирования сделан очередной шаг, связанный с «очеловечиванием» поведенческой природы холонов, «примеривших» на себя одну из шляп мышления Эдварда де Боно. Процессы согласования «точек зрения» и «взаимных уступок» участников определяются цветом их шляп мышления, наиболее полно описывающих мотивы и логику шаблонов поведения людей.
- 2) Проанализированы онтологические модели ситуаций на примере корпуса заданий на проверку знания личных местоимений в английском языке.

Список источников

- [1] **Гинзбург, А.Н.** Компьютерный контроль знаний – миф или реальность / А.Н. Гинзбург, В.М. Дёмкин // Проблемы управления и моделирования в сложных системах: Труды XII Международной конф. (21–23 июня 2010 г., Самара, Россия). – Самара: Самарский НЦ РАН, 2010. – С. 501–504.
- [2] **Виттих, В.А.** Парадигма ограниченной рациональности принятия решений-2 / В.А. Виттих // Вестник Самарского государственного технического университета. Серия «Технические науки». – 2010. – № 2(26). – С. 23–31.
- [3] **Гинзбург, А.Н.** Адаптивный подход к управлению процессом разработки систем автоматического тестирования решения задач по программированию. Результаты и перспективы разработки / А.Н. Гинзбург, В.М. Дёмкин, А.Н. Рыбалкин, С.В. Сатунин // Проблемы управления и моделирования в сложных системах: Труды IX Международной конф. (22–28 июня 2007 г., Самара, Россия). – Самара: Самарский НЦ РАН, 2007. – С. 603–613.
- [4] **Антропов, А.А.** Архитектура Web-ориентированной системы автоматического контроля знаний дисциплины программирования / А.А. Антропов, Е.В. Баранов, А.Н. Гинзбург, В.М. Дёмкин, А.В. Мотолин // Проблемы управления и моделирования в сложных системах: Труды XI Международной конф. (22–24 июня 2009 г., Самара, Россия). – Самара: Самарский НЦ РАН, 2009. – С. 658–663.
- [5] **Смирнов, С.В.** Онтологическое моделирование в ситуационном управлении / С.В. Смирнов // Онтология проектирования. – 2012. – №2(4). – С. 16–24.
- [6] **Де Боно, Э.** Шесть шляп мышления / Э. де Боно. – СПб.: Питер Паблишинг, 1997. – 256 с. – (Серия «Тренировка ума»).

ONTOLOGICAL MODELS OF SITUATIONS APPLIED TO THE TASK OF COMPUTER-AIDED TESTING OF FOREIGN LANGUAGES PROFICIENCY

V.M. Dyomkin, A.V. Sosnin, S.S. Susmanova

*National Research University Higher School of Economics, Nizhny Novgorod Campus, Russia
vdyomkin@hse.ru*

Abstract

The paper discusses modern approaches towards the design of complicated intellectual computer systems assessing foreign language proficiency, e.g. checking students' academic progress in a higher educational establishment. The paper

provides insight into the means of developing ontology-based situation models for the tasks requiring that a person's command in English could be assessed, which is done within the paradigm of bounded rationality. The paper demonstrates the characteristic behavioral features of the holons, each of which assumes the role of one of E. de Bono's six «thinking hats» on the example of the English personal pronouns. The processes of negotiating the «opinions» and «mutual concessions» are defined by the color of the thinking hat that does the best job of describing people's motives and behavior pattern logic.

Key words: computer-aided knowledge assessment, open and closed text, testing system, ontological approach, holon, actor.

References

- [1] **Ginzburg, A.N.** Komp'yuternyj kontrol' znaniy – mif ili real'nost' [Computer-aided knowledge assessment – myth or reality] / A.N. Ginzburg, V.M. Dyomkin // Problemy upravleniya i modelirovaniya v slozhnykh sistemakh: Proceedings of the XII international conference (21-23 June 2011, Samara, Russia). – Samara: SamNTC RAS, 2010. – pp. 501–504. (In Russian).
- [2] **Vitikh, V.A.** Paradigma ogranichennoj ratsional'nosti prinyatiya reshenij-2 [The bounded rationality paradigm for decision making -2] / V.A. Vitikh // Bulletin of the Samara State Technical University. «Technical sciences» series. – 2010. – No. 2(26). – P. 23–31. (In Russian).
- [3] **Ginzburg, A.N.** Adaptivnyj podkhod k upravleniyu protsessom razrabotki sistem avtomaticheskogo testirovaniya resheniya zadach po programmirovaniyu. Rezul'taty i perspektivy razrabotki [Adaptive management approach to the development process of automatic test solutions applied to programming problems. Results and prospects of development] / A.N. Ginzburg, V.M. Dyomkin, A.N. Rybalkin, S.V. Satunin // Problemy upravleniya i modelirovaniya v slozhnykh sistemakh: Proceedings of the IX international conference (22-28 June 2007, Samara, Russia). – Samara: SamNTC RAS, 2007. – pp. 603–613. (In Russian).
- [4] **Anthropov, A.A.** Arkhitektura Web-orientirovannoj sistemy avtomaticheskogo kontrolya znaniy distsipliny programmirovaniya [Architecture of a Web-oriented computer-aided knowledge assessment for programming] / A.A. Anthropov, E.V. Baranov, A.N. Ginzburg, V.M. Dyomkin, A.V. Motolin // Problemy upravleniya i modelirovaniya v slozhnykh sistemakh: Proceedings of the XI international conference (22-24 June 2009, Samara, Russia). – Samara: SamNTC RAS, 2007. – pp. 658–663. (In Russian).
- [5] **Smirnov, S.V.** Ontologicheskoe modelirovanie v situatsionnom upravlenii [Ontological modeling in situational management] / S.V. Smirnov // Ontology of Designing. – 2012. – №2(4). – С. 16-24. (In Russian).
- [6] **De Bono, E.** Shest' shlyap myshleniya [Six thinking hats] / E. de Bono. – Saint-Petersburg: Piter publishing, 1997. – 256 p. – («Mind training» series). (In Russian).

Сведения об авторах



Дёжкин Валерий Матвеевич, 1951 г. рождения. Окончил Горьковский государственный университет им. Н.И. Лобачевского в 1974 г., к.т.н. (1991), доцент (2006). Доцент кафедры информационных систем и технологий Национального исследовательского университета «Высшая школа экономики» – Нижний Новгород, доцент кафедры прикладной математики Нижегородского государственного технического университета, старший научный сотрудник Института прикладной физики РАН. В списке научных трудов более 80 работ в области автоматизации научных исследований и применения информационных технологий в образовании.

Valeriy Matveevich Dyomkin (b. 1951) graduated from the Lobachevskiy State University, Gorky, in 1974, PhD in Computer Science (1991). He is an Assistant Professor at National Research University Higher School of Economics, Nizhny Novgorod Campus, Department of Information Systems and Technologies, he holds a part-time position of Assistant Professor at State Technical University, Nizhny Novgorod, Department of Applied Mathematics, he is Senior Researcher at Institute of Applied Physics, RAS. He is author/co-author of over 80 scientific papers on automation of scientific research and application of information technologies in education.



Соснин Алексей Владимирович, 1982 г. рождения. Окончил Нижегородский государственный лингвистический университет им. Н.А. Добролюбова в 2004 г., к.филолог.н. (2007). Доцент кафедры иностранных языков Национального исследовательского университета «Высшая школа экономики» – Нижний Новгород. В списке научных трудов более 30 работ в области семантики, когнитивной лингвистики и литературоведения.

Alexey Vladimirovich Sosnin (b. 1982) graduated from Nizhny Novgorod Linguistics University in 2004, got a PhD in English philology in 2007. He is Assistant Professor at National Research University Higher School of Economics, Nizhny Novgorod, Department of Foreign Languages. He is author/co-author of over 30 scientific papers on semantics, cognitive linguistics and literary studies.



Сусманова Софья Сергеевна, 1990 г. рождения. Окончила Национальный исследовательский университет «Высшая школа экономики» в 2012 г., бакалавр прикладной математики и информатики.

Sofya Sergeevna Susmanova (b. 1990) graduated from the National Research University Higher School of Economics in 2012, the bachelor of applied mathematics and computer science.

УДК 519.61

СТАБИЛЬНОСТЬ НЕОПРЕДЕЛЁННЫХ ЗАДАЧ ОПТИМИЗАЦИИ

В.И. Левин

Пензенский государственный технологический университет, г. Пенза, Россия
vilevin@mail.ru

Аннотация

Рассмотрена задача оптимизации неполностью определённых функций, т.е. функций с параметрами, заданными лишь с точностью до интервала. Дан обзор существующих подходов к решению задач оптимизации неполностью определённых функций с различными видами неопределённости. Описана математическая постановка задачи оптимизации функции с интервальными параметрами и метод ее решения путем сведения к двум задачам оптимизации полностью определённых функций, т.е. функций с точно известными параметрами (метод детерминизации). Показано, что решение проблемы оптимизации неполностью определённых функций требует также рассмотрения задачи определения устойчивости оптимума к варьированию значений параметров функции. В связи с этим введены понятия макроустойчивости и микроустойчивости задачи оптимизации полностью определенной функции. Даны необходимые и достаточные условия макроустойчивости задачи оптимизации полностью определенной функции. Приведен алгоритм проверки макроустойчивости. Дан пример проверки макроустойчивости конкретной задачи с помощью этого алгоритма (задача о назначениях). Приведен также алгоритм проверки микроустойчивости задачи оптимизации полностью определенной функции. Для решения указанных задач используются методы интервальной математики.

Ключевые слова: оптимизация систем, неопределенность, устойчивость оптимума, варьирование параметров, интервальная математика, макроустойчивость, микроустойчивость.

Введение

Сегодня в мире имеется обширная литература по оптимизации различных систем с детерминированными параметрами – технических, экономических и т.д. Соответствующие задачи формулируются как задачи математического программирования с целевыми функциями и функциями ограничений, параметры которых являются детерминированными величинами. При этом на практике чаще встречаются системы с недетерминированными параметрами. Оптимизация таких систем обычно формализуется в виде задач математического программирования с целевыми функциями и функциями ограничений, параметры которых суть различные недетерминированные величины: случайные, нечеткие, интервальные и т.д. Эти задачи сложнее детерминированных. Они требуют обобщения понятия экстремума функции, выяснения условия его существования, связанных с недетерминированностью параметров функции, и создания специальных методов поиска экстремума таких функций.

Известно три различных подхода к решению недетерминированных задач математического программирования: детерминированный, вероятностный [1] и интервальный [2]. Детерминированный подход заключается в решении задачи для определенных значений ее параметров, выбранных внутри заданных областей неопределенности. Вероятностный подход состоит в решении задачи для усредненных (ожидаемых, в смысле математического ожидания) значений ее параметров, что предполагает задание вероятностной меры внутри их областей неопределенности. Оба указанных подхода объединяет предварительная детерминизация параметров задачи, выполняемая перед ее оптимальным решением. В отличие от них, интервальный подход не предполагает никакой детерминизации параметров, которые задаются в ин-

тервальной форме. В данном подходе оптимальное решение задачи проводится на основе прямого сравнения недетерминированных значений целевой функции, соответствующих различным значениям вектора аргументов, и выбора оптимального значения данной функции. Достоинства и недостатки указанных подходов рассмотрены в [1–8]. Соответствующий краткий обзор дан в работе [9].

Изложенные подходы объединяет одна существенная черта – все они предназначаются для решения задач оптимизации, в которых параметры целевых функций и функций ограничений точно неизвестны. Поэтому может оказаться, что действительные значения параметров задачи отличаются от тех, которые были приняты в процессе отыскания решения. В этом случае для того, чтобы найденное оптимальное решение задачи имело содержательный смысл, нужно, чтобы оно еще обладало следующим свойством: при небольшом варьировании значений параметров задачи ее оптимальное решение должно по-прежнему существовать. При этом точка, в которой достигается оптимум целевой функции, может переместиться из исходного положения в новое положение, которое, однако, должно быть близко к исходному. Другими словами, требуется, чтобы найденное оптимальное решение неполностью определенной (недетерминированной) задачи математического программирования было устойчивым относительно небольших количественных изменений ее параметров.

1 Постановка задачи

В процессе постановки настоящей задачи мы отталкиваемся от постановки задачи в нашей работе [9].

Пусть задана некоторая непрерывная функция n вещественных переменных

$$(1) \quad y = F(x_1, \dots, x_n),$$

где параметры (коэффициенты) ее явного представления $p_k, k = \overline{1, l}$, известны точно. Будем рассматривать функцию (1) в ограниченной области, определяемой системой ограничений

$$(2) \quad \Phi_i(x_1, \dots, x_n) \leq b_i, \quad i = \overline{1, m},$$

имеющей точные параметры $q_s, s = \overline{1, t}$, явного представления функций ограничений Φ_i и правые части b_i . Тогда относительно функции (1) можно сформулировать полностью определенную задачу условной оптимизации (математического программирования)

$$(3) \quad F(x_1, \dots, x_n) = \max,$$

при условии

$$(4) \quad \Phi_i(x_1, \dots, x_n) \leq b_i, \quad i = \overline{1, m}.$$

Теперь предположим, что в задаче оптимизации (3), (4) параметры явного представления целевой функции F и функций ограничений Φ_i , а также правые части ограничений b_i известны не точно, а приближенно. Тогда, в соответствии со сказанным в Введении, мы должны вместе с задачей условной оптимизации (3), (4) рассматривать еще задачу проверки устойчивости решения задачи (3), (4) относительно небольших изменений ее параметров.

В отличие от существующих методов [5] изучения устойчивости решения задач оптимизации, будем рассматривать все возможные количественные значения каждого параметра задачи как единое целое. Это позволит задавать все возможные количественные значения параметров задач оптимизации в теоретико-множественных терминах. Простейший способ такого задания состоит в том, чтобы задать совокупность указанных значений параметров задачи в

виде соответствующих числовых интервалов. Преимущество этого подхода к изучению устойчивости решения задач оптимизации в том, что возникает возможность изучать устойчивость с помощью хорошо разработанных методов интервальной математики [10].

