



УДК 004.056.53

ПОДДЕРЖАНИЕ ИНТЕРНАЛЬНОЙ НЕПРОТИВОРЕЧИВОСТИ АЛГЕБРАИЧЕСКИХ БАЙЕСОВСКИХ СЕТЕЙ С ЛИНЕЙНОЙ И ЗВЕЗДЧАТОЙ СТРУКТУРОЙ

Н.А. Харитонов^а

^а Санкт-Петербургский институт информатики и автоматизации Российской академии наук (СПИИРАН),
Санкт-Петербург, 199178, Российская Федерация
Адрес для переписки: nikita.kharitonov95@yandex.ru

Информация о статье

Поступила в редакцию 26.09.18, принята к печати 28.10.18

doi: 10.17586/2226-1494-2018-18-6-1108-1117

Язык статьи – русский

Ссылка для цитирования: Харитонов Н.А. Поддержание интернальной непротиворечивости алгебраических байесовских сетей с линейной и звездчатой структурой // Научно-технический вестник информационных технологий, механики и оптики. 2018. Т. 18. № 6. С. 1108–1117. doi: 10.17586/2226-1494-2018-18-6-1108-1117

Аннотация

Предмет исследования. При работе с алгебраическими байесовскими сетями необходимо обеспечивать непротиворечивость оценок вероятностей составляющих такие сети элементов. Существует несколько подходов к автоматизации поддержания непротиворечивости, различающихся вычислительной сложностью (временем исполнения). Эта сложность зависит от структуры сети и от выбранного вида непротиворечивости. Выполнено основанное на статистических оценках сравнение времени поддержания интернальной непротиворечивости в алгебраических байесовских сетях с линейной и звездчатой структурой и времени поддержания непротиворечивости фрагмента знаний, покрывающих такие сети. **Метод** основан на сокращении числа переменных и условий в задачах линейного программирования, решение которых обеспечивает поддержание интернальной непротиворечивости. Проведен эксперимент, демонстрирующий различия между временем поддержания непротиворечивости для различных по глобальной структуре представлений алгебраических байесовских сетей. **Основные результаты.** Представлена улучшенная версия алгоритма поддержания интернальной непротиворечивости. Упрощены решаемые задачи линейного программирования в сравнении с предыдущей версией алгоритма. Сформулированы и доказаны две теоремы, уточняющие оценки числа переменных и условий в решаемых задачах линейного программирования, а также количества самих задач. Выполнен эксперимент, показавший, что предложенная программная реализация превосходит по скорости работы программную реализацию для полного фрагмента знаний. **Практическая значимость.** Полученные результаты могут найти применение в машинном обучении алгебраических байесовских сетей (в том числе синтезе их глобальных структур). Предложенный метод позволяет при обучении и дальнейшей обработке сети оптимально синтезировать глобальные ее структуры, для которых достаточно использовать поддержание интернальной непротиворечивости. Благодаря использованию метода эти процессы будут иметь приемлемую вычислительную сложность.

Ключевые слова

алгебраические байесовские сети, интернальная непротиворечивость, задача линейного программирования, теоретические оценки, эмпирические оценки, фрагмент знаний

Благодарности

Работа выполнена в рамках проекта по государственному заданию СПИИРАН № 0073-2018-0001, при финансовой поддержке РФФИ, проект №18-01-00626 – Методы представления, синтеза оценок истинности и машинного обучения в алгебраических байесовских сетях и родственных моделях знаний с неопределенностью: логико-вероятностный подход и системы графов.

