

Отзыв на автореферат

диссертации Забейайло Михаила Ивановича
«Комбинаторные средства формализации эмпирической индукции»,
представленной на соискание ученой степени доктора физико-математических наук
по специальности 05.13.17 – Теоретические основы информатики

Проблематика построения формализованных компьютерных моделей рассуждений была и остается одним из традиционных направлений исследований в области искусственного интеллекта. Дополнительную актуальность таким работам придает растущая потребность в разработке современных компьютерных инструментов интеллектуального анализа данных, которые были бы пригодны для использования в тех плохо формализованных областях исследования так называемых Big Data, где к настоящему времени еще не сложился свой универсальный аппарат анализа эмпирических зависимостей (например: в науках о живом, анализе социального поведения и др.). Особую роль здесь играют задачи формализации специального класса эвристик, ассоциируемых с так называемой проблемой эмпирической индукции – обучения на коллекциях прецедентов.

В диссертационной работе М.И.Забейайло задача формализации эмпирической индукции рассматривается в сравнительно редко изучаемой версии – при анализе примеров и контрпримеров (представленных нечисловыми описаниями - множествами признаков и некоторыми видами отношений на них, возможно, дополненными векторами числовых значений существенных параметров). При этом свойства прецедентов определяются теми или иными деталями описания их внутренней структуры (см., например, лекарственные свойства конкретного соединения, определяемые характерными комбинациями элементов его химической структуры и т.п.). Особенность развиваемого автором подхода – в уточнении сходства изучаемых прецедентов не в традиционных метрических терминах (например, как близости, выражаемой в терминах расстояния и т.п.), а как алгебраической операции, характеризующей структурную схожесть описаний анализируемых объектов. Операции с такими структурными описаниями (выделение возможных сходств) порождают комбинаторные объекты специальной природы (диаграммы вложимости классов эквивалентности прецедентов), анализ свойств которых и является основной целью развиваемого автором комплекса математических моделей методов и алгоритмов.

Автором предложена алгебраическая формализация комплекса эвристик (индукции в стиле Д.С.Милля, аналогии в стиле Д.Пойа и абдукции в стиле Ч.С.Пирса), используемого в рамках процедурной конструкции успешно развиваемого В.К. Финном и его школой так называемого ДСМ-метода автоматического извлечения зависимостей из структурных описаний исходных прецедентов. Обеспечена возможность обработки унифицированными средствами различных типов данных. Построены оценки вычислительной сложности порождения диаграмм вложимости классов эквивалентности прецедентов, характеризующих целый ряд возникающих здесь комбинаторных задач как трудно-разрешимые (т.е. принадлежащие известным классам трудных переборных проблем). Предложен комплекс моделей и алгоритмов, позволяющих организовать целенаправленное порождение таких диаграмм в режиме управляемых последовательных приближений (метод последовательных приближений – приближенный ДСМ-метод). Представлена возможность реализации разработанного приближенного ДСМ-метода в режиме параллельных вычислений. Работоспособность предлагаемых математических моделей методов и алгоритмов демонстрируется на примерах решения прикладных задач интеллектуального анализа данных в нескольких различных предметных областях (анализе

физиологических активностей химических соединений – канцерогенеза, токсичности и др., компьютерном анализе коллекций полнотекстовых документов, прогнозировании нефтегазоносности территорий и др.).

По теме диссертации автором подготовлены 37 публикаций, 17 из которых – в изданиях из перечня ВАК.

По автореферату следует сделать несколько замечаний:

1. Следовало бы явным образом указать, что в анализируемых автором (см. стр. 15-17) покрытиях классов сходства $T(a)$ классами эквивалентности $E_v(a)$ приходится иметь дело с формирующими соответствующие классы эквивалентности всего лишь *подмножествами* исходной выборки прецедентов. (Т.е. речь здесь не идет о *разбиении* исходного множества прецедентов на классы по *одному* отношению эквивалентности. Наоборот, рассматривается *система отношений эквивалентности*, порождаемых соответствующими подмножествами v , каждое из которых является общим для всех элементов каждого из таких поднаборов прецедентов. При этом каждое из таких отношений эквивалентности порождает всего один класс эквивалентности, в который собраны *все те и только те* прецеденты из исходной обучающей выборки, которые участвуют в порождении соответствующего v).
2. Описание процедур формализации сходства на векторах числовых значений параметров (см. стр. 20-22) носит достаточно общий характер и, по-видимому, нуждается в нескольких пояснительных примерах. Это в значительной мере упростило бы понимание формируемой автором формальной конструкции.

Однако данные замечания не носят существенного характера и не могут повлиять на общую положительную оценку автореферата и диссертационной работы в целом. Судя по автореферату, рассматриваемая диссертация «Комбинаторные средства формализации эмпирической индукции» полностью соответствует требованиям Положения о порядке присуждения ученых степеней ВАК Минобрнауки РФ, а ее автор - М.И.Забежайло – заслуживает присуждения ему ученой степени доктора физико-математических наук по специальности 05.13.17 – Теоретические основы информатики.

Генеральный директор Института точной
механики и вычислительной техники имени
С.А.Лебедева Российской академии наук,
д.ф-м.н., профессор

А.В.Князев

30 ноября 2016 года

ИТМ и ВТ: 119991, Москва, Ленинский проспект, д. 51
Контактный тел.: +7(495)649-12-70
Эл. почта: info@ipmce.ru

Подпись д.ф-м.н., проф. Князева А.В. заверяю

на начальнике отдела кадров
Александр Е.В. Шней