Итак, совместно с полностью определенной задачей (3), (4) мы должны рассмотреть производную от нее интервальную задачу условной оптимизации

$$(5) \quad \tilde{F}(x_1, \dots, x_n) = \max,$$

при условии

$$(6) \quad \tilde{\Phi}_i(x_1, \dots, x_n) \leq \tilde{b}_i, i = \overline{1, m} \}.$$

Целевая функция $\tilde{F}$ интервальной задачи оптимизации (5), (6) получается из целевой функции F искомой, полностью определенной задачи оптимизации (3), (4) путем замены ее точных параметров $p_k, k = \overline{1, l}$, соответствующими интервальными параметрами $\tilde{p}_k = [p_{k1}, p_{k2}], k = \overline{1, l}$. Аналогично, любая функция ограничений $\tilde{\Phi}_i, i = \overline{1, m}$, интервальной задачи (5), (6) получается из соответствующей функции $\Phi_i, i = \overline{1, m}$, исходной полностью определенной задачи (3), (4) заменой ее точно известных параметров $q_{si}, s = \overline{1, t}, i = \overline{1, m}$, соответствующими интервальными параметрами $\tilde{q}_{si} = [q_{si1}, q_{si2}], s = \overline{1, t}, i = \overline{1, m}$. Так же интервальные параметры $\tilde{b}_i, i = \overline{1, m}$, в ограничениях интервальной задачи (5), (6) заменяют собой соответствующие точно известные параметры $b_i, i = \overline{1, m}$ в ограничениях исходной, детерминированной задачи оптимизации (3), (4).

Будем называть полностью определенную задачу условной оптимизации (математического программирования) (3), (4) макроустойчивой, если она имеет решение и, кроме того, имеет решение производная от нее интервальная задача оптимизации (5), (6).

Далее, будем называть полностью определенную задачу условной оптимизации (математического программирования) (3), (4) микроустойчивой, если она макроустойчива и, сверх того, существует пара решений (x', x'') , где $x' = (x'_1, \dots, x'_n)$ – некоторая точка решения задачи оптимизации (3), (4), а $x'' = (x''_1, \dots, x''_n)$ – некоторая точка решения задачи (5), (6), расстояние между которыми $D(x', x'')$ не превосходит заданной достаточно малой величины d . Задача настоящего исследования – разработать алгоритмы определения макро- и микроустойчивости полностью определенных задач условной оптимизации типа (3), (4).

2 Математический аппарат

В основу решения поставленной задачи положим аппарат интервальной математики [10], где алгебраические операции над интервальными числами $\tilde{a} = [a_1, a_2], \tilde{b} = [b_1, b_2], \dots$ вводятся как следующие теоретико-множественные конструкции

$$(7) \quad \tilde{a} + \tilde{b} = \{a + b \mid a \in \tilde{a}, b \in \tilde{b}\}, \tilde{a} - \tilde{b} = \{a - b \mid a \in \tilde{a}, b \in \tilde{b}\}, k\tilde{a} = \{ka \mid a \in \tilde{a}\}, \dots$$

и т.д. Другими словами, любая операция над интервалами определяется на основе соответствующей операции над точечными величинами, при условии, что конкретные значения величин пробегают все возможные значения из соответствующих интервалов. Из введенных алгебраических операций над интервалами вытекают простые правила выполнения операций:

$$\begin{aligned}
 [a_1, a_2] + [b_1, b_2] &= [a_1 + b_1, a_2 + b_2], \\
 [a_1, a_2] - [b_1, b_2] &= [a_1 - b_2, a_2 - b_1]; \\
 (8) \quad k[a_1, a_2] &= \begin{cases} [ka_1, ka_2], & k > 0, \\ [ka_2, ka_1], & k < 0; \end{cases} \\
 [a_1, a_2] \cdot [b_1, b_2] &= [\min_{i,j}(a_i \cdot b_j), \max_{i,j}(a_i \cdot b_j)]; \\
 [a_1, a_2] / [b_1, b_2] &= [a_1 \cdot a_2] \cdot [1/b_2, 1/b_1].
 \end{aligned}$$

Введем теперь операции сравнения интервальных чисел. Единственное разумное здесь, согласно [2, 8, 10] – реализовать операцию сравнения интервалов на теоретико-множественном уровне, подобно алгебраическим операциям над интервалами (7). Так, введем операции взятия максимума $\vee$ и минимума $\wedge$ интервальных чисел $\tilde{a} = [a_1, a_2]$ и $\tilde{b} = [b_1, b_2]$ в виде конструкций

$$(9) \quad \tilde{a} \vee \tilde{b} = \{a \vee b \mid a \in \tilde{a}, b \in \tilde{b}\}, \quad \tilde{a} \wedge \tilde{b} = \{a \wedge b \mid a \in \tilde{a}, b \in \tilde{b}\}.$$

Свойства введенных операций сравнения интервальных чисел (9) определяются ниже-следующими теоремами (подробнее см. [2, 8, 10]).

Теорема 1. Для сравнимости двух интервалов $\tilde{a} = [a_1, a_2]$ и $\tilde{b} = [b_1, b_2]$ и их нахождения между собой в отношении $\tilde{a} \geq \tilde{b}$ необходимо и достаточно, чтобы одноименные границы этих интервалов удовлетворяли условиям

$$(10) \quad a_1 \geq b_1, \quad a_2 \geq b_2,$$

а для сравнимости этих интервалов и их нахождения между собой в отношении $\tilde{a} \leq \tilde{b}$ – чтобы удовлетворялись следующие условия:

$$(11) \quad a_1 \leq b_1, \quad a_2 \leq b_2.$$

Теорема 2. Для несравнимости двух интервалов $\tilde{a} = [a_1, a_2]$ и $\tilde{b} = [b_1, b_2]$, т.е. для того, чтобы они не находились ни в отношении $\tilde{a} \geq \tilde{b}$, ни в отношении $\tilde{a} \leq \tilde{b}$, необходимо и достаточно, чтобы одноименные границы интервалов удовлетворяли условиям

$$(12) \quad a_1 < b_1, a_2 > b_2 \quad \text{или} \quad b_1 < a_1, b_2 > a_2.$$

Теорема 3. Для существования в системе интервалов $\tilde{a}(1) = [a_1(1), a_2(1)]$, $\tilde{a}(2) = [a_1(2), a_2(2)]$,... максимального интервала необходимо и достаточно, чтобы его границы располагались относительно одноименных границ всех остальных интервалов согласно следующим условиям

$$(13) \quad a_1(1) \geq a_1(2), a_1(1) \geq a_1(3), \dots; \quad a_2(1) \geq a_2(2), a_2(1) \geq a_2(3), \dots$$

Условия-неравенства (13) записаны для конкретного случая, когда максимальным является интервал $\tilde{a}(1)$, что не ограничивает общности.

Теорема 4. Для существования в системе интервалов $\tilde{a}(1) = [a_1(1), a_2(1)]$, $\tilde{a}(2) = [a_1(2), a_2(2)]$,... минимального интервала необходимо и достаточно, чтобы его границы были расположены относительно одноименных границ всех остальных интервалов согласно условиям

$$(14) \quad a_1(1) \leq a_1(2), a_1(1) \leq a_1(3), \dots; \quad a_2(1) \leq a_2(2), a_2(1) \leq a_2(3), \dots$$

Условия (14), аналогично условиям (13), записаны для случая, когда минимальным является интервал $\tilde{a}(1)$, что не ограничивает общности.

Теоремы 1–4 примечательны тем, что сводят сравнение интервальных чисел к сравнению границ соответствующих интервалов.

3 Макроустойчивость задачи условной оптимизации

Обратимся к полностью определенной задаче условной оптимизации (3), (4) и опишем метод установления макроустойчивости этой задачи. Задача (3), (4) по определению (см. раздел 1) является макроустойчивой, если она сама и производная от нее интервальная задача условной оптимизации (5), (6) имеют решения. Существование решения полностью определенной задачи условной оптимизации (3), (4) обычно можно установить с помощью общеизвестных методов решения задач математического программирования, решая соответствующую задачу [11–13]. Сложнее обстоит дело с проверкой существования решения неполностью определенной (интервальной) задачи условной оптимизации (5), (6). Здесь эффективным оказывается применение детерминизационного метода решения задач интервальной оптимизации [2, 8, 14].

Интервальная задача условной оптимизации (5), (6) имеет интервальную целевую функцию $\tilde{F}(x_1, \dots, x_n)$, интервальные функции ограничений $\tilde{\Phi}_i, \overline{1, m}$, в левых частях ограничений и интервальные параметры $\tilde{b}_i, i = \overline{1, m}$, в правых частях. Используя формулы элементарных преобразований интервалов (8), функции $\tilde{F}$ и $\tilde{\Phi}_i$ можно представить явно в интервальной форме. Так же можно представить и параметры $\tilde{b}_i$. Все эти представления записываются в виде

$$(15) \quad \begin{aligned} \tilde{F}(x_1, \dots, x_n) &= [F_1(x_1, \dots, x_n), F_2(x_1, \dots, x_n)], \\ \tilde{\Phi}_i(x_1, \dots, x_n) &= [\Phi_{i1}(x_1, \dots, x_n), \Phi_{i2}(x_1, \dots, x_n)], \quad i = \overline{1, m}, \\ \tilde{b}_i &= [b_{i1}, b_{i2}], \quad i = \overline{1, m}. \end{aligned}$$

Алгоритм получения представлений (15) покажем на примере одной достаточно общей задачи оптимизации типа (5), (6).

Пример 1. Рассмотрим сначала частный случай общей интервальной задачи условной оптимизации (5), (6) – интервальную задачу линейного программирования:

$$\begin{aligned} \tilde{c}_1 x_1 + \dots + \tilde{c}_n x_n &= \max, \\ \tilde{a}_{i1} x_1 + \dots + \tilde{a}_{in} x_n &\leq \tilde{b}_i, \quad i = \overline{1, m}, \quad x_1 \geq 0, \dots, x_n \geq 0. \\ \tilde{c}_j &= [c_{j1}, c_{j2}], \quad j = \overline{1, n}; \quad \tilde{a}_{ij} = [a_{ij,1}, a_{ij,2}], \quad i = \overline{1, m}, \quad j = \overline{1, n}; \quad \tilde{b}_i = [b_{i1}, b_{i2}]. \end{aligned}$$

Здесь интервальная целевая функция $\tilde{F}(x_1, \dots, x_n) = \tilde{c}_1 x_1 + \dots + \tilde{c}_n x_n$, интервальные функции ограничений $\tilde{\Phi}_i(x_1, \dots, x_n) = \tilde{a}_{i1} x_1 + \dots + \tilde{a}_{in} x_n, i = \overline{1, m}$, и наконец, интервальные параметры $\tilde{b}_i$ в правых частях ограничений: $\tilde{b}_i = [b_{i1}, b_{i2}]$. Подставляя в выражения $\tilde{F}, \tilde{\Phi}_i$ параметры $\tilde{c}_j, \tilde{a}_{ij}$ в явной форме интервалов и умножая эти интервалы на неотрицательные переменные $x_1, \dots, x_n$, согласно формуле (8), получим необходимые явные интервальные представления вида (15) для рассматриваемой нами задачи линейного программирования

$$\begin{aligned} \tilde{F}(x_1, \dots, x_n) &= [F_1(x_1, \dots, x_n) = \tilde{c}_{11} x_1 + \dots + \tilde{c}_{n1} x_n, F_2(x_1, \dots, x_n) = \tilde{c}_{12} x_1 + \dots + \tilde{c}_{n2} x_n], \\ \tilde{\Phi}_i(x_1, \dots, x_n) &= [\Phi_{i1}(x_1, \dots, x_n) = \tilde{a}_{i1,1} x_1 + \dots + \tilde{a}_{i1,n} x_n, \Phi_{i2}(x_1, \dots, x_n) = \tilde{a}_{i2,1} x_1 + \dots + \tilde{a}_{i2,n} x_n], \quad i = \overline{1, m}, \\ \tilde{b}_i &= [b_{i1}, b_{i2}], \quad i = \overline{1, m}. \end{aligned}$$

Учитывая полученные представления (15), всю интервальную задачу (5), (6) также можно записать в явном интервальном виде

$$(16) \quad [F_1(x_1, \dots, x_n), F_2(x_1, \dots, x_n)] = \max,$$

$$(17) \quad [\Phi_{i1}(x_1, \dots, x_n), \Phi_{i2}(x_1, \dots, x_n)] \leq [b_{i1}, b_{i2}], \quad i = \overline{1, m}.$$

Согласно представлению (16), (17), задача (5), (6) заключается в том, чтобы найти максимум интервальной функции в области, ограниченной системой интервальных неравенств.

От интервального представления задачи (16), (17) перейдем к ее эквивалентному представлению в виде пары полностью определенных (детерминированных) задач условной оптимизации, которое уже поддается решению. Для этого сначала по теореме 3 представим интервальное уравнение (16) в виде эквивалентной пары детерминированных уравнений

$$(18) \quad F_1(x_1, \dots, x_n) = \max, \quad F_2(x_1, \dots, x_n) = \max.$$

Далее, по теореме 1 представим систему интервальных неравенств (17) в виде эквивалентной системы обычных детерминированных неравенств

$$(19) \quad \Phi_{i1}(x_1, \dots, x_n) \leq b_{i1}, \quad \Phi_{i2}(x_1, \dots, x_n) \leq b_{i2}, \quad i = \overline{1, m}.$$

Соединив пару уравнений оптимизации (18) с системой неравенств-ограничений (19), получаем совокупность двух полностью определенных задач условной оптимизации вида (3), (4)

$$(20) \quad \begin{cases} F_1(x_1, \dots, x_n) = \max, \\ \Phi_{i1}(x_1, \dots, x_n) \leq b_{i1}, \quad i = \overline{1, m}, \\ \Phi_{i2}(x_1, \dots, x_n) \leq b_{i2}, \quad i = \overline{1, m}, \end{cases}$$

$$(21) \quad \begin{cases} F_2(x_1, \dots, x_n) = \max, \\ \Phi_{i1}(x_1, \dots, x_n) \leq b_{i1}, \quad i = \overline{1, m}, \\ \Phi_{i2}(x_1, \dots, x_n) \leq b_{i2}, \quad i = \overline{1, m}, \end{cases}$$

эквивалентную исходной интервальной задаче условной оптимизации (5), (6). Задачу (20) будем называть нижней граничной задачей интервальной задачи (5), (6), а задачу (21) – ее верхней граничной задачей. Для получения решения интервальной задачи (5), (6) нам нужно решить ее нижнюю (20) и верхнюю (21) граничные задачи. В общем случае решение нижней граничной задачи имеет вид $\{M_n(x), F_{1,\max}\}$, а верхней граничной задачи – вид $\{M_v(x), F_{2,\max}\}$.