MAINTAINING OF INTERNAL CONSISTENCY OF ALGEBRAIC BAYESIAN NETWORKS WITH LINEAR AND STELLATE STRUCTURE

N.A. Kharitonov^а

^а Saint Petersburg Institute for Informatics and Automation RAS (SPIIRAS), 199178, Saint Petersburg, Russian Federation
Corresponding author: nikita.kharitonov95@yandex.ru

Article info

Received 26.09.18, accepted 28.10.18

doi: 10.17586/2226-1494-2018-18-6-1108-1117

Article in Russian

For citation: Kharitonov N.A. Maintaining of internal consistency of algebraic Bayesian networks with linear and stellate structure. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2018, vol. 18, no. 6, pp. 1108–1117 (in Russian). doi: 10.17586/2226-1494-2018-18-6-1108-1117

Abstract

Subject of Research. When working with algebraic Bayesian networks, it is necessary to ensure their correctness in terms of the consistency of the probability estimates of their constituent elements. There are several approaches to automating the maintenance of consistency, characterized by their computational complexity (execution time). This complexity depends on the network structure and the chosen type of consistency. The time for internal consistency maintenance in algebraic Bayesian networks with linear and stellate structure is compared with the time for consistency maintenance of a knowledge pattern covering such networks. The comparison is based on statistical estimates. **Method.** The essence of the method lies in reducing the number of variables and conditions in linear programming problems which solution ensures the maintenance of internal consistency. An experiment was carried out demonstrating the differences between the time of consistency maintenance for different algebraic Bayesian networks with a global structure. **Main Results.** An improved version of the algorithm for internal consistency maintenance is presented. Solvable linear programming problems are simplified in comparison with the previous version of the algorithm. Two theorems are formulated and proved, refining the estimates of the number of variables and conditions in the linear programming problems to be solved, as well as the number of the problems themselves. An experiment is performed, which showed that the proposed software implementation of internal consistency maintenance is superior in working time to software implementation of the consistency maintenance of a complete knowledge pattern. **Practical Relevance.** The results obtained can be applied in machine learning of algebraic Bayesian networks (including the synthesis of their global structures). The proposed method provides optimal synthesis of global network structures for which it is enough to use the maintenance of internal consistency during learning and further network processing. Owing to the method application these processes will have acceptable computational complexity.

Keywords

algebraic Bayesian networks, internal consistency, linear programming problem, theoretical estimates, empirical estimates, knowledge pattern

Acknowledgements

The research was carried out in the framework of the project on SPIIRAS state assignment No. 0073-2018-0001, with the financial support of the RFBR (project No. 18-01-00626 Methods of representation, synthesis of truth estimates and machine learning in algebraic Bayesian networks and related knowledge models with uncertainty: the logic-probability approach and graph systems).

Введение

Чтобы упростить восприятие формальных построений используем фразеологизмы с дальнейшим приданием одной из их сторон математической интерпретации. Рассмотрим пример: «Если утром трава сухая – к ночи жди дождя». С математической точки зрения эта фраза описывает причинно-следственную связь¹: «При выполнении условия X – «утром трава сухая» выполняется утверждение Y – «ночью будет дождь»». В данном утверждении допущена неточность – приметы «работают» не всегда. Иначе говоря, условие X не гарантирует наверняка наступления утверждения Y .

Далеко не всегда вероятность наступления заявленного события легко описать одним числом, причиной чего может служить потеря или отсутствие части данных, различие экспертных мнений, сложность «перевода» естественного языка в цифры. Последнее иллюстрирует приведенный пример: с точки зрения одного человека, данная примета означает, что дождь будет с вероятностью 87 %, с точки зрения другого – 43 %. Одним из возможных решений данной проблемы является использование неточных (интервальных) вероятностей [4–7]. Для рассматриваемого примера вероятность дождя будет находиться в интервале $[0,4;1]$.

Данный подход используется в алгебраических байесовских сетях (АБС) [1, 8–10], относящихся к одному из классов вероятностных графических моделей [11–14]. Структурно АБС представлены ненаправленными графами с идеалами конъюнктов в узлах, каждому из которых задана вероятность. Последняя может быть как точной, так и интервальной. Для оптимизации работы АБС разбиваются на части, называемые фрагментами знаний (ФЗ). Фрагмент знаний представляет собой полностью представленный в сети идеал конъюнктов.

