

Министерство науки и высшего образования РФ
Федеральное государственное автономное
образовательное учреждение высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»
Институт математики и фундаментальной информатики
Кафедра высшей и прикладной математики

УТВЕРЖДАЮ

Заведующий кафедрой

_____ / С.Г. Мысливец

« ____ » _____ 2022 г.

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ
РАЗРАБОТКА АЛГОРИТМОВ ПОИСКА ФОРМАЛЬНЫХ
ПОНЯТИЙ В НЕЧЁТКИХ КОНТЕКСТАХ

Направление 01.04.02 Прикладная математика и информатика

Магистерская программа 01.04.02.01 Математическое моделирование

Руководитель	доцент, кандидат физико- математических наук	Д.В. Семенова
Выпускник		В.К. Гольшев
Нормоконтролер		Т.Н. Шипина

Красноярск 2022

АННОТАЦИЯ

Цель работы — разработка алгоритмов поиска формальных понятий в нечётких контекстах.

В результате исследований проведён обзор существующих подходов обобщения анализа формальных понятий на нечёткие контексты, разобраны подробные примеры для каждого подхода, разработаны алгоритмы FuzzyCbO и DirectCbO. Создан комплекс программ, реализующий алгоритмы Янга, Белоглавека, FuzzyCbO и DirectCbO, проведены вычислительные эксперименты для сравнения реализованных алгоритмов.

Ключевые слова: нечёткий анализ формальных понятий, нечёткий формальный контекст, решётка формальных понятий, алгоритм поиска формальных понятий.

ABSTRACT

The research aim is to devise algorithms for computing formal concepts from fuzzy contexts.

As a result of the research, a survey of existing approaches for generalization of formal concept analysis to fuzzy contexts was carried out, detailed examples for each approach were calculated, and FuzzyCbO and DirectCbO algorithms were developed. A software package has been created that implements the Yang, Belohlavek, FuzzyCbO and DirectCbO algorithms, and computational experiments have been carried out in order to compare the implemented algorithms.

Keywords: fuzzy formal concept analysis, fuzzy formal context, formal concept lattice, algorithm for computing formal concepts.

Содержание

Введение	3
1 Нечёткий анализ формальных понятий: обзор подходов	7
1.1 Классический анализ формальных понятий	7
1.2 Основные определения и обозначения нечёткого анализа формальных понятий	9
1.3 Постановка задачи нахождения всех формальных понятий в нечёткого контексте	14
1.4 Выводы по первой главе	18
2 Алгоритмы нахождения множества нечётких формальных понятий	19
2.1 Метод α -сечения	19
2.2 Обобщение алгоритма Гантера	23
2.3 Обобщение алгоритма Close-by-One	24
2.4 Выводы по второй главе	25
3 Вычислительные эксперименты	27
3.1 Описание программного комплекса	27
3.2 Влияние выбора операций на производительность алгоритмов подхода П1	27
3.3 Влияние выбора порога на производительность алгоритмов подхода П7	29
3.4 Выводы по третьей главе	30
Заключение	31
Список использованных источников	32
Список таблиц	38
Список иллюстраций	39
Приложение А. Алгоритмы анализа формальных понятий	40
Приложение Б. Демонстрационные примеры	42
Приложение В. Алгоритмы нечеткого анализа формальных понятий	54

ВВЕДЕНИЕ

Анализ формальных понятий (АФП) позволяет построить онтологическую модель предметной области, используя алгебраическую теорию решёток [1]. Предметная область задаётся формальным контекстом, который можно представить в виде объектно-признаковой таблицы (матрицей, заполненной нулями и единицами). С помощью АФП строится решётка формальных понятий, полученная из заданного формального контекста. Такие решётки можно визуализировать в виде диаграммы Хассе, явно отображающей иерархическую структуру решётки и предметной области. Формальное понятие в АФП соответствует философскому значению понятия, т.к. понятие — это пара множеств объектов и признаков (объёма и содержания); при увеличении объёма уменьшается содержание, и наоборот; также формальные понятия упорядочены отношением «быть более общим понятием» [8].