При этом $M_n(x)$, $M_v(x)$ – множества точек решения $x = (x_1, \dots, x_n)$ нижней и верхней граничных задач, а $F_{1,\max}$, $F_{2,\max}$ – полученные максимальные значения целевых функций этих задач. Решение интервальной задачи оптимизации (5), (6) формируется из решений ее нижней и верхней граничных задач и имеет вид

$$(22) \quad \{x^* \in M_n(x) \cap M_v(x); \quad \tilde{F}_{\max} = [F_{1,\max}, F_{2,\max}] \}.$$

Согласно (22), в качестве точки решения x^* интервальной задачи оптимизации (5), (6) выбирается любая точка из пересечения множеств точек решения ее нижней и верхней граничных задач, а в качестве максимального значения интервальной целевой функции $\tilde{F}_{\max}$ – интервал от максимального значения целевой функции нижней граничной задачи $F_{1,\max}$ до максимального значения целевой функции верхней граничной задачи $F_{2,\max}$.

Из выполненного процесса построения решения интервальной задачи вида (5), (6) и определения макроустойчивости полностью определенной задачи условной оптимизации (3), (4) вытекает следующая основная теорема.

Теорема 5. Для того чтобы полностью определенная задача условной оптимизации (3), (4) была макроустойчива, необходимо и достаточно, чтобы: 1) она имела решение; 2) интер-

вальная задача оптимизации (5), (6), производная от задачи (3), (4), имела нижнюю и верхнюю граничные задачи, обладающие решениями; 3) множества решений нижней граничной задачи и верхней граничной задачи интервальной задачи оптимизации (5), (6) пересекались.

Доказательство теоремы 5 содержится непосредственно в полученном общем решении (22) интервальной задачи условной оптимизации (5), (6), при учете определения макроустойчивости полностью определенной задачи оптимизации (3), (4).

Теорема 5 определяет следующий алгоритм для проверки произвольной полностью определенной задачи условной оптимизации (3), (4), на макроустойчивость.

Шаг 1. Применяя подходящие для конкретного типа целевой функции методы решения полностью определенных задач условной оптимизации [11–13], ищем решение $x' = (x'_1, \dots, x'_n)$ задачи (3), (4). Одновременно проверяется и существование (несуществование) решения задачи.

Шаг 2. Задаваясь некоторыми подходящими значениями интервальных параметров целевой функции F , функций ограничений $\Phi_i, i = \overline{1, m}$, и правых частей ограничений $b_i, i = \overline{1, m}$ полностью определенной задачи условной оптимизации (3), (4), строим производную от нее интервальную задачу условной оптимизации (5), (6).

Шаг 3. Используя формулы интервальной математики (8), выражающие результаты элементарных преобразований интервалов, представляем целевую функцию $\tilde{F}$, функции ограничений $\tilde{\Phi}_i, i = \overline{1, m}$, а также правые части ограничений $\tilde{b}_i, i = \overline{1, m}$, интервальной задачи условной оптимизации (5), (6) в интервальной форме (15).

Шаг 4. По найденным на шаге 3 интервальным представлениям функций $\tilde{F}, \tilde{\Phi}_i, i = \overline{1, m}$, и параметров $\tilde{b}_i, i = \overline{1, m}$, формируем нижнюю (20) и верхнюю (21) граничные задачи интервальной задачи условной оптимизации (5), (6).

Шаг 5. Используя те же самые методы, что и на шаге 1, ищем решения оптимизационных задач (20) и (21). Одновременно с этим проверяем существование или несуществование решений указанных задач. Полные решения задач условной оптимизации вида (20), (21) имеют соответственно вид $\{M_n(x), F_{1, \max}\}, \{M_v(x), F_{2, \max}\}$, где $M_n(x)$ – множество точек x решения нижней, $M_v(x)$ – множество точек x решения верхней граничной задачи.

Шаг 6. Проверяется наличие (отсутствие) пересечения найденных в результате решения задач (20) и (21) множеств $M_n(x), M_v(x)$.

Итог. Если в результате работы алгоритма выяснилось, что полностью определенная задача условной оптимизации (3), (4) имеет решение, а производная от нее интервальная задача условной оптимизации (5), (6) имеет нижнюю и верхнюю граничные задачи, обладающие решениями, причем множества этих решений пересекаются, то задача оптимизации (3), (4) является макроустойчивой. В противном случае задача (3), (4) не является макроустойчивой.

Пример 2 (задача о назначениях). Имеется 3 работы и 3 исполнителя. Заданы доходы a_{ij} от выполнения любой j -й работы любым i -м исполнителем ($i, j = \overline{1, 3}$). Требуется распределить работы между исполнителями так, чтобы каждый из них выполнял ровно одну работу, и, кроме того, суммарный доход от выполнения всех работ был максимальным. Введя множество неизвестных матриц назначений $X = \|x_{ij}\|, x_{ij} \in \{0, 1\}$, где $x_{ij} = 1$, если i -й исполнитель выполняет j -ю работу, и $x_{ij} = 0$ в противном случае, задачу можно записать математически в виде

$$F(x_{ij}) \equiv \sum_{i=1}^3 \sum_{j=1}^3 a_{ij} x_{ij} = \max, \text{ при условии } \Phi_1(x_{ij}) \equiv \sum_{i=1}^3 x_{ij} = 1, j = \overline{1, 3}; \Phi_2(x_{ij}) \equiv \sum_{j=1}^3 x_{ij} = 1, i = \overline{1, 3}.$$

Видим, что наша задача – частный случай полностью определенной задачи условной оптимизации (3), (4). Проверим эту задачу на макроустойчивость, используя изложенный алгоритм.

Шаг 1. Для определенности конкретизируем матрицу доходов $A = \|a_{ij}\|$ в виде матрицы с точно известными параметрами

$$A = \begin{vmatrix} 2 & 3 & 3 \\ 4 & 4 & 3 \\ 3 & 4 & 4 \end{vmatrix}$$

и решим нашу задачу при этих условиях. Имеется 6 различных матриц назначений X , удовлетворяющих ограничениям задачи:

$$X_1 = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix}, X_2 = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}, X_3 = \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix}, X_4 = \begin{vmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix}, X_5 = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix}, X_6 = \begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix},$$

которым соответствуют значения целевой функции $F_1=9, F_2=10, F_3=11, F_4=11, F_5=9, F_6=10$. Так что решение задачи существует, достигается на матрицах назначений X_3, X_4 и равно

$$F_{\max} = F_{X_3, X_4} = 11.$$

Шаг 2. Согласно описанию данного шага алгоритма (см. выше), задаемся подходящими значениями интервальных параметров целевой функции F нашей недетерминированной задачи о назначениях в виде заданной неполностью (с точностью до интервалов возможных значений) матрицы доходов $\tilde{A} = \|\tilde{a}_{ij}\| = \| [a_{1ij}, a_{2ij}] \|$:

$$\tilde{A} = [A_1, A_2], \text{ где } A_1 = \|\tilde{a}_{1ij}\| = \begin{vmatrix} 1 & 2 & 2 \\ 1 & 2 & 2 \\ 2 & 2 & 2 \end{vmatrix}, A_2 = \|\tilde{a}_{2ij}\| = \begin{vmatrix} 3 & 4 & 4 \\ 5 & 5 & 4 \\ 4 & 5 & 5 \end{vmatrix}.$$

Имеем производную от решенной полностью определенной задачи условной оптимизации интервальную задачу условной оптимизации типа (5), (6)

$$\tilde{F}(x_{ij}) \equiv \sum_{i=1}^3 \sum_{j=1}^3 \tilde{a}_{ij} x_{ij} = \max,$$

при тех же самых условиях-ограничениях, которые существовали и для полностью определенной задачи условной оптимизации.

Шаг 3. С помощью формул (8) элементарных преобразований интервалов представляем целевую функцию производной задачи $\tilde{F}$ в интервальной форме (15)

$$\tilde{F}(x_{ij}) = \left[F_1(x_{ij}) \equiv \sum_{i=1}^3 \sum_{j=1}^3 a_{1ij} x_{ij}, F_2(x_{ij}) \equiv \sum_{i=1}^3 \sum_{j=1}^3 a_{2ij} x_{ij} \right].$$

Шаг 4. По найденному на шаге 3 интервальному представлению целевой функции $\tilde{F}$ и заданным условиям-ограничениям формируем нижнюю (20) и верхнюю (21) граничные задачи исходной интервальной задачи условной оптимизации

$$\left. \begin{aligned} F_1(x_{ij}) &\equiv \sum_{i=1}^3 \sum_{j=1}^3 a_{1ij} x_{ij} = \max, \\ \text{при условии } \Phi_1(x_{ij}) &\equiv \sum_{i=1}^3 x_{ij} = 1, j = \overline{1,3}; \Phi_2(x_{ij}) \equiv \sum_{j=1}^3 x_{ij} = 1, i = \overline{1,3} \end{aligned} \right\}; \left. \begin{aligned} F_2(x_{ij}) &\equiv \sum_{i=1}^3 \sum_{j=1}^3 a_{2ij} x_{ij} = \max, \\ \text{при тех же самых условиях} \end{aligned} \right\}.$$

Шаг 5. Тем же методом, что и на шаге 1, находим решения нижней и верхней граничных задач. В нашем случае решение нижней граничной задачи существует, достигается на матрицах назначений X_1, X_2, X_3, X_4 и равно $F_{1,\max} = F_{1(X_1, X_2, X_3, X_4)} = 5$. Решение верхней граничной

задачи тоже существует, достигается на матрицах назначений X_3, X_4 и равно $F_{2,\max} = F_{2(X_3, X_4)} = 14$.

Шаг 6. Проверяем наличие пересечения множеств точек решения нижней и верхней граничных задач интервальной задачи

$$M_H \cap M_B = \{X_1, X_2, X_3, X_4\} \cap \{X_3, X_4\} = \{X_3, X_4\} \neq \emptyset, \text{ т.е. пересечение непусто.}$$

Итог. Исходная полностью определенная задача условной оптимизации типа (3), (4) имеет решение. Производная от нее интервальная задача типа (5), (6) имеет нижнюю и верхнюю граничные задачи, обладающие решениями, причем множества точек решения этих задач пересекаются. Таким образом, заданная полностью определенная задача условной оптимизации типа (3), (4) является макроустойчивой.

4 Микроустойчивость задачи условной оптимизации

Снова обратимся к полностью определенной задаче условной оптимизации (3), (4) и опишем метод установления ее микроустойчивости. Из определения микроустойчивости (раздел 1) вытекает следующий алгоритм проверки задачи (3), (4) на микроустойчивость.

Шаг 1. С помощью 6-шагового алгоритма, изложенного в п. 4, проверяем задачу (3), (4) на макроустойчивость. В случае отрицательного результата (задача (3), (4) не макроустойчива) конец алгоритма, с выводом: задача (3), (4) не является микроустойчивой. При положительном результате проверки (задача (3), (4) макроустойчива) переход к шагу 2.

Шаг 2. Выбираем некоторую произвольную точку решения $x' = (x'_1, \dots, x'_n)$ задачи (3), (4), найденную на шаге 1. После этого добавляем к ней какую-либо точку решения $x'' = (x''_1, \dots, x''_n)$ соответствующей интервальной задачи (5), (6), также найденную на шаге 1. В результате получаем пару решений (x', x'') указанных двух задач.

Шаг 3. Вычисляем величину расстояния $D(x', x'')$ между точками решения x', x'' указанных двух задач, используя для этого формулу

$$(23) \quad D(x', x'') = \sqrt{(x'_1 - x''_1)^2 + \dots + (x'_n - x''_n)^2}.$$

Шаг 4. Проверяем выполнение неравенства, сравнивающего расстояние $D(x', x'')$ с некоторой изначально заданной достаточно малой величиной d :

$$(24) \quad D(x', x'') \leq d,$$

Если условие (24) выполнено, задача оптимизации (3), (4) объявляется микроустойчивой и конец алгоритма. В противном случае совершается переход к шагу 2, в котором теперь к точке решения $x' = (x'_1, \dots, x'_n)$ задачи (3), (4), найденной на шаге 1, добавляется какая-то другая точка решения $x'' = (x''_1, \dots, x''_n)$ задачи (5), (6) из числа найденных на шаге 1. В результате получаем новую пару решений (x', x'') и т.д.

Итог. Если в результате работы алгоритма после некоторого достаточного числа шагов получена пара решений (x', x'') , удовлетворяющая неравенству (24), процедура останавливается и задача (3), (4) объявляется микроустойчивой. В противном случае процедура также останавливается, но задача (3), (4) признается не обладающей свойством микроустойчивости.

Заключение

В статье показано, что проблема оптимизации неполностью определенных функций не может ограничиться только отысканием точки оптимума и значения в ней нашей функции, но и должна включать в себя задачу определения устойчивости найденного оптимума. Последнее означает, что при небольшом варьировании параметров оптимизируемой функции ее оптимум должен по-прежнему существовать и находиться в точке, близкой к точке исходного оптимума. Для установления устойчивости оптимума неполностью определенных функций предложена специальная эффективная методика, которая основана на аппарате интервальной математики. Несколько иные подходы к решению рассмотренной проблемы можно найти в [16–24].

