

УДК 510.67, 519.24

О метриках многозначных логических высказываний и приложения метрик в базах знаний

Викентьев Александр Александрович

Новосибирский государственный университет, Институт математики им. С. Л. Соболева, e-mail: vikent@math.nsc.ru

В статье рассматриваются модели и формулы многозначных логик, имеющие различные применения в теории моделей и базах знаний, и, в частности для записи многозначных высказываний экспертов (логической базы знаний). С использованием методов математической логики и теории моделей для многозначных логик получены теоремы о богатых совокупностях формул (типов) без предположения стабильности (полной, метрической) многозначной модели, а с условием на класс ее расширений с той же (непрерывной) теорией рассмотрены и ортогональные случаи. Далее находятся «расстояния» на формулах (высказываниях) и степени (меры) нетривиальности (недоверности) формулы как мера неверности формулы на классе рассматриваемых многозначных моделей (возможных миров). Изучены свойства введенных расстояний и мер нетривиальности (недоверности) формул на классах многозначных моделей. Предложены различные способы задания на классах эквивалентных формул метрик, степеней нетривиальности (недоверности) и установлены для них полезные свойства, которые используются в алгоритмах кластеризации, построении решающих функций, распознавании образов и вопроса об эталонах в искусственном интеллекте. Из этого дается описание всевозможных метрик при фиксированном классе конечнозначных моделей фиксированной многозначной логики и/или ее ослабления (уменьшением списка аксиом).

Ключевые слова: теория моделей, двукардинальные(богатые) формулы(типы), многозначные модели непрерывной логики, расстояния на логических формулах –высказываниях экспертов(базы знаний), релевантные модели, метрики на классах эквивалентности, меры недоверности(нетривиальности), кластеризация, распознавание образов, искусственный интеллект.

В настоящее время изучаются теоретико-модельные метрики на множествах высказываний (базы знаний) многозначных и простейших семантически определенных релевантных логик с помощью релевантного класса моделей от переменных и их отрицаний. Решен вопрос об описании всех таких метрик, которые получаются с привлечением конечного релевантного класса моделей. Доклад посвящен также обобщению и уточнению результатов и теорем о двукардинальных (богатых) множествах типов, доказанных ранее в стабильном случае или с условиями стабильности, на случай многозначной и многосортной теории [1–3]. Некоторые из них вошли в диссертацию автора «Теории с покрытием и формульные подмножества» (ИМ СО РАН, Новосибирск, 1992, 134 с.) для семейств формул, а также опубликованы в сборнике Two cardinal theorems for sets of types in stable theory, посвященном 90-летию академика А. Д. Тайманова (Казахстан, Алма-Ата, 2007, с. 67–69), которые были представлены в Алма-Ате и Новосибирске на ежегодных Мальцевских чтениях с 2006 г., в том числе к 100-летию акад. А. И. Мальцева и др.

На случай богатых семейств типов над параметрами модели многосортной, многозначной теории s -компактными (насыщенными, однородными) измеримыми и вычислимыми моделями со свойством отделимости новых элементов, реализующих вычислимые типы (над малыми подмножествами модели) совместных с этими множествами, от элементов вложенной модели и наличия реализаций в большей (с богатым семейством) модели вполне определенных, вычисляемых (стабильных) типов или неразличимых элементов. Стабильность теорий не предполагается, а многие известные теоремы в обычном случае получаются как следствия. Также дан обзор имеющихся результатов по применению многозначных исчислений для кластеризации множеств высказываний.

Основными инструментами доказательств являются теоремы типа компактности, развитая техника современной теории моделей, вычислимости, в частности, для логических исчислений, локальной стабильности и наличия (даже локально) подходящих компактных измеримых (нужных малых мощностей *var kappa*) моделей теории со свойствами *var kappa*-отделимости над реализациями семейств стабильных (определимых, рекурсивных) типов.

Продолжено изучение спектра и числа двукардинальных и предельных моделей в классе теорий с покрытиями введенных автором. Рассмотрены вопросы определимости систем с метрикой в наследственно конечных надстройках, и о мощностях типово определимых подмножеств и их свойств двукардинальности. Интерес к этим вопросам и таким моделям имеет и прикладной характер в поиске наиболее информативных (нетривиальных) типов (формул) прикладной теории, логических закономерностей для кластеризации и упорядочения таких «знаний» с помощью привлечения конечных, многозначных, метрических или измеримых систем и расстояний между множествами моделей.