Анализ формальных понятий берёт своё начало в работе Р. Вилле [51]. Первые 10 лет своего существования АФП развивался в основном небольшой группой исследователей и студентов Р. Вилле в Германии. Поначалу АФП представлял только математический интерес, однако теперь он имеет приложения в большом количестве областей, таких как лингвистика и анализ текстов [25, 44, 45, 47], программная инженерия [30], искусственный интеллект [35], информационный поиск и другое [40–42, 48].

Классический анализ формальных понятий работает с бинарными контекстами в виде $(0,1)$ -матриц. На практике при измерении свойств объектов возникают ситуации, при которых оценка истинности утверждения «объект g обладает признаком m » является расплывчатой [11, 12]. Для оценивания таких суждений привлекают математический аппарат различных теорий неопределённости, таких как теория нечётких множеств, грубых множеств, теории вероятностей или их комбинаций [23, 43].

После появления и развития теории нечётких множеств [5, 10, 54] начинаются попытки обобщения АФП на нечёткие контексты. Внедрение нечёткой логики в классический анализ формальных понятий привело к возникновению и активному развитию нового направления интеллектуального анализа дан-

ных — нечёткого анализа формальных понятий (НАФП) [7]. Решётки нечётких понятий впервые были введены в работе [24]. В работах [18, 46] были заложены основы теории нечёткого анализа формальных понятий (НАФП). В последствии Белоглавок активно развивал НАФП [15, 17, 20, 21]. В то же время появлялись и другие способы обобщения АФП на нечёткие контексты [43].

Задача нахождения множества всех формальных понятий для бинарного контекста является комбинаторной задачей перечислительного типа и является $\#P$ -полной [36]. Высокая вычислительная сложность задачи АФП объясняется тем, что количество получаемых формальных понятий в общем случае экспоненциально зависит от размера рассматриваемого формального контекста. Для решения задачи АФП разработано большое количество алгоритмов [34]. Вычислительная сложность задачи НАФП нахождения множества всех нечётких формальных понятий превышает сложность задачи АФП, т.к. количество получаемых нечётких понятий для нечёткого контекста может в разы превышать число понятий для аналогичного бинарного контекста такого же размера. На данный момент актуальна разработка быстрых и эффективных алгоритмов, позволяющих находить множество всех нечётких формальных понятий. В работе [13] описывается два способа нахождения данного множества: сведение задачи НАФП к задаче АФП [21] и модификация существующих алгоритмов АФП для работы с нечёткими контекстами [6, 13, 15, 53].

Цели и задачи исследования. Целью исследования является разработка алгоритмов поиска формальных понятий в нечётких контекстах.

Для достижения цели были поставлены и решены следующие задачи.

1. Провести обзор существующих подходов обобщения анализа формальных понятий на нечёткие контексты.
2. Разработать и теоретически обосновать модификацию алгоритма Close-by-One для метода α -сечения.
3. Разработать и теоретически обосновать модификацию алгоритма Close-by-One для подхода Белоглавека.
4. Создать комплекс программ, реализующий алгоритмы Янга, Белоглавека, FuzzyCbO, DirectCbO, для проверки результативности разработанных

алгоритмов на случайных нечётких контекстах.

Апробация работы. Результаты работы докладывались и обсуждались на семи конференциях:

- XV Международной конференции студентов, аспирантов и молодых ученых «Перспектив Свободный — 2019» (Красноярск, 2019, диплом I степени),
- VII Международной молодежной научной конференции «Математическое и программное обеспечение информационных, технических и экономических систем» (Томск, 2019, диплом лауреата),
- XVIII Международной конференции имени А.Ф. Терпугова «Информационные технологии и математическое моделирование» (Саратов, 2019, диплом победителя),
- XIX Международной конференции имени А.Ф. Терпугова «Информационные технологии и математическое моделирование» (Томск, 2020, диплом победителя),
- Международной научной конференции «Математическое и программное обеспечение информационных, технических и экономических систем» (Томск, 2020, диплом лауреата),
- XVII Международной конференции студентов, аспирантов и молодых ученых «Перспектив Свободный — 2021» (Красноярск, 2021, диплом II степени),
- VIII Международной молодежной научной конференции «Математическое и программное обеспечение информационных, технических и экономических систем» (Томск, 2021, диплом лауреата).