Список источников

- [1] *Первозванский, А.А.* Математические модели в управлении производством / А.А. Первозванский. - М.: Наука, 1975 – 616 с.
- [2] *Левин, В.И.* Интервальное дискретное программирование / В.И. Левин // Кибернетика и системный анализ. – 1994. – №6. - С. 91–103.
- [3] *Libura, M.* Integer Programming Problems with Inexact Objective Function / M. Libura // Control and Cybernetics. - 1980. - Vol. 9. - №4. – P. 189–202.
- [4] *Тимохин, С.Г.* О задачах линейного программирования в условиях неточных данных / С.Г. Тимохин, А.В. Шапкин // Экономика и математические методы. - 1981 – Т. 17. - №5. - С. 955–963.
- [5] *Роцин, В.А.* Вопросы решения и исследования одного класса задач неточного целочисленного программирования / В.А. Роцин, Н.В. Семенова, И.В. Сергиенко // Кибернетика. - 1989 – № 2. - С. 42–46.
- [6] *Семенова Н.В.* Решение одной задачи обобщенного целочисленного программирования / Н.В. Семенова // Кибернетика. – 1984. - № 5. – С. 25–31.
- [7] *Вошинин, А.П.* Оптимизация в условиях неопределенности / А.П. Вошинин, Г.Р. Сотиров. – М.: Изд-во МЭИ, 1989. – 224 с.
- [8] *Левин, В.И.* Интервальные методы оптимизации систем в условиях неопределенности / - Пенза: Изд-во Пензенского технологического института, 1999. – 95 с.
- [9] *Левин, В.И.* Оптимальное проектирование в условиях неопределенности. Метод детерминизации / В.И. Левин // Онтология проектирования. – 2013. - №3. – С. 41–52.
- [10] *Алефельд, Г.* Введение в интервальные вычисления / Г. Алефельд, Ю. Херцбергер. - М.: Мир, 1987 – 360 с.
- [11] *Юдин, Д.Б.* Задачи и методы линейного программирования / Д.Б. Юдин, Е.Г. Гольдштейн. - М.: Советское радио, 1964 – 350 с.
- [12] *Корбут, А.А.* Дискретное программирование / А.А. Корбут, Ю.Ю. Финкельштейн. - М: Наука, 1969 – 280 с.
- [13] *Левин, В.И.* Структурно-логические методы исследования сложных систем / В.И. Левин. - М.: Наука, 1987 – 304 с.
- [14] *Левин, В.И.* Дискретная оптимизация в условиях интервальной неопределенности / В.И. Левин // Автоматика и телемеханика. – 1992. - № 7. – С. 97–106.
- [15] *Левин, В.И.* Нелинейная оптимизация в условиях интервальной неопределенности / В.И. Левин // Кибернетика и системный анализ. – 1999. – № 2. - С. 35–46.
- [16] *Шашихин, В.Н.* Оптимизация интервальных систем / В.Н. Шашихин // Автоматика и телемеханика. – 2000. - №11. – С. 17–24.
- [17] *Ащепков, Л.Т.,* Универсальные решения интервальных задач оптимизации и управления / Л.Т. Ащепков, Д.В. Давыдов. - М.: Наука, 2006. – 285 с.
- [18] *Островский, Г.М.* Технические системы в условиях неопределенности. Анализ гибкости и оптимизация / Г.М. Островский, Ю.М. Волин. - М.: Бином, 2008. – 325 с.
- [19] *Боргест, Н.М.* Автоматизация предварительного проектирования самолета / Н.М. Боргест. - Самара: Изд-во КуАИ, 1992. – 92 с.
- [20] *Островский, Г.М.* Оптимизация технических систем / Г.М. Островский, Н.Н. Зиятдинов, Т.В. Лаптева. - М.: Кнорус, 2012. – 252 с.
- [21] *Tsoukias, A.* A Characterization of PQI Interval Orders / A. Tsoukias, P. Vincke // Discrete Applied Mathematics. – 2003. - №127(2). – P. 387–397.
- [22] *Davydov, D.V.* Identification of Parameters of Linear Interval Controllable Systems with the Interval Observation / D.V. Davydov // Journal of Computer and Systems Sciences International. - 2008. - Vol. 47. - №6. – P. 861–865.
- [23] *Ozturk, M.* Positive and Negative Reasons in the Interval Comparisons: Valued PQI Interval Orders / M. Ozturk, A. Tsoukias // LAMSADE-CNRS: Universite Paris Dauphine. 2004. – 7 p.

- [24] **Shashihin, V.N.** Solution of Interval Matrix Game / V.N. Shashihin // Journal of Computer and Systems Sciences International. – 2001. – Vol. 40. – №5. – P. 97–104.

STABILITY OF UNCERTAIN OPTIMIZATION PROBLEMS

V.I. Levin

Penza State Technological University, Penza, Russia
 vilevin@mail.ru

Abstract

The problem of optimization of not completely defined functions, i.e. functions with parameters set only up to interval is considered. A review of existing approaches to solving optimization problems of incompletely specified functions with different kinds of uncertainty is given. The mathematical formulation of the problem of optimizing the functions with interval parameters and the method of its solution by reducing it to two optimization problems of completely defined functions i.e. functions with exactly known parameters (the determination method) is described. It is shown that the solution of the optimization problem of not fully defined functions also requires consideration of the problem of determining the optimum stability to variations in the values of the function parameters. In this regard, we introduce notions of macrostability and microstability of optimization problem of fully defined functions. Necessary and sufficient conditions for the macrostability of optimization problem of completely specified functions are given. An algorithm for checking of macrostability is presented. An example of checking macrostability of particular problem with this algorithm (assignment problem) is given. We also present an algorithm for checking microstability of optimization problem of fully defined function. To solve such problems methods of interval mathematics are used.

Key words: *problem of system optimization, uncertainty, stability of the optimum, variation of parameters, interval mathematics, macrostability, microstability.*

References

- [1] **Pervozvanskij A.A.** Matematicheskie modeli v upravlenii proizvodstvom [Mathematical models in production management] / A.A. Pervozvanskij. - Moscow: Nauka, 1975. - 616 p. (In Russian)
- [2] **Levin, V.I.** Interval'noe diskretnoe programmirovaniye [Interval discrete programming] / V.I. Levin // Kibernetika i sistemnyj analiz. – 1994. – No. 6. – P. 91–103. (In Russian)
- [3] **Libura, M.** Integer Programming Problems with Inexact Objective Function / M. Libura // Control and Cybernetic. - 1980. - Vol. 9. - №4. – P. 189–202.
- [4] **Timokhin, S.G.** O zadachakh linejnogo programmirovaniya v usloviyakh netochnykh dannykh [On the linear programming problem under conditions of inaccurate data] / S.G. Timokhin, A.V. Shapkin // Ekonomika i matematicheskie metody. - 1981. - Vol. 17. - No. 5. - P. 955–963. (In Russian)
- [5] **Roshhin, V.A.** Voprosy resheniya i issledovaniya odnogo klassa zadach netochnogo tselochislennogo programmirovaniya [Questions and research solutions of a class of non-precision integer programming] / V.A. Roshhin, N.V. Semenova, I.V. Sergienko // Kibernetika. - 1989. - No. 2. - P. 42–46. (In Russian)
- [6] **Semenova, N.V.** Reshenie odnoj zadachi obobshhennogo tselochislennogo programmirovaniya [Solution to one of the problems of generalized integer programming] / N.V. Semenova // Kibernetika. - 1984. – No. 5. - P. 25–31. (In Russian)
- [7] **Voschinin, A.P.** Optimizatsiya v usloviyakh neopredelennosti [Optimization under uncertainty] / A.P. Voschinin, G.R. Sotirov. – Moscow: MEI publ., 1989. – 224 p. (In Russian)
- [8] **Levin, V.I.** Interval'nye metody optimizatsii sistem v usloviyakh neopredelennosti [Interval methods of optimization of systems under uncertainty] / V.I. Levin. Penza: Penza technological institute publ., 1999. - 95 p. (In Russian)
- [9] **Ahlefeld, G.** Vvedenie v interval'nye vychisleniya [Introduction to interval computation] / G. Ahlefeld, J. Herzberger. - Moscow: Mir, 1987. - 360 p. (In Russian)
- [10] **Yudin, D.B.** Zadachi i metody linejnogo programmirovaniya [Objectives and methods of linear programming] / D.B. Yudin, E.G. Gol'dshtejn. - Moscow: Sovetskoe radio, 1964. - 350 p. (In Russian)

- [11] **Korbut, A.A.** Diskretnoe programmirovaniye [Discret programming] / A.A. Korbut, Yu.Yu. Finkel'shtejn. - Moscow: Science, 1969. - 280 p. (In Russian)
- [12] **Levin, V.I.** Strukturno-logicheskie metody issledovaniya slozhnykh sistem [Structural-logical methods for complex systems research] / V.I. Levin. - Moscow: Nauka, 1987. - 304 p. (In Russian)
- [13] **Levin, V.I.** Diskretnaya optimizatsiya v usloviyakh interval'noj neopredelennosti [Discrete optimization under interval uncertainty] / V.I. Levin // Avtomatika i telemekhanika. - 1992. - No 7. - P. 97–106. (In Russian)
- [14] **Levin, V.I.** Nelineynaya optimizatsiya v usloviyakh interval'noj neopredelennosti [Nonlinear optimization under interval uncertainty] / V.I. Levin // Kibernetika i sistemnyy analiz. - 1999. - No. 2. – P. 35–46. (In Russian)
- [15] **Shashikhin, V.N.** Optimizatsiya interval'nykh sistem [Interval systems optimization] / V.N. Shashikhin // Avtomatika i telemekhanika. - 2000. - No. 11. – P. 17–24. (In Russian)
- [16] **Ashhepkov, L.T.** Universal'nye resheniya interval'nykh zadach optimizatsii i upravleniya [Universal solutions interval optimization problems and management] / L.T. Ashhepkov, D.V. Davydov. - Moscow: Science, 2006. - 285 p. (In Russian)
- [17] **Ostrovskij, G.M.** Tekhnicheskie sistemy v usloviyakh neopredelennosti. Analiz gibkosti i optimizatsiya. [Technical systems under conditions of uncertainty. Flexibility analysis and optimization] / G.M. Ostrovskij, Yu.M. Volin. - Moscow: Binom, 2008. - 325 p. (In Russian)
- [18] **Borgest, N.M.** Avtomatizatsiya predvaritel'nogo proektirovaniya samoleta. Uchebnoe posobie. [Automation of the preliminary design of the aircraft. Textbook.] / N.M. Borgest. – Samara: Samara aviation institute publ., 1992. – 92 p. (In Russian)
- [19] **Ostrovskij, G.M.** Optimizatsiya tekhnicheskikh sistem [Optimization of technical systems] / G.M. Ostrovskij, N.N. Ziyatdinov, T.V. Lapteva. - Moscow: Knorus, 2012. – 252 p.(In Russian)
- [20] **Tsoukias, A.** A Characterization of PQI Interval Orders / A. Tsoukias, P. Vincke // Discrete Applied Mathematics. – 2003. - №127(2). – P. 387–397.
- [21] **Davydov, D.V.** Identification of Parameters of Linear Interval Controllable Systems with the Interval Observation / D.V. Davydov // Journal of Computer and Systems Sciences International. - 2008. - Vol. 47. - №6. – P. 861–865.
- [22] **Ozturk, M.** Positive and Negative Reasons in the Interval Comparisons: Valued PQI Interval Orders / M. Ozturk, A. Tsoukias // LAMSADE-CNRS: Universite Paris Dauphine. 2004. – 7 p.
- [23] **Shashihin, V.N.** Solution of Interval Matrix Game / V.N. Shashihin // Journal of Computer and Systems Sciences International. – 2001. - Vol. 40. - №5. – P. 97–104.

Сведения об авторе



Левин Виталий Ильич окончил Каунасский политехнический институт, Открытый университет Израиля. Доктор технических наук, профессор, PhD, Full Professor, заведующий кафедрой математики (1975–2000), советник ректора по науке (2006–2011) Пензенского государственного технологического университета, профессор Московского университета им. С.Ю. Витте (с 2003 г.). В списке научных трудов сотни работ (в том числе десятки монографий) по логике; математическому моделированию в технике, экономике, социологии, принятию решений; оптимизации; теории надежности; истории науки; проблемам образования. Действительный член МАИ, ЕАИ, МАНЭБ и АСН, заслуженный деятель науки Российской Федерации, лауреат международных премий «Соросовский профессор», Международный эксперт в области социологии конфликта и рейтингования университетов.

Vitaly Ilyich Levin graduated from Kaunas Polytechnical Institute, Open University of Israel, Doctor of Engineering Science, Professor, PhD, Full Professor. Head of Mathematics Department (1975–2000), the scientific counselor of rector (2006–2011) of Penza State Technological University, professor of Moscow University named after S.J. Vitte (since 2003). He is the autor of hundreds of publications (among them dozens of monographs), in logic; mathematical modelling of engineering, economics, sociology; optimization; reliability; history of science; education problems. The member of IIA, EIA, IAELP, ASS. Honoured scientist of Russia, Laureate of International Prizes “Soros Professor”, International Reviewer in Sociology of conflicts and University ranking.

УДК 519.5

ПРОСТОЙ И УНИВЕРСАЛЬНЫЙ МЕТОД ПРИНЯТИЯ РЕШЕНИЙ В ПРОСТРАНСТВЕ КРИТЕРИЕВ «СТОИМОСТЬ–ЭФФЕКТИВНОСТЬ»

С.А. Пиявский

Самарский государственный архитектурно-строительный университет, Самара, Россия
spiyav@mail.ru

Аннотация

В онтологию принятия многокритериальных решений вводится новое понятие шансов оптимальности решения. Его плодотворность иллюстрируется на практически значимой задаче сравнения альтернатив в пространстве критериев «стоимость–эффективность», в которой это понятие позволяет получить комплексную количественную оценку качества решений без обязательного ввода какой-либо дополнительной информации от ЛПР, кроме, при его желании, «расплывчатых» суждений типа «лучше» или «важнее». При этом показано, что единственной непрерывной оценочной функцией, позволяющей выявить в любом наборе многокритериальных альтернатив все Парето-оптимальные решения, является минимаксная свёртка Ю.Б. Гермейера. Мера множества значений неопределённых коэффициентов этой свёртки, при которых оценочная функция принимает на некотором решении лучшее значение, чем на прочих решениях из исходного набора альтернатив, определяет шанс оптимальности этого решения. Разработаны простой расчётный алгоритм и наглядный геометрический метод вычисления шансов оптимальности.

Ключевые слова: многокритериальная оптимизация, принятие решений, ЛПР, шансы оптимальности, альтернатива, неопределённость, Парето-оптимальность, стоимость, эффективность.