Все это служит для введения новых метрик на классах неэквивалентных формул (логических высказываний экспертов или типов (совместных совокупностей высказываний) на измеримых подклассах измеримых, вычисляемых (метрических, измеримых) моделей, необходимых для разработки алгоритмов распознавания образов, поиска закономерностей, обнаружения редких событий и кластеризации многозначных формул-знаний в различных логиках, например Лукасевича.

В настоящее время найдены различные теоретико-модельные новые метрики, разработаны методы кластеризации по метрикам для конечных множеств формул в различных логических исчислениях с привлечением различных подклассов релевантных моделей (частично совместно с аспирантом Авиловым М. С., Фефеловой В. В., Таганашкиной Л. А.), изучены различные индексы качества для сравнений и способы введения коллективных метрик. Проведены, вначале с Фефеловой В. В., модельные эксперименты с помощью

разработанной программы, поддерживающей все необходимые алгоритмы и нахождение по найденным расстояниям кластеризаций. Показано, что коллективные расстояния имеют более высокие индексы кластеризаций по сравнению с другими введенными метриками. В дальнейшем планируется использование лучших кластеризаций для локальной структуризации баз знаний. Новым шагом в приложениях наших подходов является использование различных малых конечных классов релевантных моделей, предложенных на Мальцевских чтениях автором доклада, и взятие по ним (в конечном числе) коллективного расстояния, которое обеспечивает эффективную и оптимальную коллективную кластеризацию данных множеств формул. Ответ на поставленный выше вопрос важен для полного учета различных расстояний для небольших конечных (малых) подклассов класса релевантных моделей, и будет использоваться в поисках лучшей из (разумных) возможных кластеризаций.

Теорема 1. *Найдены все способы задания теоретико-модельных метрик на формулах многозначной релевантной логики высказываний в классе конечных многозначных релевантных моделей (и при фиксированном размере N многозначной логики).*

Это позволяет проводить алгоритмы классификации таких высказываний, уменьшает трудоемкость алгоритма согласования решений и построения решающих функций.

Работа выполнена при финансовой поддержке РФФИ (проект № 18-07-600-а).

Список литературы

- [1] Vikentiev A. A. Distances and Degrees of Uncertainty in Many-Valued Propositions of Experts and Application of These Concepts in Problems of Pattern Recognition and Clustering // Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications. 2014. Vol. 24, N 4. P. 409-421.
- [2] Викентьев А. А., Фефелова В. В. Новые модельные расстояния и меры достоверности формул логики Лукасевича в автоматической кластеризации высказываний баз данных // Algebra and Model Theory 10. Collection of papers / eds.: A. G. Pinus, K. N. Ponomaryov, S. V. Sudoplatov, and E. I. Timoshenko. Novosibirsk : NSTU, 2015. P. 197–209.
- [3] Vikentiev A. A., Avilov M. C. New Model Distances and Uncertainty Measures for Multivalued Logic // Artificial Intelligence: Methodology,

Systems, and Applications. C. Dichev, G. Agre (eds.). Lecture Notes on Computer Science, LNCS 9883, 2016. P. 89–98.

On the Metrics on Multi-Valuable Logical Expression and Application of the Metrics in Databases

Vikentiev Alexandr Alexandrovich

Novosibirsk State University, Sobolev Institute of Mathematics SB RAS, e-mail: vikent@math.nsc.ru

The article discusses models and formulas of multivalued logics having various applications in model theory and knowledge bases, and, in particular, for recording ambiguous statements of experts (logical knowledge base). Using methods mathematical logic and model theory for multi-valued logics, theorems on rich collections of formulas (types) are obtained without assuming the stability of the (complete, metric) model, but with the condition on the class of its extensions with the same (continuous) theory, orthogonal cases are considered. We consider formulas of manyvalued logic. Further are « distances » on formulas (statements) and degrees (measures) nontriviality (uncertainty) of the formula, as a measure of the formula's infidelity on the class models under consideration (possible worlds). The properties of the introduced distances and measures of nontriviality (uncertainty) of formulas on classes of multivalued models. Methods are proposed assignment on classes of nonequivalent formulas of metrics, degrees of nontriviality (uncertainty) and set useful properties for them, which are used in algorithms clustering, the construction of deciding functions, pattern recognition and the question of standards in artificial intelligence. From this description of all possible metrics is given for a fixed class of finite-valued models of a fixed many-valued logic and / or its weakening (by reducing the list of axioms).

Keywords: model theory, two-cardinal (rich) formulas (types), multivalued models of continuous logic, distances on logical formulas—expressions of experts (knowledge bases), relevant models, metrics on equivalence classes, measures of uncertainty(non-triviality), clustering, pattern recognition, artificial intelligence.