Проиллюстрируем принцип, согласно которому строится АБС, добавив к первой примете еще две: «Красного цвета солнце на восходе – к большому ветру» и «Беловатые облака над водой, быстро исчезающие при восходе солнца, – к ясной погоде». Если мы хотим узнать, с какой вероятностью сегодня будет грибной дождь (дождь и ясная погода) или с какой вероятностью будет ясный, но ветреный день, нам достаточно построить АБС над двумя фрагментами знаний по два атома в каждом (рис. 1, а).

Стоит отметить, что более корректно было бы вместо фразы «Сухая трава утром» оставить описание «Дождь», «Беловатые облака над водой» – солнце, «Красное солнце на восходе» – ветер, поскольку с формальной точкой зрения конъюнкция «дождь и солнце» («грибной дождь») является конъюнкцией

¹ Более осторожная и общая точка зрения состоит в том, что рассматриваемый фразеологизм описывает некоторую взаимосвязь или взаимозависимость между утверждениями, тонко указывая на недетерминированность такой связи: «жди дождя», но не «будет дождь» [1–3].

именно дождя и именно солнца, а не сухой травы утром (из которой следует дождь) и беловатых облаков над водой (из которых следует солнце). Однако, в целях придания примеру некой словесной изящности в этом и следующем примерах АБС представлена именно таким образом.

Для получения информации о том, будет ли завтра грибной дождь с ветром, АБС можно достроить до полного фрагмента знаний, представленного на рис. 1, б.

Для работы с таким фрагментом знаний также нужно иметь оценки вероятностей того, что утром трава сухая, облака над водой беловатые и солнце красное на восходе. Их можно получить, например, на основе наблюдений за некоторый временной промежуток.

На основе построенной АБС можно получать информацию о некоторой пропозициональной формуле. Данный процесс назван априорным выводом [1]. Для приведенного выше примера такой «формулой» может быть: «Утром будет красное солнце на восходе и грибной дождь с безветренной погодой».

Также в АБС при поступлении информации об истинности некоторой входящей пропозициональной формулы (в теории называемой свидетельством) существующие оценки могут меняться. Данный процесс назван апостериорным выводом [1, 2]. Например, если утром была сухая трава, но дождя не было, вероятность наступления грибного дождя и дождя с ветром несколько уменьшится на фоне возрастания вероятности того, что трава сухая утром.

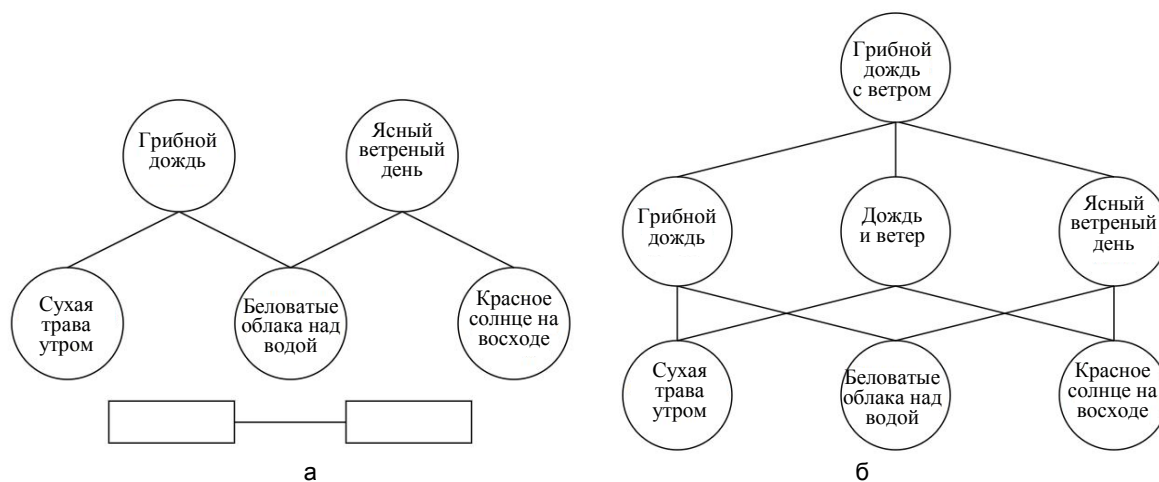


Рис. 1. Пример алгебраической байесовской сети с двумя фрагментами знаний (а) и полного фрагмента знаний (б)