Введение

Понятийный аппарат, используемый в онтологии принятия оптимальных решений, непрерывно растёт. Это вызвано увеличением сложности решаемых практических задач и связанным с этим расширением фронта теоретических исследований. В дополнение к введённым нами в [1] и [2] понятиям *расширенной Парето-оптимальности*, *прогрессивных решений*, *уровня прогрессивности решений* и *уверенных суждений* лица, принимающего решение (ЛПР) в настоящей статье предлагается новое понятие **шансов оптимальности многокритериальных альтернатив**. Это понятие, по мнению автора, имеет прозрачный смысл, доступный пониманию ЛПР с любым уровнем математической подготовки. В то же время оно позволяет уверенно решать слабо формализованные задачи многокритериального выбора при полном отсутствии дополнительно вводимой информации. Понятие применимо в задачах любой размерности и степени неопределённости, однако в настоящей статье оно иллюстрируется на простой и потому наглядной задаче двухкритериального выбора, что позволяет показать его ценность.

1 Постановка задачи

Пусть задано m вариантов решений, характеризуемых количественными значениями двух критериев: стоимости и эффективности. Обозначим через $i = 1, \dots, m$ номер варианта решения, а через c_i , e_i - значения для него количественных критериев стоимости и эффективности. Учитывая, что содержательный характер вариантов решения в данной постановке не

имеет значения, будем для краткости называть вариантом решения вектор его критериев эффективности.

Поставим задачу разработать простой, универсальный, максимально обоснованный метод, который позволяет ЛПР, не привлекая никакой дополнительной информации, получить объективную скалярную количественную оценку степени оптимальности каждого варианта решения; при этом, при желании, ЛПР может уточнять постановку задачи своими «расплывчатыми» суждениями типа «лучше» или «важнее».

2 Описание метода

Благодаря простоте применения, предлагаемый метод может быть использован «вручную» любыми лицами, заинтересованными в объективной сравнительной оценке различных вариантов решений по критериям «стоимость–эффективность», даже если они владеют только элементарной математикой.

2.1 Выбор наилучшего из нескольких вариантов решения

Пусть требуется выбрать наилучшее решение из шести возможных решений, представленных в таблице 1. Содержательный смысл решений может быть любой (это различные проекты танков или варианты городов, в которых планируется провести отпуск и т.д.), существенны лишь две характеристики каждого решения, представленные во втором и третьем столбцах таблицы 1: стоимость (затраты на его реализацию) и эффект, который будет при этом получен.

Таблица 1 – Исходные данные примера

Вариант решения	Стоимость (у.е.)	Эффективность (у.е.)	Нормированная стоимость	Нормированный дефект (потери)
1	2	3	4	5
А	100	16	0	1
Б	165	20	0,65	0,6
В	120	17	0,2	0,9
Г	200	26	1	0
Д	175	23	0,75	0,3
Е	150	18	0,5	0,8
min	100	16	0	0
max	200	26	1	1

Сложность выбора наилучшего решения для заинтересованного в этом человека или организации (ЛПР) состоит в том, что не существует единственного «объективного» способа сравнения между собой двух вариантов решений, один из которых, например, по стоимости лучше, а по эффективности хуже другого. Таких способов, в одинаковой степени разумных, бесконечно много. Разные способы будут приводить к разным результатам выбора.

Идея предлагаемого нами метода состоит в том, чтобы перебрать (теоретически) все «разумные» способы сравнения, при каждом из них определить наилучшее решение и окончательно принять за наилучшее решение то, за которое «проголосует» большинство перебираемых способов. Реализовать эту идею удалось весьма простым алгоритмом, который изложим по шагам.

Шаг 1. Исключение заведомо неоптимальных вариантов решений. Проверим, что в таблицу 1 не включён ни один заведомо неоптимальный вариант решения, т.е. такой, который по обоим показателям уступает каким-либо другим вариантам решений из таблицы.

Шаг 2. Нормализация исходных данных. Найдём в каждом из столбцов 2 и 3 таблицы 1 наибольшее и наименьшее значения (**min** и **max**), а затем рассчитаем для каждого варианта решения две новые характеристики, *по каждой из которых ЛПР стремится к минимуму*: нормированную стоимость и нормированный дефект (потери), отражающие относительные значения стоимости и эффективности.

Нормированная стоимость_i =

$$\frac{(\text{стоимость}_i - \text{минимальная стоимость})}{(\text{максимальная стоимость} - \text{минимальная стоимость})}$$

Нормированный дефект_i =

$$1 - \frac{(\text{эффективность}_i - \text{минимальная эффективность})}{(\text{максимальная эффективность} - \text{минимальная эффективность})}$$

Результаты расчёта показаны в 4-м и 5-м столбцах таблицы 1.

Шаг 3. Расчёт индексов стоимости и дефекта. Построим таблицу 2, расположив в ней результаты расчётов в порядке возрастания нормированной стоимости и перенумеровав их в этом порядке (при этом значения нормированного дефекта автоматически расположатся в порядке убывания).

Таблица 2 – Нормализованные исходные данные и расчётные параметры примера

Номер варианта решения	Нормированная стоимость (с)	Нормированный дефект (d)	Индекс стоимости	Индекс дефекта = 1/индекс стоимости	Интервалы индексов	Шанс оптимальности варианта решения
1	2	3	4	5	6	7
1 (А)	0	1	0,2	5	0,2	10
2 (В)	0,2	0,9	0,555	1,8	0,355	18
3 (Е)	0,5	0,8	0,812	1,231	0,257	13
4 (Б)	0,65	0,6	1,25	0,8	0,388	19
5 (Д)	0,75	0,3	3,333	0,3	0,5	25
6 (Г)	1	0		0	0,3	15

Полученная задача принятия решения показана на рисунке 1. По каждому из критериев ЛПР стремится к минимизации.

Заполним столбец с названием «Индекс стоимости». В первую его строчку занесем частное от деления нормированной стоимости варианта решения 2 на нормированные потери предшествующего варианта решения (с номером 1) и так далее вплоть до предпоследней клетки столбца, как показано линиями на таблице 2.

В ячейки столбца «Индекс дефекта» таблицы 2 внесем результаты деления единицы на числа, стоящие в тех же строках столбца «Индекс стоимости». В последнюю клетку столбца «Индекс дефекта» внесем ноль.

Шаг 4. Расчёт шансов оптимальности вариантов решений. Сначала заполним столбец 6 таблицы 2, для удобства прежде выделив в столбцах 4 и 5 числа, не большие единицы.

В первую строку столбца 6 занесем число из первой строки индекса стоимости. Затем, идя от первой строки столбца 6 *вниз*, будем вносить в очередную ячейку *разность* отмеченных подчеркиванием индексов стоимости, стоящих в текущей и последующей (с большим

номером) клетках индекса стоимости (столбец 4). Так, во вторую строку занесем число 0,355 как разность между 0,555 и 0,2. Продолжаем этот процесс, пока не закончатся отмеченные подчёркиванием числа.

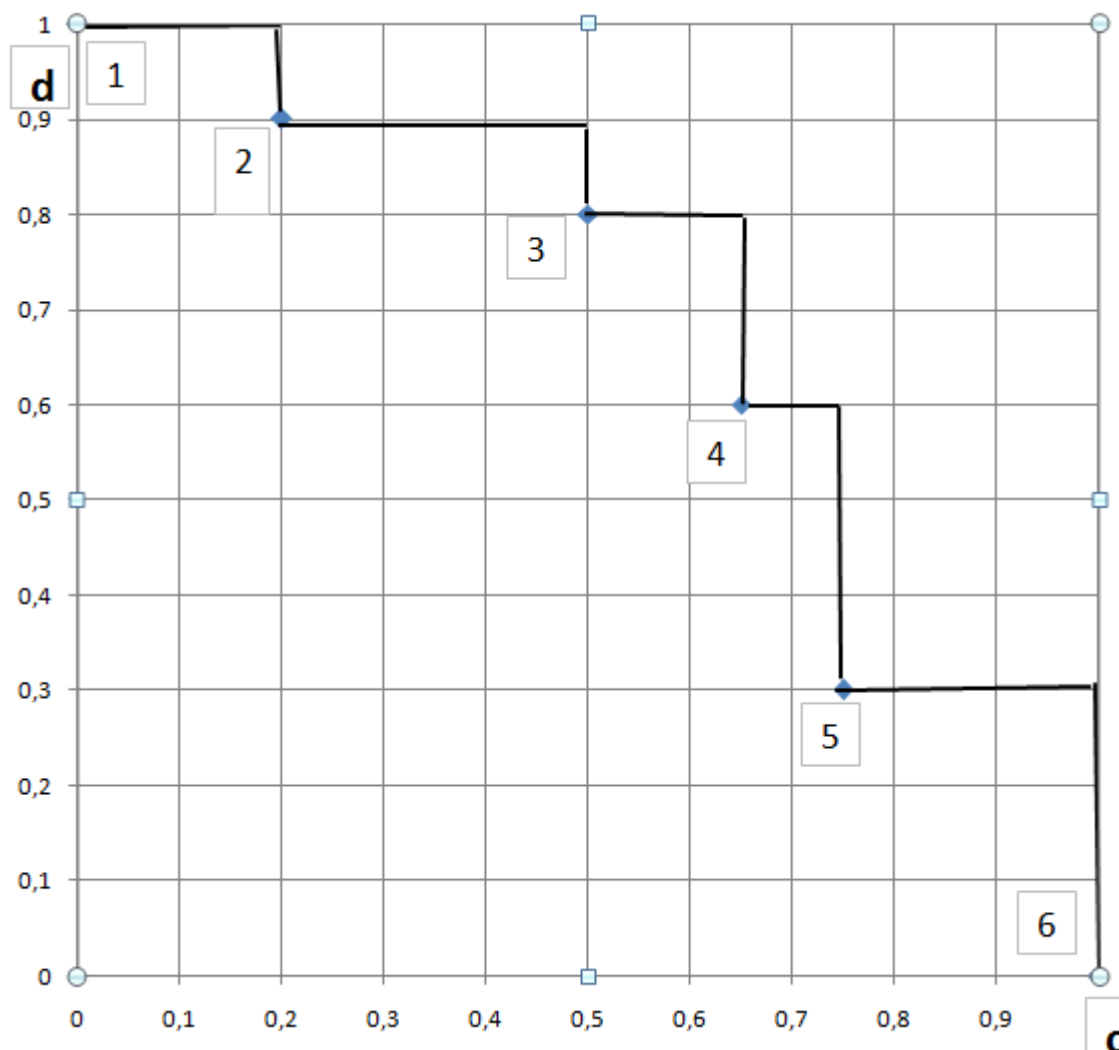


Рисунок 1 – Пример задачи выбора решений (варианты 1–6) в критериях «нормированная стоимость–нормированный дефект (эффективность)»

Аналогично, идя от последней строки столбца 6 *вверх*, будем вносить в очередную ячейку *разность* отмеченных подчёркиванием индексов дефекта, стоящих в предыдущей (с меньшим номером) и текущей ячейках индекса дефекта (столбец 5). Так, в последнюю строку занесём число 0,3 как разность между 0,3 и 0. Продолжаем этот процесс, пока не закончатся отмеченные подчёркиванием числа.

В редких случаях процесс завершится полным заполнением всех клеток столбца 6, тогда в ячейке, в которой «встретились» оба процесса заполнения (сверху и снизу), окажется единица. Как правило же, одна клетка окажется незаполненной. В нее нужно внести число, которое дополняет сумму всех чисел столбца 6 до числа 2. В нашем примере это число равно

$$2 - 0,2 - 0,355 - 0,257 - 0,5 - 0,3 = 0,388.$$

Теперь заполним последний столбец таблицы 2 (столбец 7), перенеся в него числа столбца 6, умноженные на 50 (рекомендуется округлить их до ближайших целых чисел так, чтобы сумма всех чисел в столбце 7 составила 100).

Числа, стоящие в столбце 7 таблицы 2, характеризуют целесообразность выбора соответствующих им решений в качестве оптимальных. В рассматриваемом примере, таким образом, наилучшим решением является решение 5 (Д).

2.2 Выбор наилучшего решения, когда один показатель «важнее» другого

Возможно, ЛПР может с уверенностью уточнить свое понимание ситуации выбора, считая, что один из показателей «важнее» другого, например, стоимость решения для него имеет большее значение, чем эффективность (при этом не требуется пытаться определить, «насколько» важнее, это никогда невозможно с уверенностью сделать).

В этом случае в столбце «Интервалы индексов» следует оставить лишь те числа, которые получены в результате «регулярной вниз/вверх» работы с выделенными подчеркиванием числами в столбце индексов «*важнейшего*» показателя (поэтому, когда эффективность «важнее», подчеркнутое значение 0,8 в столбце «Индекс дефекта» не будет оставлено). Затем следует вычислить сумму этих чисел и записать разность единицы и этой суммы в ячейке, ближайшей к оставленным числам этого столбца. В остальные ячейки столбца вписать нули. Окончательно в столбец «Шанс оптимальности вариантов решений» вписать числа из столбца «Интервалы индексов», умноженные на 100. Как показано в таблицах 3 и 4, если стоимость важнее эффективности, наилучшим оказывается решение 2 (В), а если эффективность «важнее» стоимости – по-прежнему, решение 5 (Д), но шансы его признания оптимальным вдвое возрастают по сравнению со случаем, когда показатели представлялись ЛПР эквивалентными (50% против 25%).

Таблица 3 – Вычисление шансов оптимальности решений в примере, если стоимость «важнее» эффективности

Номер варианта решения	Интервалы индексов	Шанс оптимальности варианта решения
1 (А)	0,2	20
2 (В)	0,355	36
3 (Е)	0,257	26
4 (Б)	0,188	19
5 (Д)	0	0
6 (Г)	0	0

Таблица 4 – Вычисление шансов оптимальности решений в примере, если эффективность «важнее» стоимости

Номер варианта решения	Интервалы индексов	Шанс оптимальности варианта решения
1 (А)	0	0
2 (В)	0	0
3 (Е)	0	0
4 (Б)	0,2	20
5 (Д)	0,5	50
6 (Г)	0,3	30

2.3 Графическое решение

Содержание описанного алгоритма удобно изобразить в графической форме, показанной на рисунке 2.