По тематике диссертации опубликовано 6 работ, индексируемых РИНЦ [2–7].

Структура и объём магистерской диссертации. Работа состоит из введения, трёх глав, заключения, списка литературы, трёх приложений. Общий объём составляет 55 страниц, включая приложения; иллюстративный материал представлен 6 рисунками и 11 таблицами; список литературы содержит 54 наименования.

Первая глава диссертации посвящена основным определениям и обозначе-

ниям АФП и НАФП, обзору подходов к обобщению АФП на нечёткие контексты, постановке задачи. Во второй главе рассматриваются существующие и разработанные в данной работе алгоритмы, программно реализованные в ходе исследования. Третья глава посвящена описанию разработанного программного комплекса, с помощью которого проводились вычислительные эксперименты для сравнения реализованных алгоритмов. Приложения содержат псевдокоды рассмотренных существующих алгоритмов, а также подробные примеры вычисления стрелочных операторов для каждого подхода НАФП.

ЗАКЛЮЧЕНИЕ

Основные результаты диссертационной работы представлены ниже.

1. Произведён обзор существующих подходов обобщения АФП на нечёткие контексты.
2. Продемонстрирована на примерах работа стрелочных операторов в различных подходах.
3. Разработан и теоретически обоснован алгоритм FuzzyCbO для метода α -сечения.
4. Разработан алгоритм DirectCbO для проверки гипотезы о возможности обобщения алгоритма Close-by-One для работы с нечёткими контекстами в рамках подхода П1.
5. Частично подтверждена выдвинутая гипотеза о возможности обобщения алгоритма Close-by-One для работы с нечёткими контекстами в рамках подхода П1.
6. Создан комплекс программ, реализующий алгоритмы Янга, Белоглавека, FuzzyCbO и DirectCbO.
7. Проведена проверка результативности реализованных алгоритмов на случайных нечётких контекстах.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Биркгоф, Г. Теория решеток : пер. с англ. / Г. Биркгоф ; Главная редакция физико-математической литературы. — Москва : Наука, 1984. — 568 с.
2. Голышев, В. К. Анализ формальных нечетких понятий / В. К. Голышев // Проспект Свободный — 2019: материалы XV Международной конференции студентов, аспирантов и молодых ученых / ответственный за выпуск С. К. Франчук. — Красноярск, 2019 — С. 902–905.
3. Голышев, В. К. Извлечение формальных понятий из нечеткого контекста / В. К. Голышев // Информационные технологии и математическое моделирование (ИТММ–2019): Материалы XVIII Международной конференции имени А.Ф. Терпугова (26–30 июня 2019 г.). — Томск: Изд-во НТЛ, 2019. — Часть 1. — С. 120–125.
4. Голышев, В. К. Методы и алгоритмы решения задачи поиска формальных понятий для бинарных и нечётких контекстов / В. К. Голышев, Д. В. Семенова // Математическое и программное обеспечение информационных, технических и экономических систем : материалы Международной научной конференции. — Томск, 2020. — С. 11–18.
5. Голышев, В. К. Методы нечёткой логики в анализе формальных понятий / В. К. Голышев // Математическое и программное обеспечение информационных, технических и экономических систем: Материалы VII Международной молодежной научной конференции. — Томск, 2019. — С. 269–276.
6. Голышев, В. К. Нечёткий анализ формальных понятий: метод α -сечения / В. К. Голышев, Д. В. Семенова // Информационные технологии и математическое моделирование (ИТММ-2020): Материалы XIX Международной конференции имени А. Ф. Терпугова (2–5 декабря 2020 г.). — Томск: Изд-во НТЛ, 2021— С. 460–465.
7. Голышев, В. К. Нечёткий анализ формальных понятий: обзор подходов на примере / В. К. Голышев // Математическое и программное обеспечение информационных, технических и экономических систем: Материалы VIII Международной молодежной научной конференции. — Томск, 2021. — С. 28–32.