При создании изменении, дополнении АБС возникает необходимость в проверке соответствия представленных в ней оценок теоретическим и предметным условиям, для чего вводится понятие непротиворечивости. При этом одной из степеней непротиворечивости является интернальная.

АБС называется интернально непротиворечивой, если для каждого конъюкта при любом скалярном значении из интервала оценки вероятности его истинности можно взять такие скалярные оценки во всех остальных конъюнктах, что получившаяся АБС будет экстернально непротиворечивой [15].

В работах [16] описаны теоремы, доказывающие, что для АБС, не имеющей циклов (ациклической АБС), из интернальной непротиворечивости следует глобальная. Большая часть алгоритмов логико-вероятностного вывода описана именно для ациклических АБС. Кроме того, существуют способы приведения циклических АБС к ациклическим, со структурой в виде дерева смежности. В связи с этим нужно в тщательно исследовать сложность алгоритма поддержания интернальной непротиворечивости, зачастую используемого для пред- или постобработки АБС в прочих алгоритмах.

На данный момент не производилось эмпирических исследований, позволяющих однозначно утверждать, что алгоритм поддержания интернальной непротиворечивости по времени работы более выгоден, чем поддержание непротиворечивости полноценного фрагмента знаний над тем же числом атомов. Таким образом, целью настоящей работы является выбор наиболее быстрого алгоритма поддержания непротиворечивости АБС с интервальными оценками вероятности истинности конъюнктов при заданных условиях на основе стохастических оценок скорости его работы.

Алгоритм поддержания интернальной непротиворечивости

Алгоритм, представленный в работах [1, 9, 15, 17], неоднократно модифицирован. Текущая реализованная версия представлена в работе [9]. Однако в ходе данного исследования был разработан алгоритм, в котором сокращено число условий в решаемой в ходе его работы задачи линейного программирования (ЗЛП) [18–20] (рис. 2).

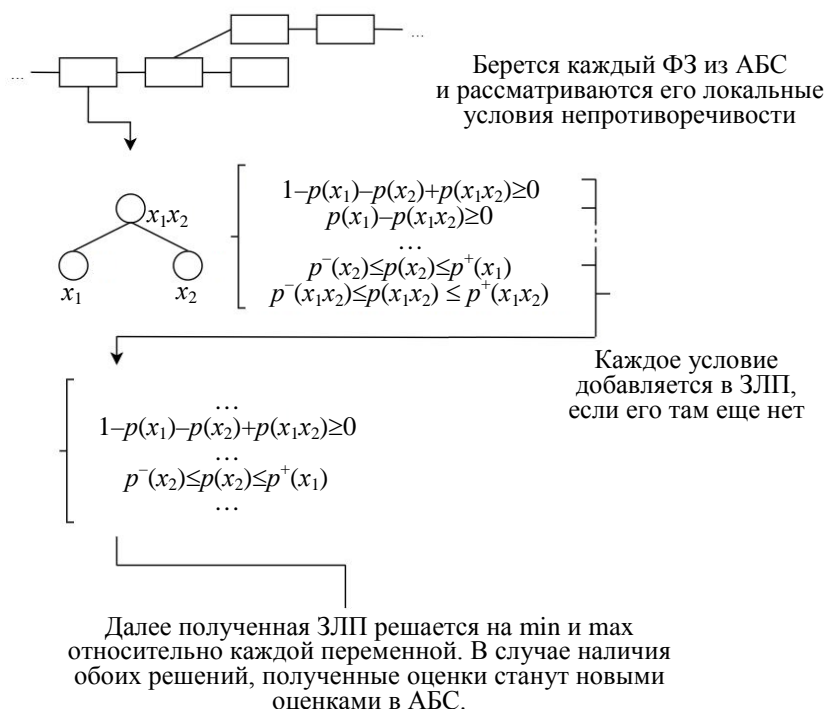


Рис. 2. Алгоритм поддержания интернальной непротиворечивости

Охарактеризуем теоретическую оценку сложности представленного алгоритма, доказав две теоремы. Их доказательство опирается на теорему 1.