Построим квадрат $ABCD$ со стороной, равной единице. Сторону BC назовем линией стоимости, а сторону DC – линией дефекта.

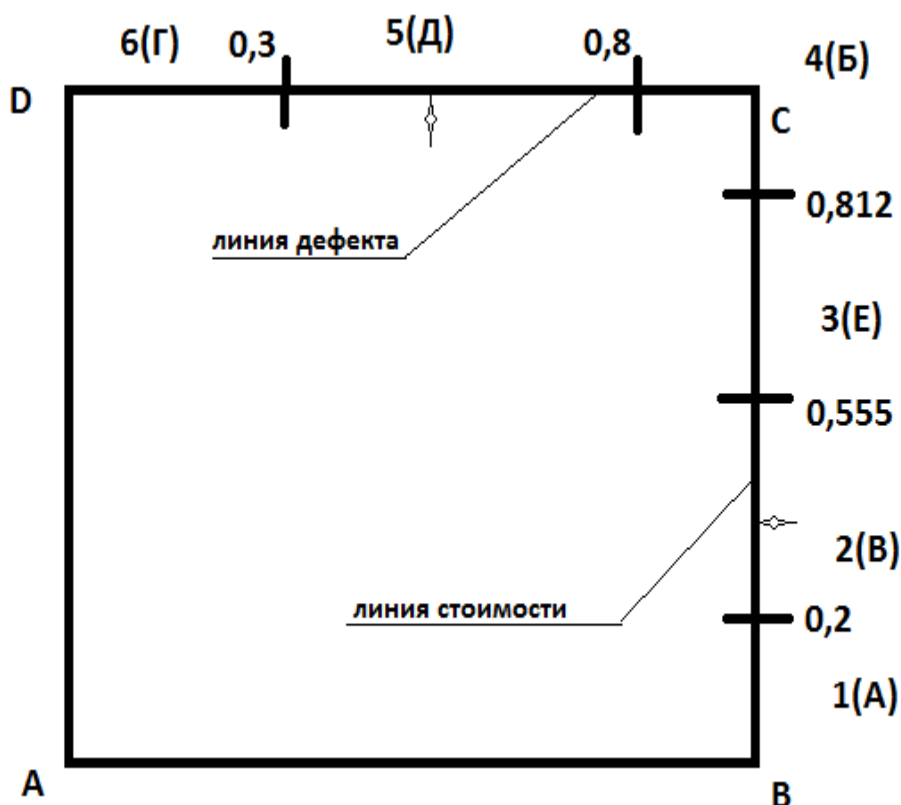


Рисунок 2 – Графический метод расчёта шансов оптимальности для вариантов решений

На линии стоимости, начиная от точки *B*, отложим значения индекса стоимости, выделенные подчеркиванием в столбце 4 таблицы 2. На линии дефекта, начиная от точки *D*, отложим значения индекса дефекта, выделенные подчеркиванием в столбце 5 таблицы 2. Пронумеруем получившиеся интервалы последовательно цифрами от 1 до 6 (очевидно, это номера сравниваемых решений), передвигаясь от точки *B* к точке *D* через точку *C*.

Если ЛПР не считает один из критериев более важным, чем другой, то длины пронумерованных интервалов, умноженные на 50, равны шансам соответствующих вариантов решений быть оптимальными.

Если же ЛПР считает, например, что критерий «стоимость» важнее критерия «эффективность», то линия менее важного критерия исключается из рассмотрения. Варианты решений, интервалы которых попадают на оставшуюся в рассмотрении линию, имеют шансы быть оптимальными. Эти шансы равны длинам интервалов, умноженным на 100.

В нашем примере, если стоимость «важнее» эффективности, то шансы быть оптимальным для варианта решения 1 (*A*) равны 20, для варианта 2 (*B*) равны 35,5, для варианта 3 (*E*) равны 25,7 (т.е. вдвое больше, чем в случае равноважных критериев); шансы оптимальности варианта решения 4 (*B*) равны $100 \cdot (1 - 0,812) = 18,8$. Варианты 5 (*D*) и 6 (*Г*) вообще не могут быть признаны оптимальными.

Если же более «важным» признан критерий эффективности, то шансы оптимальности варианта решения 4 (*B*) равны $100 \cdot (1 - 0,8) = 20$, варианта 5 (*D*) – 50, варианта 6 (*Г*) – 30; нулевые шансы оптимальности у вариантов решений 1 (*A*), 2 (*B*), 3 (*E*).

2.4 Выбор наилучшего решения, когда ЛПР может уверенно указать, что какой-то вариант решения для него лучше какого-то другого

Возможно, в некоторых случаях ЛПР может уменьшить неопределённость, имеющуюся в задаче принятия решений, указав два *не обязательно реальных* варианта решений, из которых, по его уверенному суждению, один заведомо лучше другого. Обозначим нормированные стоимость и дефект лучшего решения (*good*) через c_g и d_g , а нормированные стоимость и дефект худшего решения (*bad*) через c_b и d_b . Чтобы эта информация не была тривиальной или противоречивой соотношение значений показателей должно быть разнонаправленным, т.е., если $c_g \geq c_b$, то $d_g \leq d_b$ (случай **A**), а если $c_g \leq c_b$, то $d_g \geq d_b$ (случай **B**).

В случае **A** отложим на линии стоимости от вершины *B* отметку c_b / d_g , а на линии дефекта – от вершины *D* отметку d_b / c_g . Допустимыми следует считать точки этих линий, расположенные соответственно между этими отметками и углами квадрата соответственно *B* и *D*.

В случае **B** отложим на линии стоимости от вершины *B* отметку c_g / d_b , а на линии дефекта – от вершины *D* отметку d_g / c_b . Допустимыми следует считать точки этих линий, расположенные соответственно между этими отметками и углами квадрата соответственно *B* и *D*. Если величина отметки равна или превышает единицу, допустимыми остаются все стороны квадрата *ABCD* (рисунок 2).

Например, пусть ЛПР уверенно считает, что вариант решения (0,3; 0,1) лучше, чем вариант решения (0,2; 0,9). Это соответствует случаю **A**, и на линии стоимости следует отложить отметку $0,3 / 0,9 = 0,333$, а на линии дефекта – отметку $0,1 / 0,2 = 0,5$. Длина интервалов, содержащих допустимые точки соответствующих им вариантов решения, составляет теперь:

- для 1 (A) – 0,2;
- для 2 (B) – $0,333 - 0,2 = 0,133$;
- для 6 (Г) – 0,3;
- для 5 (Д) – $0,5 - 0,3 = 0,2$.

Разделив каждую длину на сумму этих длин и умножив на 100, получим шансы оптимальности соответствующих вариантов решений:

- для 1 (A) – 24,
- для 2 (B) – 16,
- для 6 (Г) – 36,
- для 5 (Д) – 24.

Шансы оптимальности других вариантов решений равны нулю.

В итоге приводим сводную таблицу 5, отражающую изменение шансов оптимальности различных вариантов решений в зависимости от различных суждений ЛПР.

Таблица 5 – Сводная таблица расчётов шансов оптимальности при различных уверенных суждениях ЛПР

Вариант решения	Стоимость (у.е.)	Эффективность (у.е.)	Эквивалентная значимость критериев	Стоимость «важнее» эффективности	Эффективность «важнее» стоимости	(0,3; 0,1) лучше, чем (0,2; 0,9)
1	2	3	4	5	6	7
A (1)	100	16	10	20	0	24
Б (4)	165	20	19	19	20	
В (2)	120	17	18	36	0	16
Г (6)	200	26	15	0	30	36
Д (5)	175	23	25	0	50	24
Е (3)	150	18	13	26	0	

3 Обоснование метода

3.1 Оценочная функция

Не уменьшая общности, будем считать, что среди вариантов решений нет неэффективных по Парето, и что номера вариантов решений следуют в порядке возрастания их стоимости. Нормируем значения критериев от минимального, принятого за нуль, до максимального, принятого за единицу, и перейдем от нормированной эффективности к нормированному дефекту d , дополняя нормированную эффективность до единицы.

Тогда

$$(1) \quad c_i \geq c_{i-1}, i = 2, \dots, m, c_1 = 0, c_m = 1; d_i \leq d_{i-1}, i = 2, \dots, m, d_1 = 1, d_m = 0.$$

Примем следующую *гипотезу о принципиальном существовании оценочной функции*.

Положим, что теоретически может существовать неизвестная оценочная функция, аргументами которой являются нормированные частные критерии. Она позволяет по значениям стоимости и эффективности любого варианта решения найти некоторую его скалярную оценку. Такая функция позволила бы легко решить проблему принятия многокритериального решения – лучшим был бы тот вариант решения, который имел бы наибольшее (наименьшее) значение такой оценки. Большинство существующих методов принятия многокритериальных решений направлено именно на отыскание такой функции. При этом её структура задаётся универсальной, а задача ЛПР сводится к тому, чтобы определить значение настроечных параметров этой функции, наиболее адекватных специфике конкретной решаемой задачи.

Часто в качестве оценочной функции используют линейную свёртку критериев. При этом настроечными параметрами являются коэффициенты, с которыми критерии входят в эту свёртку. Если используются нормированные значения критериев, то настроечные «весовые коэффициенты» отражают как бы сравнительную важность различных критериев с позиций ЛПР. Такая простая интерпретация привлекает пользователей, хотя показано (например, [7]), что подобная структура может «не замечать» даже некоторые Парето-оптимальные решения.

Предъявим к оценочной функции следующие требования:

- 1) она должна иметь универсальную структуру, т.е. описываться универсальными математическими средствами, а для отражения специфики решаемой задачи включать лишь один свободный числовой параметр;
- 2) она должна быть непрерывна и определена на квадрате $[0, 1] \times [0, 1]$;
- 3) функция должна быть монотонна по своим аргументам, т.е. большим значениям пар аргументов должно отвечать большее (точнее, не меньшее) значение функции;
- 4) любая Парето-оптимальная точка любой совокупности Парето-оптимальных точек на этом квадрате может рассматриваться как точка минимума этой функции при некотором значении параметра.

Покажем, что такая оценочная функция может иметь один-единственный вид (свёртка Гермейера [3]):

$$(2) \quad F = \max(x_c \cdot c, x_d \cdot d), \text{ где } \max(x_c, x_d) = 1, x_c \geq 0, x_d \geq 0.$$

В соотношении (2) присутствуют два параметра x_c , и x_d , однако они функционально связаны и потому эквиваленты одному свободному параметру.

Действительно, рассмотрим, какой вид должны иметь семейства линий уровня оценочной функции. При любом значении входящего в функцию параметра её линии уровня должны заполнять весь квадрат $[0, 1] \times [0, 1]$, не пересекаться, являться монотонно убывающими,

а их пометки (значение функции на линии уровня) должны возрастать от начала координат к периферии квадрата (рисунок 3).

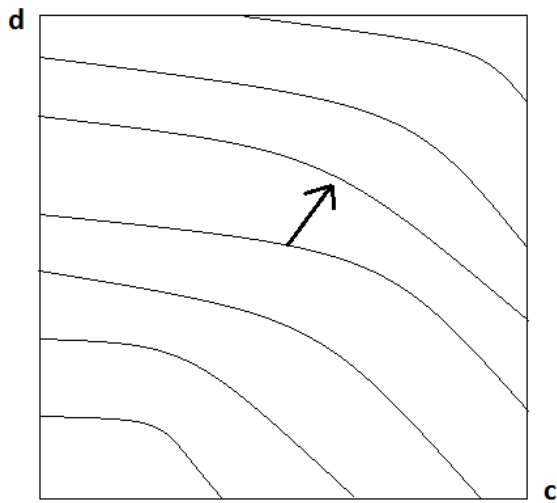


Рисунок 3 – Семейство линий уровня оценочной функции при некотором значении параметра

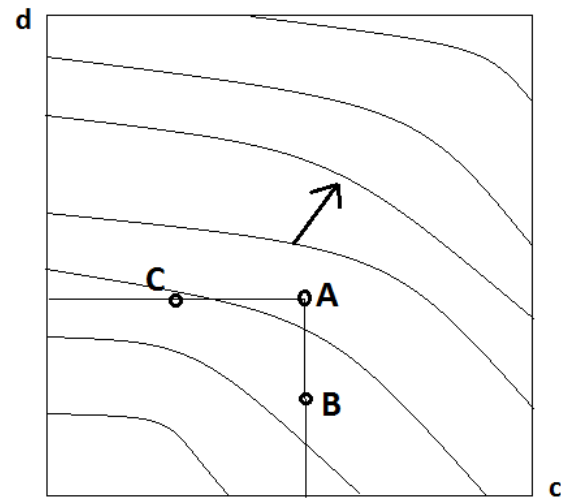


Рисунок 4 – Система трех Парето-оптимальных точек

Рассмотрим на квадрате $[0, 1] \times [0, 1]$ множество, состоящее из трёх Парето-оптимальных точек A, B и C, построенное следующим образом: A – любая точка квадрата, B и C – также две любые точки, из которых первая имеет значение абсциссы, на бесконечно малую величину большее, чем значение абсциссы точки A, а вторая – значение ординаты, на бесконечно малую величину большее, чем значение ординаты точки A. Значения же соответственно ординат и абсцисс точек B и C меньше соответствующих координат точки A на конечную величину. Рисунок 4 иллюстрирует ситуацию, когда наименьшее значение оценочная функция принимает в точке B, поэтому при значении параметра, отвечающем изображенному на этом рисунке семейству линий уровня, эта точка определяется как Парето-оптимальная во множестве точек A, B и C. Легко представить себе, причём различным образом, семейство линий уровня, при котором линия с меньшей пометкой проходит через точку C, определяя её в качестве Парето-оптимальной. А вот линия уровня, определяющая в качестве Парето-оптимальной точку A, и потому имеющая пометку не больше пометок линий уровня, проходящих через точки B и C, может выглядеть лишь единственным образом, как показано на рисунке 5 (линия Г1). Действительно, при ином виде этой линии уровня, например Г2, через точки C (например, линия Г3) и B будут проходить линии уровня с меньшими пометками, чем пометка линии Г1.

