

ОТЗЫВ НАУЧНОГО КОНСУЛЬТАНТА

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

Диссертация Забейжайло Михаила Ивановича «Комбинаторные средства формализации эмпирической индукции» представляет собой законченное самостоятельное исследование в области интеллектуального анализа данных – направления искусственного интеллекта, ориентированного на создание математических моделей, методов и алгоритмов, а также программных инструментов, позволяющих извлекать новое знание из накапливаемых эмпирических данных. Решение задач такого рода актуально и чрезвычайно важно для развития теоретических основ информатики.

Диссертационная работа М.И.Забейжайло выполнена в Лаборатории №35 «Интеллектуальный анализ данных и автоматизированная поддержка научных исследований» Института Проблем Информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН).

Задача извлечения зависимостей из эмпирических данных (задача обучения на прецедентах) давно и хорошо известна исследовательскому сообществу, в частности, как задача так называемой эмпирической индукции. В диссертации рассмотрена пока еще мало изученная, однако, весьма важная для приложений версия этой задачи, в которой исходно заданные прецеденты представляют собою нечисловые структурные объекты (характеризуемые множествами признаков и отношениями на них – например, множествами признаков, графами, цепочками символов конечного алфавита и т.п.), дополняемые также числовыми значениями ряда параметров. Важные особенности исследовательской ситуации также связаны здесь с необходимостью оперировать в том числе и с малыми (статистически незначимыми) коллекциями исходно заданных прецедентов, учитывать причинно-следственные зависимости, которые связывают структурные описания прецедентов с множествами присущих им свойств.

Традиционное для искусственного интеллекта использование эвристик существенным образом расширяет возможности компьютерных реализаций процедурной конструкции эмпирической индукции. Однако, параллельно с этим в ряд особо актуальных выдвигается проблема формализации (выделения области корректности) применения таких эвристик.

Михаилом Ивановичем Забейжайло предложен, обоснован и развит оригинальный подход – комплекс математических моделей, методов и алгоритмов,

позволяющий построить процедурную конструкцию восстановления эмпирических зависимостей из коллекций прецедентов, в основе которого – учет структурного сходства описаний прецедентов. Сходство здесь формализуется как алгебраическая операция, с ее помощью определяется соответствующее отношение сходства, затем формируются классы сходства и, наконец, реконструируются системы классов эквивалентности наборов исходно заданных прецедентов, покрывающие сформированные классы сходства. Таким образом, каждой извлекаемой из исходной коллекции прецедентов эмпирической зависимости сопоставляется (причем - взаимно-однозначным образом) некоторый класс эквивалентности. Выделен и подробно изучен особый класс комбинаторных объектов – диаграмм взаимной вложенности классов эквивалентности на множестве исходно заданных прецедентов, структурные свойства которого и определяют все алгоритмические особенности рассматриваемого процесса извлечения эмпирических зависимостей из исходных данных.

Базовым элементом развиваемой процедурной формализации эмпирической индукции является вычисление сходства сложно структурированных нечисловых объектов (описаний прецедентов из исходной выборки). Значительная часть решаемых здесь переборных задач характеризуется экспоненциально быстро растущими оценками сложности вычислений. Таким образом, центральной и критически важной проблемой оказывается оптимизация комбинаторного перебора вариантов. Помимо теоретического значения эффективные решения в области подобной оптимизации также открывают возможности для более широкого применения компьютерных моделей эмпирической индукции при решении прикладных задач интеллектуального анализа данных.

Автором диссертационной работы создан комплекс математических моделей, методов и алгоритмов, позволяющих кардинальным образом изменить существовавшую до этого ситуацию. Предложенная М.И.Забежайло математическая техника дискретной оптимизации перебора позволяет организовать формирование соответствующих диаграмм вложенности в параллельном режиме вычислений, задействовав при этом оригинальную процедурную конструкцию целенаправленного последовательного порождения все более и более точных приближенных описаний целевой диаграммы, соответствующей исходно заданной для обучения коллекции прецедентов. Таким образом, сформированы алгоритмические основания для создания прикладных программных систем интеллектуального анализа данных, которые обеспечивали бы порождение индуктивных обобщений обсуждаемого вида на коллекциях эмпирических данных произвольного размера. Так, например, для приложений ДСМ-метода автоматического порождения гипотез предложенная М.И.Забежайло техника дискретной оптимизации позволяет снять имеющиеся на практике ограничения на размеры обучающих выборок исходных данных, которые можно обрабатывать в режиме процессно-реального времени.