Теорема 1 [17]. Если АБС построена над n атомами, имеет k фрагментов знаний, построенных не более чем над r атомами, то:

- задача линейного программирования имеет не более $O(k) = k(2^r - 1)$ переменных;
- задача линейного программирования имеет не более $O(k) = k(2^r)$ условий из теории вероятности;
- задача линейного программирования имеет не более $2k(2^r - 1)$ условий из предметной области;
- в ходе поддержания интернальной непротиворечивости будет решено не более $2k(2^r - 1)$ задач линейного программирования.

В рамках описанного выше алгоритма можно ввести ограничение снизу и уточнить ограничение сверху на указанные в теореме 1 величины. Для этого будут введены переменные s_{\min}, s_{\max} , которые описывают наименьшее и наибольшее число атомов на пересечении двух фрагментов знаний для АБС.

Теорема 2. Пусть АБС построена над n атомами, имеет k фрагментов знаний, построенных не более чем над r атомами. Если число пересечений между фрагментами знаний не меньше s_{\min} , то задача линейного программирования в алгоритме имеет не более:

- $O(k) = k(2^r - 2^{s_{\min}})$ переменных;
- $k(2^r - 2^{s_{\min}})$ условий из теории вероятности;
- $2k(2^r - 2^{s_{\min}})$ условий из предметной области.

Также для поддержания интернальной непротиворечивости необходимо решить не более $k(2^r - 2^{s_{\min}})$ задач линейного программирования.

Доказательство. Пункт 1 теоремы 2 следует из процесса построения ЗЛП: повторяющиеся конъюнкты являются одними и теми же переменными для ЗЛП. Наименьшее число повторяющихся конъюнктов $k(2^{s_{\min}})$ следует из условия теоремы. Таким образом, согласно теореме 1, наибольшее количество переменных $k(2^r - 2^{s_{\min}})$.

Пункт 2 теоремы 2 доказывается следующим образом. Число условий из теории вероятности в модифицированном алгоритме уменьшается на число условий из ФЗ, образующихся на местах пересечения. Наименьшее число таких условий $k(2^{s_{\min}})$. Таким образом, основываясь на теореме 1, наибольшее количество условий из теории вероятности равно $k(2^r - 2^{s_{\min}})$.

Пункт 3 теоремы 2 следует из первого, так как число условий из предметной области равно удвоенному числу переменных.

Пункт 4 теоремы 2 также следует из первого: так как ЗЛП решается для каждой переменной, то число решаемых ЗЛП равно удвоенному количеству переменных.

■ Следующая теорема рассматривает алгебраическую сеть с введенными ограничениями на наибольшее количество атомов на пересечении ФЗ.

Теорема 3. Пусть АБС построена над n атомами, имеет k фрагментов знаний, построенных не более чем над r атомами. Если число пересечений между фрагментами знаний не более s_{\max} , то ЗЛП в алгоритме имеет не менее:

- $O(k) = k(2^r - 2^{s_{\max}})$ переменных;
- $k(2^r - 2^{s_{\max}})$ условий из теории вероятности;
- $2k(2^r - 2^{s_{\max}})$ условий из предметной области.

Также для поддержания интернальной непротиворечивости необходимо решить не менее $k(2^r - 2^{s_{\max}})$ ЗЛП.