Поскольку все три точки A, B и C являются текущими, отсюда следует, что семейство линий уровня оценочной функции при любом значении параметра должно иметь структуру, показанную на рисунке 6. А такой структурой обладает свёртка Гермейера (2). Параметром семейства линий уровня является коэффициент наклона k направляющей прямой Γ . На этой прямой

$$x_c \cdot c, = x_d \cdot d,$$

или

$$(3) \quad d / c = x_c / x_d.$$

Отметим, что аналогичные соображения относительно оценочной функции имеют место и в общей задаче принятия решений при любом конечном числе критериев $f^j, j = 1, \dots, p$:

$$(4) \quad F = \max_{j=1, \dots, p} (x^j \cdot f^j), \text{ где } \max_{j=1, \dots, p} (x^j) = 1, x^j \geq 0, j = 1, \dots, p.$$

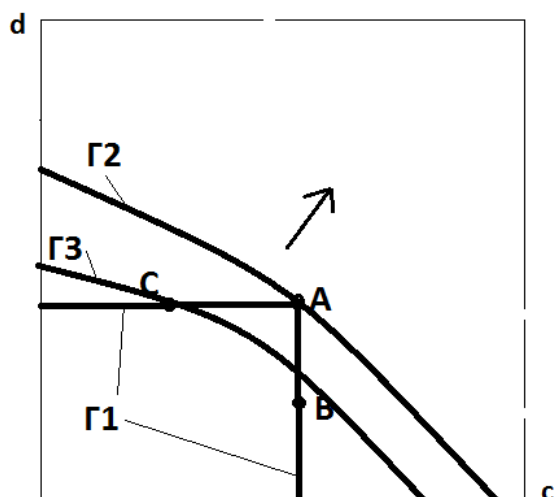


Рисунок 5 – Иллюстрация особенности линии уровня Γ_1

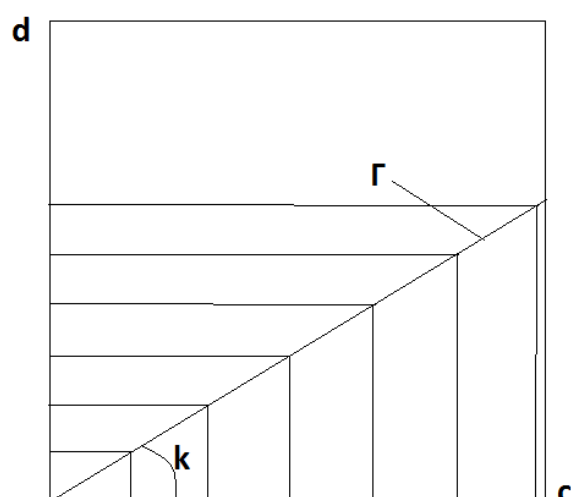


Рисунок 6 – Семейство линий уровня оценочной функции при параметре k

3.2 Основные соотношения

Принятая гипотеза о принципиальном существовании оценочной функции делает задачу принятия многокритериальных решений математически замкнутой. Расплывчатое понимание необходимости сопоставления в ней «сравнительного влияния» различных разнокачественных критериев (в рамках настоящей статьи это стоимость и эффективность) на выбор «наилучшего решения» отображается введением в рассмотрение неопределённых коэффициентов x_c , x_d , удовлетворяющих наложенным в (2) ограничениям. Эти коэффициенты определяют наклон направляющей прямой Γ (рисунок 6).

Если взглянуть на поставленную таким образом задачу принятия решения, то традиционный подход заключается в том, чтобы выявить конкретное значение этих коэффициентов, которое адекватно конкретно решаемой задаче. Мы считаем такой подход неверным теоретически и вредным практически, поскольку:

- во-первых, объективно невозможно формализовать специфику конкретной задачи в виде единственного числа (в случае двухкритериальной задачи) или $p - 1$ чисел в случае использования p частных критериев;
- во-вторых, стремление выявить это несуществующее число с помощью экспертных процедур или хитрых опросных процедур неоправданно усложняет и затягивает процесс принятия решения, а, кроме того, может подменять действительное понимание конкретики задачи ЛПРом, уполномоченным на это своим статусом, мнением определённых лоббистских групп.

Возможность внесения в решаемую задачу дополнительной информации, безусловно, должна быть сохранена, однако она должна сводиться к внесению *органичных* для ЛПР расплывчатых суждений типа «лучше–хуже», «более важно–менее важно», причем она *не должна* быть направлена на выявление, в конечном счете, конкретного значения параметра (набора параметров при более чем двух критериях) оценочной функции.

В [1] нами был обоснован учитывающий сказанное *Метод уверенных суждений* ЛПР (метод МУС), дающий простой, универсальный, максимально объективный, наиболее экономный и понятный для ЛПР способ решения задачи наилучшего выбора. Он состоит в том, чтобы в качестве комплексной оценки варианта решения рассматривать шансы его опти-

мальности – меру множества допустимых значений параметра оценочной функции, при которых эта функция принимает на нём оптимальное значение по сравнению со значениями на всех других допустимых варианта решения. В [2] разработан метод определения шансов оптимальности вариантов решений в общей задаче многокритериальной оптимизации. В задаче, рассматриваемой в данной статье, учитывая наличие в ней всего двух критериев, этот метод принимает наиболее простой и наглядный вид. На рисунке 7 показаны точки, отображающие t вариантов решений, пронумерованных от 1 до t . Проведенные на рисунке прямые ограничивают наклон направляющих линий оценочной функции, между которыми решения оказываются оптимальными.

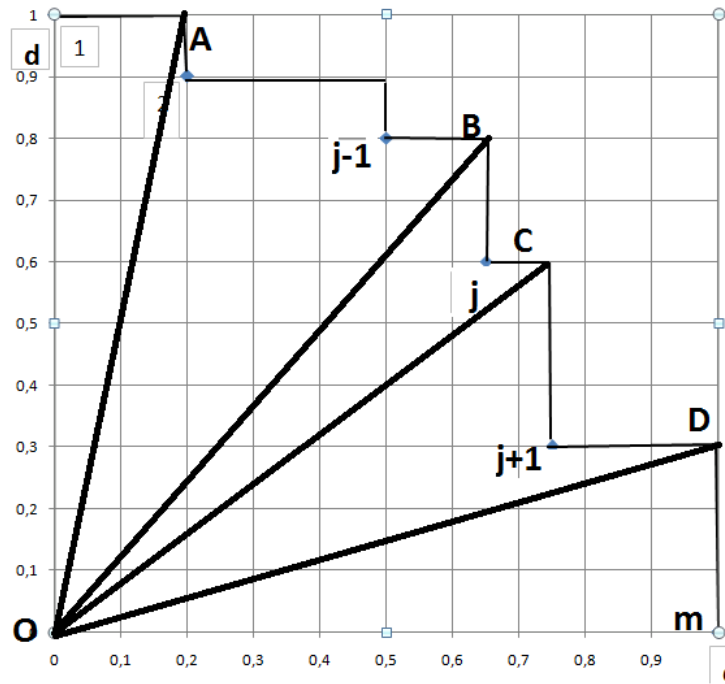


Рисунок 7 – Направляющие линии оценочной функции при расчёте шансов оптимальности вариантов решений

Для «внутреннего» варианта решения с номером j , $j = 1, \dots, t$ наклон направляющих прямых ограничен линиями OC , OB .

Это выразится неравенством

$$(5) \quad d_i / c_{i+1} \leq d / c = x_c / x_d \leq d_{i-1} / c_i$$

С учётом ограничений, наложенных в (2) на коэффициенты оценочной функции, неравенство (5) распадется на два неравенства:

- если $x_c = 1$, то $d_i / c_{i+1} \leq 1 / x_d \leq d_{i-1} / c_i$, откуда $c_i / d_{i-1} \leq x_d \leq c_{i+1} / d_i$;
- если $x_d = 1$, то $d_i / c_{i+1} \leq x_c \leq d_{i-1} / c_i$.

Для оптимальности варианта решения с номером 1 направляющая прямая Γ должна находиться между прямой OA и осью ординат на рисунке 7. Аналогично рассуждениям, проведенным для «внутреннего» варианта решения, легко показать, что этому отвечают условия

$$x_c = 1, 0 \leq x_d \leq c_2.$$

Заметим, что в данном случае неравенство (5) в предположении $x_d = 1$ приводит к условию $x_c \geq 1 / c_2$, что возможно лишь в вырожденном случае $c_2 = 1$.

Для оптимальности варианта решения с номером t направляющая прямая Γ должна находиться между прямой OD и осью абсцисс на рисунке 7. Аналогично предыдущим рассуждениям, легко показать, что этому отвечают условия

$$x_d = 1, 0 \leq x_c \leq d_{m-1}.$$

Соотношения, полученные в настоящем пункте, обосновывают все результаты, приведенные в разделе 2 статьи.

4 Выбор наилучшего из трёх решений

Особенно простые результаты в виде конечных формул получаются в задаче, в которой сравниваются всего *три* варианта решения. Как легко видеть из подраздела 3.2, в этом случае шансы оптимальности решений таковы:

Для решения, отображенного точкой 1, $F_1 = c_2 / 2$, для решения, отвечающего точке 2, $F_2 = 1 - (c_2 + d_2) / 2$, а для решения, отвечающего точке 3, $F_3 = d_2 / 2$.

На рисунке 8 показаны зоны оптимальности каждого из трех вариантов решений. Его нужно понимать следующим образом. Если точка 2 расположена в некоторой точке изображенного на рисунке квадрата, то зона, в которую эта точка попала, показывает, какое из трёх решений будет иметь наибольшее значение шанса оптимальности. Рисунок полезен тем, что показывает, насколько должно отличаться от аналогов (точек 1 и 3) решение (точка 2), для того, чтобы оно было конкурентоспособно по сравнению с ними.

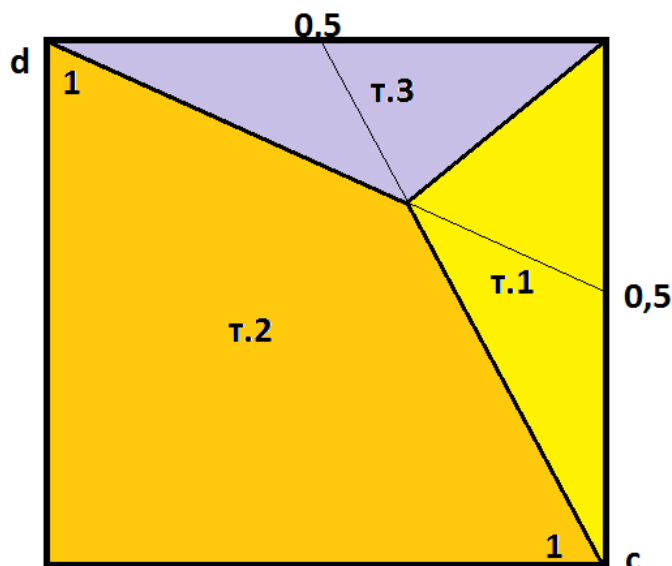


Рисунок 8 – Зоны оптимальности трёх вариантов решений

Заключение

Задаче выбора многокритериальных альтернатив присуща органическая неопределённость. Плодотворный путь в решении таких задач состоит не в том, чтобы пытаться получить от ЛПР в принципе неформализуемую дополнительную информацию о конкретной задаче, якобы устраняющую эту неопределённость, а в том, чтобы признать её органический характер и перейти к расчёту шансов каждого варианта решения оказаться оптимальным с учётом всего множества неопределённостей. Этот путь, реализация которого была начата нами сорок лет назад [4] и продолжена разработкой метода ПРИНН [5-7], продемонстрировал свои широкие возможности, предложив простое, универсальное и максимально объективное решение задачи выбора решений в критериальном пространстве «стоимость–эффективность».

Список источников

- [1] **Пиявский, С.А.** Два новых понятия верхнего уровня в онтологии многокритериальной оптимизации / С.А. Пиявский // Онтология проектирования. – 2013. – №1(7). – С. 65-85.
- [2] **Пиявский, С.А.** Прогрессивность многокритериальных альтернатив / С.А. Пиявский // Онтология проектирования. – 2013. – №4(10). – С. 60-71.
- [3] **Гермейер, Ю.Б.** Введение в теорию исследования операций / Ю.Б. Гермейер. – М.: Наука, 1971. – 383 с.
- [4] **Пиявский, С.А.** Оптимизация параметров многоцелевых летательных аппаратов / С.А. Пиявский, В.С. Брусов, Е.А. Хвилон. – М.: «Машиностроение», 1974. – 106 с.
- [5] **Смирнов, О.Л.** САПР: формирование и функционирование проектных модулей. / О.Л. Смирнов, С.А. Падалко, С.А. Пиявский. – М.: Машиностроение, 1987. – 272 с.
- [6] **Пиявский, С.А.** Методы оптимизации и принятия решений / С.А. Пиявский. – Самара: СГАСУ, 2004.
- [7] **Малышев, В.В.** Метод принятия решений в условиях многообразия способов учёта неопределённости / В.В. Малышев, Б.С. Пиявский, С.А. Пиявский // Известия РАН. Теория и системы управления. – 2010. – №1. – С. 46–61.

A SIMPLE AND UNIVERSAL METHOD OF DECISION MAKING WITHIN THE SCOPE OF CRITERIA OF «COST AND EFFICIENCY»

S.A. Piyavsky

*Samara State University of Architecture and Civil Engineering, Samara, Russia
spiyav@mail.ru*

Abstract

A new concept of chances of decision optimality is introduced to the ontology of multi-criteria decision making. Its fruitfulness is illustrated by a practically significant task of comparing the alternatives within the scope of criteria of “cost and efficiency” in which this concept makes it possible to get complex quantitative evaluation of decision quality without any compulsory input of additional information from a decision-maker., except some indeterminate judgments as “better” or “more important”, if desired. In addition to it, the author shows that the only incessant evaluation function that makes it possible to determine all Pareto-optimal decisions in any range of multi-criteria alternatives is the minimax folding of Yu.B. Germeyer. The chance of decision optimality is determined by the measure of the variety of values of indefinite coefficients in this folding when the evaluation function gets the best value within one solution than within any other solutions of the given range of alternatives. A simple calculation algorithm and a visual geometric method of evaluating optimality chances have been worked out.