Наиболее существенными результатами диссертационной работы являются:

1. Разработка алгебраической формализация для используемой процедурной конструкции извлечения зависимостей из структурных описаниях прецедентов, исходно заданных для обучения.
2. Идентификация и детальный анализ особого класса комбинаторных объектов (диаграмм взаимной вложенности классов эквивалентности описаний объектов-прецедентов), определяющих особенности алгоритмики порождения эмпирических зависимостей предлагаемым способом.
3. Доказательство оценок сложности вычислений, требуемых для решения основных переборных задач, которые возникают в процессе порождения диаграмм при обработке различных типов данных (вариантов описаний исходно заданных прецедентов).
4. Разработка метода дискретной оптимизации и управления комбинаторным перебором вариантов при формировании рассматриваемых диаграмм и анализе переносимости порождаемых эмпирических зависимостей на описания новых - ранее еще не изученных - прецедентов (метода последовательных приближений при формировании диаграмм вложенности, реализуемого в том числе и в режиме параллельных вычислений, что в совокупности позволяет снять ограничения на размеры эффективно обрабатываемых исходных коллекций прецедентов).
5. Демонстрация работоспособности предложенных математических моделей, методов и алгоритмов при решении прикладных задач в различных предметных областях.

Диссертационная работа является итогом многолетних исследований автора. М.И.Забежайло начал свои исследования в области искусственного интеллекта в конце 70-х годов.

Автор успешно выступал с научными докладами на престижных конференциях как в Российской Федерации (в том числе – и международных), так и за рубежом (в США, Франции, Венгрии и др.). Вклад М.И.Забежайло в разработку методов искусственного интеллекта получил высокие оценки специалистов. За время работы над диссертацией автор зарекомендовал себя как высококвалифицированный исследователь, способный самостоятельно ставить и целенаправленно решать сложные научные задачи.

М.И.Забежайло является известным специалистом в области искусственного интеллекта. Его работы имеют международное признание.

Представленные в диссертационном исследовании результаты послужили основой курса лекций по интеллектуальному анализу данных, ряд лет читавшегося студентам старших курсов Факультета управления и прикладной математики МФТИ.

Автор диссертационной работы является одним из создателей ДСМ-метода автоматического порождения гипотез, применяемого в России и ряде зарубежных стран.

Все теоретические результаты диссертационной работы получены М.И.Забежайло лично. Разработанные автором алгоритмы использовались при создании программного обеспечения интеллектуальных систем анализа данных в ВИНТИ РАН, ФИЦ ИУ РАН и ряде других организаций.

Совокупность теоретических результатов, полученный автором в диссертационном исследовании, можно квалифицировать как решение крупной научной задачи, научное достижение в области теоретической информатики.

По моему мнению, рассматриваемая диссертация является законченным научным трудом, имеющим высокую научную и практическую ценность. Результаты исследований прошли достоверную апробацию. Считаю, что диссертационная работа Михаила Ивановича Забежайло «Комбинаторные средства формализации эмпирической индукции» удовлетворяет требованиям ВАК Министерства образования и науки Российской Федерации, предъявляемым к докторским диссертациям, а ее автор заслуживает присвоения ему ученой степени доктора физико-математических наук по специальности 05.13.17 – Теоретические основы информатики.

Научный консультант
д.т.н., профессор,
Заслуженный деятель науки РФ

« 2 » сентября 2016 года



В.К. Финн

Подпись В.К. Финна заверяю.
Ученый секретарь ФИЦ ИУ РАН
д.т.н.



В.Н.Захаров

02.09.2016