Доказательство аналогично доказательству предыдущей теоремы: число повторяющихся конъюнктов не более $k(2^{s_{\max}})$, т.е. наименьшее количество переменных равно $k(2^r - 2^{s_{\max}})$.

Остальные пункты следуют из первого.

■ Кроме того, на основе данных двух теорем можно сформулировать следствие для случая, когда количество атомов на пересечении фрагментов знаний постоянно.

Следствие. Пусть АБС построена над n атомами, имеет k фрагментов знаний, построенных не более чем над r атомами. Если число пересечений s между фрагментами знаний постоянно, то ЗЛП в Алгоритме 1 имеет:

- $O(k) = k(2^r - 2^s)$ переменных;
- $k(2^r - 2^s)$ условий из теории вероятности;
- $2k(2^r - 2^s)$ условий из предметной области.

Также для поддержания интернальной непротиворечивости необходимо решить не менее $k(2^r - 2^s)$ ЗЛП.

Доказательство напрямую следует из теорем 2, 3: $s = s_{\min} = s_{\max}$.

■ Также в этой части стоит отметить, что при поддержании непротиворечивости фрагмента знаний над n атомами справедливо [17], что задача линейного программирования имеет

- $2^n - 1$ переменных;
- не более 2^n условий из теории вероятности.

Описание эксперимента

С целью компаративного анализа (получения эмпирических оценок) скоростей поддержания интернальной и глобальной непротиворечивости были проведены два эксперимента. В ходе каждого над одинаковым количеством атомов были взяты фрагмент знаний (иллюстрирует поддержание глобальной непротиворечивости без использования интернальной) и три АБС. Эксперименты различались структурой АБС, однако в обоих случаях они строились над фрагментами знаний с одинаковым количеством конъюнктов как внутри ФЗ, так и на пересечениях. В каждом из экспериментов АБС строились над фрагментами знаний над:

- двумя атомами и одним атомом на пересечениях;
- тремя атомами и одним атомом на пересечениях;
- тремя атомами и двумя атомами на пересечениях.

Число атомов в целом варьировало от 3 до 8. Верхняя граница обусловлена тем, что поддержание непротиворечивости для фрагмента знаний над алфавитом из 9 атомов занимает время, неприемлемо большое в рамках данного эксперимента (более 10 ч). Отметим, что при этом поддержание интернальной непротиворечивости АБС (с рассмотренными глобальными структурами) над алфавитом из 9 и более атомов остается вычислительно доступным, однако составляет материал отдельного исследования, в том числе потому, что там требуется более тонкие техники обработки не только абсолютных, но и относительных показателей, позволяющих производить компаративный анализ разных случаев.

Границы интервалов во фрагменте знаний и АБС задавались случайным образом с наложенными ограничениями: левая граница интервала не может быть больше правой.

Для каждого вида структуры вычисления проводились 50 раз. При этом и ФЗ, и АБС создавались заново (во избежание влияния того, что ФЗ или АБС уже были приведены к непротиворечивому виду). Измерялось время, необходимое только для поддержания непротиворечивости (измерения времени, необходимого для построения ФЗ или АБС не производилось). После получения измерений были взяты их

математическое ожидание, дисперсия и среднее квадратичное отклонение. Данные величины характеризуют полученные измерения следующим образом, исходя из предположения о том, что распределение измерений близко к нормальному: около 95 % измерений попадают в интервал, нижняя граница которого равна разнице математического ожидания и удвоенного среднее квадратичного отклонения, а верхняя – их сумме.

Далее по полученным измерениям строился график, и приводились выводы, полученные на основании эксперимента.

Эксперимент 1

Использовались АБС, имеющие цепную структуру. Примеры ФЗ и АБС, построенных над пятью атомами, приведены на рис. 3–6.