8. Игнатов, Д. И. Анализ формальных понятий : от теории к практике / Д. И. Игнатов, Р. Э. Яворский // Анализ изображений, сетей и текстов. — 2012. — С. 1–12.
9. Кузнецов, С. О. Быстрый алгоритм построения всех пересечений объектов из конечной полурешетки / С. О. Кузнецов // Научно-техническая информация. — 1993. — № 1. — С. 17–20.
10. Новак, В. Математические принципы нечёткой логики / В. Новак, И. Перфильева, И Мочкорж // Пер. с англ.; под ред. Аверкина А. Н. — М.: ФИЗМАТЛИТ, 2006. — 352 с.
11. Смирнов, С. В. Онтологический анализ предметных областей моделирования / С. В. Смирнов // Известия Самарского научного центра РАН. — 2001. — Т. 3, № 1— С. 62–70.
12. Офицеров, В. П. Нечёткий анализ формальных понятий при разработке онтологий / В. П. Офицеров, С. В. Смирнов // Онтология проектирования. — 2017. — Т. 4, № 26— С. 487–495.
13. Belohlavek, R. Algorithms for fuzzy concept lattices / R. Belohlavek // Proc. Fourth Int. Conf. on Recent Advances in Soft Computing. — 2002. — P. 200–205.
14. Belohlavek, R. A note on variable threshold concept lattices: Threshold-based operators are reducible to classical concept-forming operators / R. Belohlavek // Information Sciences. — 2007. — Vol. 177, Iss. 15. — P. 3186–3191.
15. Belohlavek, R. Computing the Lattice of All Fixpoints of a Fuzzy Closure Operator / R. Belohlavek, D. Bernard, J. Outrata, V. Vychodil // IEEE Transactions on Fuzzy Systems. — 2010. — Vol. 18, Iss. 3. — P. 546–557.
16. Belohlavek, R. Formal concept analysis and linguistic hedges / R. Belohlavek, V. Vychodil // International Journal of General Systems. — 2012. — Vol. 41, Iss. 5. — P. 1–30.
17. Belohlavek, R. Fuzzy Closure Operators / R. Belohlavek // Journal of Mathematical Analysis and Applications. — 2001. — V. 262. — P. 473–489.
18. Belohlavek, R. Fuzzy Concepts and Conceptual Structures: Induced Similarities / R. Belohlavek // Joint Conf. Inf. Sci. 1998, Durham, NC. —