Key words: multi-criteria optimization, decision making, decision-maker, chances of optimality, alternative, indeterminacy, Pareto-optimality, cost, efficiency.

References

- [1] **Piyavsky, S.A.** Dva novykh ponyatia verkhnego urovnya v ontologii mnogokriterialnoi optimizatsii [Two New Upper Level Concepts for the Multi-criteria Optimization Ontology] / S.A. Piyavsky // Ontology of Designing. - 2013. – №1(7). – P. 65-85. (In Russian)
- [2] **Piyavsky, S.A.** Progressivnost mnogokriterialnykh alternativ [Progressivity of Multi-criteria Alternatives] / S.A. Piyavsky // Ontology of Designing. - 2013. – №4(10). – P. 60-71. (In Russian)
- [3] **Germeyer, Yu.B.** Vvedenie v teory issledovaniya operatsii [Introduction to Operation Research] / Yu.B. Germeyer. – М.: Nauka, 1971. – 383 p. (In Russian)
- [4] **Piyavsky, S.A.** Optimizatsia parametrov mnogotselevykh letatelnykh apparatov [Optimization of Parameters of Multipurpose Aircrafts] / S.A Piyavsky, V.S. Brusov, E.A. Khvilon. – М.: Mashinostroenie, 1974. – 106 p. (In Russian)
- [5] **Smirnov, O.L.** SAPR: formirovanie I funktsionirovanie proektnykh modulei [CAD: Formation and Functioning of Design Modules] / O.L. Smirnov, S.A. Padalko, S.A. Piyavsky – М.: Mashinostroenie, 1987. – 272 p. (In Russian)

- [6] **Piyavsky, S.A.** Metody optimizatsii i prinyatia reshenii [Methods of Optimization and Decision Making] / S.A. Piyavsky. – Samara: Samara State University of Architecture and Civil Engineering, 2004. (In Russian)
- [7] **Malyshev, V.V.** Metod prinyatia reshenii v usloviakh mnogoobrazia sposobov ucheta neopredelennosti [The Method of Decision Making in conditions of the Diversity of Indeterminacy Account Methods] / V.V. Malyshev, B.S Piyavsky, S.A. Piyavsky // Izvestia RAN. Teoria I sistemy upravleniya [Journal of Computer and Systems Sciences International]. – 2010. - №1. – P. 46-61. (In Russian)
-

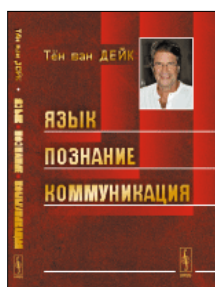
Сведения об авторе



Пиявский Семен Авраамович. Окончил факультет летательных аппаратов Куйбышевского авиационного института в 1964 году, аспирантуру при кафедре динамики полета Московского авиационного института им. С.Орджоникидзе в 1967 году. Доктор технических наук, профессор, заведующий кафедрой прикладной математики и вычислительной техники Самарского государственного архитектурно-строительного университета. Почетный работник высшей школы РФ, академик Академии наук о Земле и Академии нелинейных наук. Опубликовал более 350 научных работ в области системного анализа, методов оптимизации и принятия решений, математического моделирования, образовательных систем и технологий. Основные научные результаты: онтологии образовательного процесса, методы многоэкстремальной оптимизации, решения краевых задач для систем обыкновенных дифференциальных уравнений, принятия решений в условиях неустраняемой неопределённости, оптимизации сетей ИСЗ, оптимизации многоцелевых систем летательных аппаратов; теория многоцелевых систем, компьютерная технология технического творчества, теория оптимального управления развитием научных способностей молодежи, новые формы организации образовательного процесса в высшей школе в условиях развитой инфокоммуникационной среды, матричная структура студенческого коллектива и др.

Semyon Avraamovich Piyavsky. Graduated from the Aircraft faculty of Kuibyshev Aviation Institute in 1964 and the graduate school at the Flight Dynamics Department at the Moscow Aviation Institute named after Ordzhonikidze in 1967. Doctor of Technical Sciences, Professor, Head of the Department of Applied Mathematics and Computer Science at Samara State University of Architecture and Civil Engineering. Honored Worker of Higher School of Russia, Academician of the Academy of Earth Sciences and Academy of Nonlinear Sciences. He has published over 350 scientific papers on system analysis, optimization techniques and decision-making, mathematical modeling, education systems and technologies. Basic scientific results: education ontologies, Multiple-optimization techniques, solution of boundary value problems for systems of ordinary differential equations, decision making under fatal uncertainty, computer technology of engineering creation, the optimal control theory of young people' academic abilities development, new forms of organization of educational process in higher education in advanced info-communications environment, the matrix structure of the student groups, etc.

Рекомендуемые издания 2014 года по тематике журнала



Тён ван Дейк. Язык. Познание. Коммуникация.

Пер. с англ. Изд.2. - М.: ЛИБРОКОМ, 2015. - 320 с.

В книгу включены труды известного европейского учёного Т.А. ван Дейка - одного из основателей лингвистики текста. Данные работы посвящены лингвистической прагматике, пониманию и порождению текста, функционированию языка в системах массовой информации. Автор детально рассматривает воздействие социокультурных факторов на механизм употребления языка. Исследование коммуникации проводится с точки зрения социальной психологии на основе большого эмпирического материала. Книга представляет интерес для лингвистов, психологов, социологов, этнологов и других специалистов, занимающихся проблемами функционирования языка в сфере массовой коммуникации, а также для широкого круга читателей.



Васильев В.В. Сознание и вещи:

очерк феноменалистической онтологии.

- М.: Книжный дом «ЛИБРОКОМ», 2014. - 340 с.

Сознание остается одной из главных загадок для философии и экспериментальной науки. Признавая успехи экспериментальных исследований сознания, автор проводит свои изыскания в концептуальном ключе, пытаясь прояснить структуру и соотношение базовых убеждений о мире и о самих себе. Все мы верим в то, что прошлый опыт можно использовать для прогнозов на будущее, что в мире не бывает беспричинных событий и что физические объекты независимы от нашего сознания. Установив соотношение этих убеждений, мы сможем уточнить онтологический статус сознания и понять отношение между ментальным и физическим.



Гартман Э. Сущность мирового процесса, или философия бессознательного. Метафизика бессознательного.

Пер. с нем. Изд. стереотип. - М.: КРАСАНД, 2014. - 440 с.

Вниманию читателей предлагается классический труд выдающегося немецкого философа Э. Гартмана (1842-1906), принесший ему мировое признание. Книга посвящена проблеме бессознательного, которое, по мнению автора, выступает как неотъемлемый элемент человеческой психики, источник жизни и ее движущая сила. Автор отталкивается от эмпирических фактов и выводит из них свою основную гипотезу, от которой через ряд выводов и соображений приходит к метафизической сущности и ко всем главным положениям своего мировоззрения. Книга предназначена философам, психологам, историкам науки, всем заинтересованным читателям.



Хакен Г. Информация и самоорганизация: макроскопический подход к сложным системам.

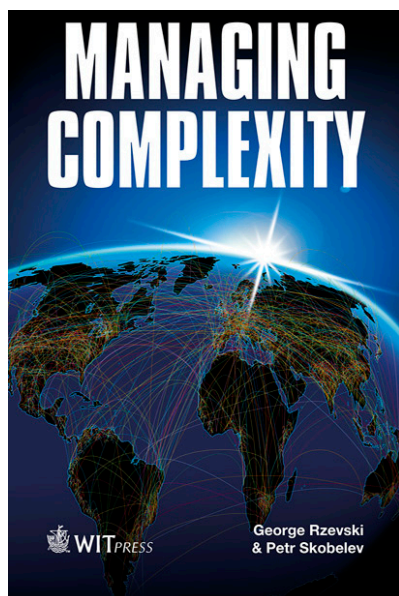
Пер. с англ. Изд.3, испр. и знач. доп. - М.: УРСС: ЛЕНАРД, 2014. - 320 с.

В книге немецкого физика-теоретика, одного из основоположников синергетики Германа Хакена, развит оригинальный подход к описанию сложных макроскопических систем. Исследована взаимосвязь информации и самоорганизации на основе принципа максимума информационной энтропии в применении к широкому кругу неравновесных процессов. На качественно новом уровне рассмотрен синергетический подход к проблеме распознавания образов, а также сформулированы принципы синергетического компьютера. В новом издании рассматривается взаимосвязь между шенноновской информацией и семантической информацией. Добавлены новые главы, посвященные фундаментальным вопросам, связанным с квантовой информацией и квантовыми вычислениями.



Новиков А.М., Новиков Д.А. Методология: словарь системы основных понятий. Изд. стереотип. - М.: УРСС, 2015. - 208 с.

Настоящий словарь, написанный академиком РАО А.М.Новиковым и членом-корреспондентом РАН Д.А.Новиковым, содержит около 300 статей, раскрывающих содержание основных понятий методологии. Системность совокупности приводимых понятий обеспечивается покрытием предметной области, полнотой и непротиворечивостью связей между ними. Работа предназначена для научных и практических работников, а также студентов, аспирантов и докторантов; в первую очередь - для преподавателей вузов и институтов повышения квалификации для использования при подготовке курсов лекций по теории систем, системному анализу, методологии научных исследований, инновационной деятельности, проектированию систем, управлению проектами и т.д.



Managing Complexity

G. Rzevski, The Open University, UK and

P. Skobelev, Software Engineering Company «Smart Solutions» Ltd., Russia

Book Description

Managing Complexity is the first book that clearly defines the concept of Complexity, explains how Complexity can be measured and tuned, and describes the seven key features of Complex Systems:

1. Connectivity, 2. Autonomy, 3. Emergency, 4. Nonequilibrium,
5. Non-linearity, 6. Self-organisation, 7. Co-evolution

The thesis of the book is that complexity of the environment in which we work and live offers new opportunities and that the best strategy for surviving and prospering under conditions of complexity is to develop adaptability to perpetually changing conditions. An effective method for designing adaptability into business processes using multi-agent technology is presented and illustrated by several extensive examples, including adaptive, real-time scheduling of taxis, sea-going tankers,

road transport, supply chains, railway trains, production processes and swarms of small space satellites. Additional case studies include adaptive servicing of the International Space Station; adaptive processing of design changes of large structures such as wings of the largest airliner in the world; dynamic data mining, knowledge discovery and distributed semantic processing.

Finally, the book provides a foretaste of the next generation of complex issues, notably, The Internet of Things, Smart Cities, Digital Enterprises and Smart Logistics

<http://amzn.com/1845649362>

Contents

PART 1 Fundamentals

1 What is complexity?

- Defining Complexity
- Complex Versus Complicated
- Complexity and Uncertainty
- The Seven Criteria of Complexity
- Negative and Positive Aspects of Complexity
- Evolution Favours Complexity
- Co-Evolution of Technology, Economy and Society
- Complexity and Information Society
- Complexity and Philosophy

2 A method for managing complexity

- Coping with External Complexity
- Tuning Internal Complexity
- Modelling Complexity
- Adaptability
- Designing Adaptive Business Processes

3 Multi-agent technology

- Fundamentals
- MAS for Adaptive Resource Allocation
- Knowledge Base
- Virtual World
- Decision-Making
- Agent Negotiations
- Architecture
- Multi-Agent Platform
- Main Features of Our MAS
- Multi-Agent Software as a Complex Adaptive System
- Comparing Multi-Agent Software with Conventional Programs

4 Emergent intelligence

- Fundamentals
- Evidence of Intelligent Behaviour
- Thermodynamics of the Virtual World

PART 2 Commercial applications

5 Adaptive scheduling of seagoing tankers

- The Problem
- The Solution
- Results

6 Adaptive scheduling of taxis

7 Adaptive scheduling of car rentals

8 Adaptive scheduling of road transport

9 Adaptive data mining

10 Adaptive semantic processing

11 Adaptive detection of clashes caused by design changes

12 Adaptive scheduling of supply networks

13 Adaptive scheduling of services for the international space station

14 Adaptive scheduling of a fleet of satellites

15 Adaptive scheduling of high-speed railways

16 Adaptive scheduling of manufacturing

17 Adaptive management of service teams

18 Adaptive project management

PART 3 A roadmap into the future

19 A vision and ideas

- A Shift from Personal to Business Applications
- The I o T
- Digital Enterprise
- Smart City
- Smart Logistics
- The story of managing complexity

References

Индекс 29151

Подписка на журнал 2015

Продолжается подписка на номера журнала

«Онтология проектирования» 2014 и 2015 гг.

Стоимость подписки на год через редакцию журнала 2000 руб.

Подробности на сайте журнала.

Для подписчиков журнала, оформивших подписку на 2014 и 2015 годы, **бонус** от редакции журнала в виде книги, рекомендованной к изданию редакционно-издательским советом федерального государственного автономного образовательного учреждения высшего профессионального образования «Самарский государственный аэрокосмический университет имени академика С.П.Королева (национальный исследовательский университет) (СГАУ)» в качестве учебного пособия для студентов, обучающихся по программам высшего профессионального образования по направлению подготовки магистров (Авиастроение).



Книга может быть полезна аспирантам, научным сотрудникам, инженерам и всем тем, кого интересует суть процесса проектирования.

Пособие по одноименной дисциплине «Онтология проектирования» подготовлено на кафедре конструкции и проектирования летательных аппаратов СГАУ. В нём излагаются теоретические основы новой учебно-научной дисциплины: истоки онтологии, терминологические соглашения, принципы проектирования, психология проектирования, онтология как спецификация концептуализации, языки описания и инструменты проектирования онтологий. На основе системного подхода, достижений информационных технологий и современной философии обосновывается целесообразность дальнейшего развития онтологического анализа начального этапа жизненного цикла сложных систем – этапа проектирования.

Второе издание пособия (1-е вышло в 2010 году) переработано и дополнено. В него вошли материалы, опубликованные автором в период 2010-2014 г.г. в научных журналах и трудах международных научных конференций.

Ontologists and designers of all countries and subject areas, join us!



Издательство «Новая техника» - Publisher «New Engineering» Ltd

Россия, 443010, Самара, ул.Фрунзе 145 - 145, Frunze Str., Samara, 443010, Russia