Численные значения измерений сведены в табл. 1 ($N_{\text{АБС}}$ – общее число атомов в АБС, $N_{\text{ФЗ}}$ – число атомов во фрагментах знаний, $N_{\text{пер}}$ – число атомов на пересечениях). Кроме него, были вычислены значения дисперсии и среднее квадратичное отклонение. Последнее для всех экспериментов не превышало 10 % от среднего, следовательно большая часть измерений находилась в интервале от 80 % до 120 % среднего значения. Все данные представлены в миллисекундах.

$N_{\text{АБС}}$	$N_{\text{АБС}} = N_{\text{ФЗ}}$	$N_{\text{ФЗ}} = 3, N_{\text{пер}} = 2$	$N_{\text{ФЗ}} = 3, N_{\text{пер}} = 1$	$N_{\text{ФЗ}} = 2, N_{\text{пер}} = 1$
3	93	73	86	47
4	187	117	–	71
5	420	167	170	107
6	1144	225	–	123
7	5078	292	225	166
8	38317	346	–	203

Таблица 1. Эксперимент 1. Время работы алгоритма поддержания интернальной непротиворечивости для АБС с цепной структурой, мс

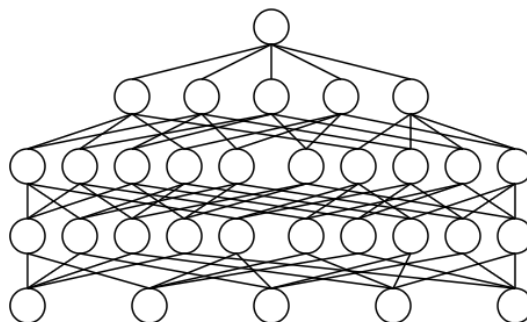


Рис. 3. Фрагмент знаний над пятью атомами

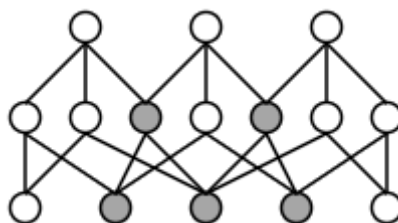


Рис. 4. АБС над тремя фрагментами знаний, каждый над тремя атомами, два атома в пересечении

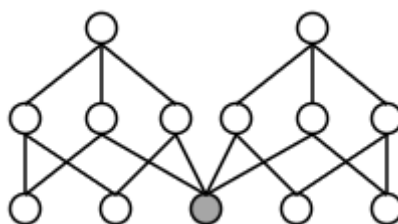


Рис. 5. АБС над двумя фрагментами знаний, каждый над тремя атомами, один атом в пересечении

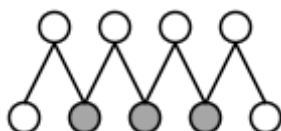


Рис. 6. АБС над 4 фрагментами знаний, каждый над двумя атомами, один атом в пересечении

Для АБС, состоящей из фрагмента знаний с тремя атомами и имеющей по одному атому на пересечениях, значения вычислены только при таком числе атомов в АБС, при котором ее можно полностью «заполнить» фрагментами знаний.

Данные эксперимента приведены на рис. 7, 8. Поскольку различия между поддержанием противоречивости ФЗ и АБС не позволяют увидеть различия между поддержанием интернальной непротиворечивости различных АБС, их результаты представлены отдельно.