1998. — V. I. — P.179–182.
19. Belohlavek, R. Fuzzy Concept Lattices Constrained by Hedges / R. Belohlavek, V. Vychodil // *J. Adv. Comput. Intell. Intell. Informatics.* — 2007. — Vol. 11. — P. 536–545.
 20. Belohlavek, R. Fuzzy Galois Connections / R. Belohlavek // *Mathematical Logic Quarterly* 45. — 1999. — P. 497–504.
 21. Belohlavek, R. Reduction and a Simple Proof of Characterization of Fuzzy Concept Lattices / R. Belohlavek // *Fundamenta Informaticae.* — 2001. — Vol. 46, Iss. 4. — P. 277–285.
 22. Belohlavek, R. Thresholds and Shifted Attributes in Formal Concept Analysis of Data with Fuzzy Attributes / R. Belohlavek, J. Outrata, V. Vychodil // *Conceptual Structures: Inspiration and Application. ICCS 2006. Lecture Notes in Computer Science.* — 2006. — Vol. 4068. — P. 117–130.
 23. Belohlavek, R. What is a fuzzy concept lattice? / R. Belohlavek, V. Vychodil // *Proc. CLA 2005, 3rd Int. Conference on Concept Lattices and Their Applications*, 162. — 2005.— P. 34–45.
 24. Burusco, A. The study of L-fuzzy concept lattice / A. Burusco, R. Fuentes-Gonzalez // *Mathware & Soft Computing.* — 1994. — Vol. 3. — P. 209–218.
 25. Carpineto, C. Using Concept Lattices for Text Retrieval and Mining / C. Carpineto, G. Romano // *Formal Concept Analysis. Lecture Notes in Computer Science.* — 2005. — Vol. 3626. — P. 288–303.
 26. Elloumi, S. A multi-level conceptual data reduction approach based on the Lukasiewicz implication / S. Elloumi, J. Jaam, A. Hasnah, A. Jaoua, I. Nafkha // *Information Sciences.* — 2004. — Vol. 163, Iss. 4. — P. 253–262.
 27. Ganter, B. Formal concept analysis: Mathematical Foundations / B. Ganter, R. Wille // *Springer.* — 1999. — P. 284.
 28. Georgescu, G. Concept Lattices and Similarity in Non-Commutative Fuzzy Logic / G. Georgescu, A. Popescu // *Fundamenta Informaticae.* — 2002. — Vol. 55, Iss. 1. — P. 23–54.
 29. Georgescu, G. Non-Commutative Fuzzy Galois Connections / G. Georgescu, A. Popescu // *Soft Computing.* — 2003. — Vol. 7. — P. 458–467.

30. Hesse, W. Formal Concept Analysis Used for Software Analysis and Modelling / W. Hesse, T. Tilley // Formal Concept Analysis. Lecture Notes in Computer Science. — 2005. — V. 3626. — P. 288–303.
31. Krajci, S. A generalized concept lattice / S. Krajci // Logic Journal of IGPL. — 2005. — Vol. 13, Iss. 5. — P. 543–550.
32. Krajci, S. Cluster Based Efficient Generation of Fuzzy Concepts / S. Krajci // Neural Network World. — 2003. — Vol. 13, Iss. 5. — P. 521–530.
33. Krajci, S. Every concept lattice with hedges is isomorphic to some generalized concept lattice / S. Krajci // CEUR Workshop Proceedings. — 2009. — P. 1–9.
34. Kuznetsov, S. O. Comparing performance of algorithms for generating concept lattices / S. Kuznetsov, S. Obiedkov // J. Exp. Theor. Artif. Intell. — 2002. — V. 14. — P. 189–216.
35. Kuznetsov, S. Machine Learning and Formal Concept Analysis / S. Kuznetsov // Lecture Notes in Artificial Intelligence. — № 2961. — 2005. — P. 287–312.
36. Kuznetsov, S. O. Mathematical aspects of concept analysis / S. O. Kuznetsov // Journal of Mathematical Sciences. — 1996. — Vol. 80, Iss. 2. — P. 1654–1698.
37. Medina, J. Formal concept analysis via multi-adjoint concept lattices / J. Medina, M. Ojeda-Aciego, J. Ruiz-Calvino // Fuzzy Sets and Systems. — 2009. — Vol. 160, Iss. 2. — P. 130–144.
38. Medina, J. On Multi-adjoint Concept Lattices: Definition and Representation Theorem / J. Medina, M. Ojeda-Aciego, J. Ruiz-Calvino // ICFCA 2007. Lecture Notes in Computer Science. — 2007. — Vol. 4390. — P. 197–209.
39. Medina, J. Relating Generalized Concept Lattices and Concept Lattices for Non-Commutative Conjunctors / J. Medina, M. Ojeda-Aciego, J. Ruiz-Calvino // Applied Mathematics Letters. — 2008. — Vol. 21, Iss. 12. — P. 1296–1300.
40. Poelmans, J. Formal Concept Analysis in Knowledge Discovery: a Survey / J. Poelmans, P. Elzinga, S. Viaene, G. Dedene // Conceptual Structures: From Information to Intelligence. — 2010. — P. 139–153.


