

ОТЗЫВ

официального оппонента

на диссертацию Забежайло Михаила Ивановича

«Комбинаторные средства формализации эмпирической индукции»,
представленную на соискание ученой степени доктора физико-математических наук по
специальности 05.13.17 – теоретические основы информатики

Диссертационная работа М.И. Забежайло **является результатом многолетних работ**, посвященных исследованиям и математическим моделям эвристических алгоритмов по выявлению закономерностей в данных. Это важная область имеет значительные приложения, которые стимулируют развитие математических подходов. В работе М.И. Забежайло в отличие от статистических подходов, которые широко используются для анализа числовых данных, рассматривается задача влияния структуры исследуемых объектов на свойства этих объектов. При этом число экспериментальных данных не очень велико, а обоснованность гипотез о выявленных закономерностях должна быть определена и оценена. **Актуальность темы диссертации** не вызывает сомнений. Подобные задачи возникают в науках о жизни при анализе влияния структуры генома на различные заболевания, в фармакологии для определения структур, определяющих лечебные свойства, в задачах анализа сложных устройств, в социологии. Это всё задачи, в решении которых современными методами математическая составляющая является значительной и определяющей. При этом, хотя имеются значительное число алгоритмических подходов, используемых при решении этих задач, ощущается потребность в общих теоретических подходах к ним. В диссертационной работе М.И. Забежайло используется подход предложенный профессором В.К. Финном логико-индуктивного анализа данных. В работах М.И. Забежайло сделано много для обоснования и развития этого подхода в течение многих лет сотрудничества с группой В.К. Финна.

Диссертация состоит из введения, восьми глав, заключения и списка литературы.

Первая глава носит вводный характер. В ней обсуждается общая проблема анализа данных и место в ней метода индуктивного подхода выявления закономерностей в зависимости от структуры исследуемых объектов, примеров и контрпримеров наличия групп свойств у объектов.

В главе 2 дается общее полуформальное описание ДСМ подхода к анализу данных, предложенного В.К.Финном. Формальное описание этого подхода вместе с обобщениями, предложенными автором диссертации дается в **главе 3**. Вводится и исследуется операция сходства на объектах, обладающих различными структурами и наборами свойств. Конструкции ДСМ-метода выявления и обоснования закономерностей на основе примеров и контрпримеров формулируются на основе общей алгебраической операции сходства исследуемых объектов. При этом рассматривается приложение общей предложенной схемы к объектам, представленным в виде множеств, графов и числовых векторов.

Глава 4 посвящена исследованию сложности рассмотренных в предыдущих главах алгоритмов. Здесь приводятся как результаты автора, так его коллег, занимающихся использованием ДСМ-метода, и исследующих структуры и алгоритмы, связанные с этим методом. Показано, что некоторые алгоритмы ДСМ-метода могут порождать элементы структур, число которых растет экспоненциально в зависимости от основных параметров модели – числа элементов в универсуме и общего числа свойств. Отсюда в диссертации делается вывод о необходимости модификации рассмотренных алгоритмов, построении направленного перебора, который позволит выдавать последовательно приближенные ответы по выявлению причинно-следственных связей.

В главе 5 автор предлагает и исследует приближенный ДСМ-метод. На основании операции сходства исследуемых объектов автором предлагается направленный перебор и построение подструктур для индуктивного построения и проверки гипотез причинно-

следственного отношения. Более того, алгоритм обеспечивает параллельную обработку по различным веткам подструктур. Исследуется корректность предложенного алгоритма.

Некоторые свойства элементов ДСМ-метода изучаются в **главе 6**.

В **главе 7** приводится демонстрационный пример, на котором показывается работа всех элементов ДСМ-метода и его обобщений, введенных автором. В **главе 8** рассматриваются примеры приложения полученных методов и алгоритмов к задачам фармакологии, классификации документов, прогноза нефтегазоносности региона, экспертизы сложных технических устройств.

Предлагаемый подход является новым, а форма представления материала работы, полученные выводы – достоверными. Научные положения, выводы и рекомендации, сформулированные в диссертации, являются обоснованными в достаточной степени. Предложенные автором решения аргументированы и оценены в сравнении с другими известными решениями.

В диссертационной работе М.И. Забейайло получены следующие основные результаты:

- Исследованы возможности организации обучения на прецедентах и автоматического извлечения зависимостей вида **ОБЪЕКТ=>СВОЙСТВА** из *данных нечислового характера*, базирующиеся на использовании комплекса эвристик *эмпирической индукции* (индуктивного обучения, выводов по аналогии и абдуктивного объяснения) и алгебраической *операции сходства объектов*.
- Процедурная конструкция ДСМ-метода автоматического порождения зависимостей расширена на *новые типы данных* (структурные описания *нечисловых объектов*, дополненные *числовыми значениями*).
- Разработана модификация ДСМ-метода (последовательных приближений) использующая направленный перебор и построение подструктур для индуктивного построения и проверки гипотез причинно-следственного отношения. Алгоритм

обеспечивает параллельную обработку по различным веткам подструктур. Исследована корректность предложенного алгоритма.

- Показана *практическая значимость и эффективность* предложенных математических моделей, методов и алгоритмов, подтвержденная результатами, полученными при решении прикладных задач в различных предметных областях.

Результаты диссертационной работы изложены **в семнадцати статьях, опубликованных в журналах, входящих в перечень ВАК РФ**, и докладывались на российских и международных конференциях.

Автореферат адекватно отражает содержание диссертационной работы.

Недостатки работы.

Основным недостатком работы является сложность текста. Читать работу непросто. Для пояснений автор некоторые формальные математические утверждения и определения пытается изложить на естественном языке. В тексте приводятся цитаты из философских работ для аргументирования некоторых эвристических алгоритмов. Это приводит к неточностям и к тому, что точные формальные определения приводятся позже, чем они используются.

Перечисленные недостатки не снижают ценности диссертационной работы Забежайло М.И. и не влияют на ее положительную оценку.

Диссертационная работа «Комбинаторные средства формализации эмпирической индукции» соответствует критериям, установленным Положением о присуждении ученых степеней, утвержденным Постановлением Правительства РФ от 24 сентября 2013 года №842: автором исследований разработаны теоретические положения, совокупность которых можно квалифицировать как научное достижение в области теории методов анализа данных. Диссертация обладает внутренним единством, содержит новые научные результаты и рекомендации по их практическому применению.

Работа полностью соответствует требованиям ВАК РФ, предъявляемым к диссертациям на соискание ученой степени доктора физико-математических наук по специальности 05.13.17. – теоретические основы информатики, а ее автор, **М.И.Забейайло**, заслуживает присуждения ему ученой степени доктора физико-математических наук по данной специальности.

Официальный оппонент:

доктор физико-математических наук, профессор, заведующий кафедрой математики, логики и интеллектуальных систем в гуманитарной сфере Федерального государственного бюджетного образовательного учреждения высшего профессионального образования «Российский государственный гуманитарный университет» (РГГУ)

Адрес: 125993, ГСП-3, Москва, ул. Чаянова, дом 15

Тел.: 8-495-250-63-29

E-mail: ebeniamin@yandex.ru

08 июня 2015



Бениаминов Евгений Михайлович

Подпись заверяю.
Ученый секретарь



Л.В. Тропкина