41. Poelmans, J. Formal concept analysis in knowledge processing: A survey on applications / J. Poelmans, S. O. Kuznetsov, D. I. Ignatov, G. Dedene // *Expert Systems with Applications*. — 2013. — Vol. 40, Iss. 16. — P. 6538–6560.
42. Poelmans, J. Formal Concept Analysis in Knowledge Processing: a Survey on Models and Techniques / J. Poelmans, S. O. Kuznetsov, D. I. Ignatov, G. Dedene // *Expert Systems with Applications*. — 2013. — Vol. 40, Iss. 16. — P. 6601–6623.
43. Poelmans, J. Fuzzy and rough formal concept analysis: a survey / J. Poelmans, D. I. Ignatov, S. O. Kuznetsov, G. Dedene // *International Journal Of General Systems* — 2014.— P. 105–134.
44. Poelmans, J. Human-Centered Text Mining: a New Software System / J. Poelmans, P. Elzinga, A. A. Neznanov, G. Dedene, S. Viaene, S. O. Kuznetsov // *Advances in Data Mining. Applications and Theoretical Aspects*. — 2012. — P. 258–272.
45. Poelmans, J. Text Mining Scientific Papers: a Survey on FCA-Based Information Retrieval Research / J. Poelmans, D. I. Ignatov, S. Viaene, G. Dedene, S. O. Kuznetsov // *Advances in Data Mining. Applications and Theoretical Aspects*. — 2012. — P. 273–287.
46. Pollandt, S. *Fuzzy Begriffe* / S. Pollandt ; Springer. — 1997.
47. Priss, U. Linguistic Applications of Formal Concept Analysis / U. Priss // *Formal Concept Analysis. Lecture Notes in Computer Science*. — 2005. — V. 3626. — P. 149–160.
48. Priss, U. Formal Concept Analysis in Information Science / U. Priss // *Annual Review of Information Science and Technology*. — 2006. — V. 40, Iss. 1. — P. 521–543.
49. Shao, M. W. Information Granularity Lattices / M. W. Shao, W. X. Zhang // *Proceedings of the Sixth International Conference on Machine Learning and Cybernetics*. — 2007. — P. 3728–2733.
50. Wille, R. Conceptual Knowledge Processing in the Field of Economics / R. Wille // *Formal Concept Analysis. Lecture Notes in Computer Science*. — V. 3626. — P. 226–249.

51. Wille, R. Restructuring lattice theory: an approach based on hierarchies of concepts / R. Wille // *Ordered sets* — 1982. — P. 445–470.
52. Yahia, S. B. Discovering Knowledge from Fuzzy Concept Lattice / S. B. Yahia, A. Jaoua // *Data Mining and Computational Intelligence*. — 2001. — P. 167–190.
53. Yang, K. M. Fuzzy Concept Mining based on Formal Concept Analysis / K. M. Yang, E. H. Kim, S. H. Hwang, S. H. Choi // *Int. J. of Computers*. — 2008. — Vol. 2, Iss. 3. — P. 279–290.
54. Zadeh, L. Fuzzy sets / L. Zadeh // *Information and control*. — 1965. — Vol. 8, Iss. 3. — P. 338–353.

Министерство науки и высшего образования РФ
Федеральное государственное автономное
образовательное учреждение высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»
Институт математики и фундаментальной информатики
Кафедра высшей и прикладной математики

УТВЕРЖДАЮ

Заведующий кафедрой


 / С.Г. Мысливец
« 17 » июня 2022 г.

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ
РАЗРАБОТКА АЛГОРИТМОВ ПОИСКА ФОРМАЛЬНЫХ
ПОНЯТИЙ В НЕЧЁТКИХ КОНТЕКСТАХ

Направление 01.04.02 Прикладная математика и информатика

Магистерская программа 01.04.02.01 Математическое моделирование


Руководитель


17.06.22

доцент, кандидат физико-
математических наук


Д.В. Семенова

Выпускник


17.06.22

В.К. Гольшев

Нормоконтролер


24.06.22

Т.Н. Шипина

Красноярск 2022