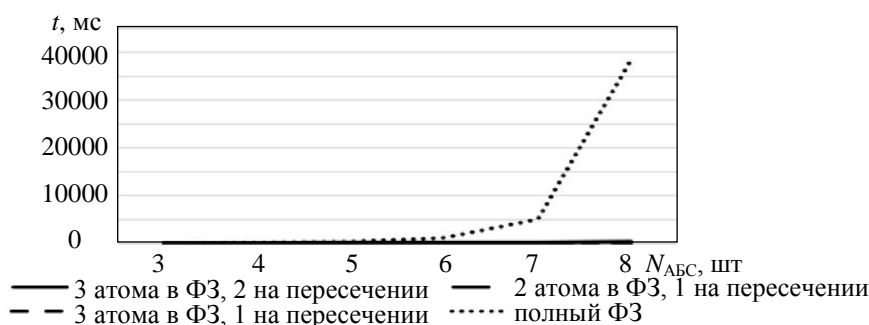


Рис. 7. Эксперимент 1. С поддержанием непротиворечивости фрагмента знаний

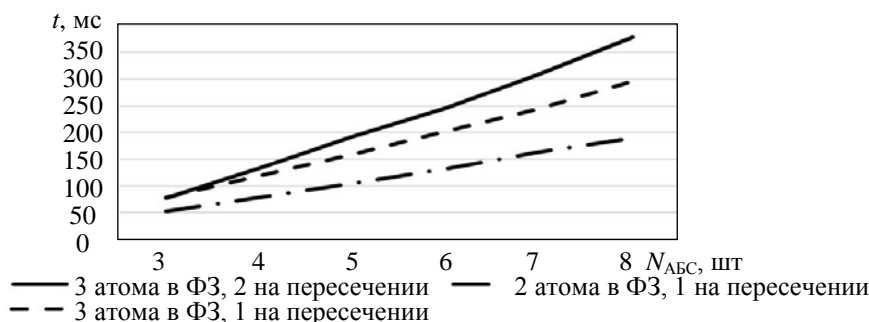


Рис. 8. Эксперимент 1. Без поддержания непротиворечивости фрагмента знаний

Вывод. На основании представленных результатов можно сделать вывод о том, что алгоритм поддержания интернальной непротиворечивости для АБС с цепной структурой по времени работы линейно зависит от количества атомов, над которыми построена АБС. Также графики наглядно показывают, насколько использование интернальной непротиворечивости оптимизирует время поддержания глобальной непротиворечивости.

Эксперимент 2

Эксперимент 2 аналогичен первому, с той разницей, что в нем исследовалось время поддержания непротиворечивости АБС со звездчатой структурой. Примеры таких АБС представлены на рис. 9.

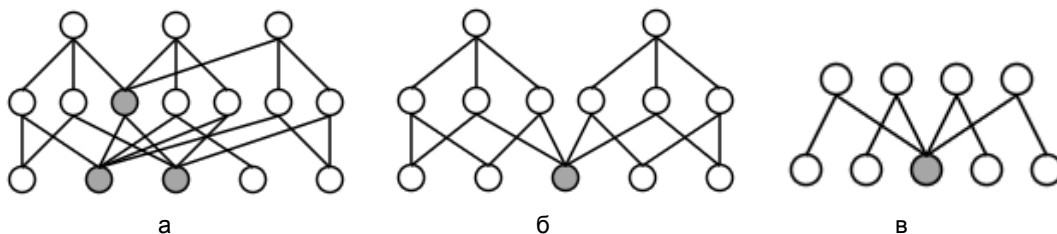


Рис. 9. АБС: над пятью атомами с тремя атомами во фрагменте знаний и двумя на пересечении (а); с тремя атомами во фрагменте знаний и с одним на пересечении (б); с двумя атомами во фрагменте знаний и одним на пересечении (в)

Были получены: математическое ожидание, дисперсия, среднеквадратичное отклонение. Последнее, как и в эксперименте 1, не превысило 10 % от среднего.

Представленные в табл. 2 и на рис. 10, 11 аналогичны представленным в эксперименте 1.

$N_{\text{АБС}}$	$N_{\text{АБС}} = N_{\text{ФЗ}}$	$N_{\text{ФЗ}} = 3, N_{\text{пер}} = 2$	$N_{\text{ФЗ}} = 3, N_{\text{пер}} = 1$	$N_{\text{ФЗ}} = 2, N_{\text{пер}} = 1$
3	93	77	79	53
4	187	133	—	78
5	420	193	159	104
6	1144	246	—	132
7	5078	309	245	163
8	38317	378	—	189

Таблица 2. Эксперимент 2. Время работы алгоритма поддержания интернальной непротиворечивости для АБС со звездчатой структурой, мс

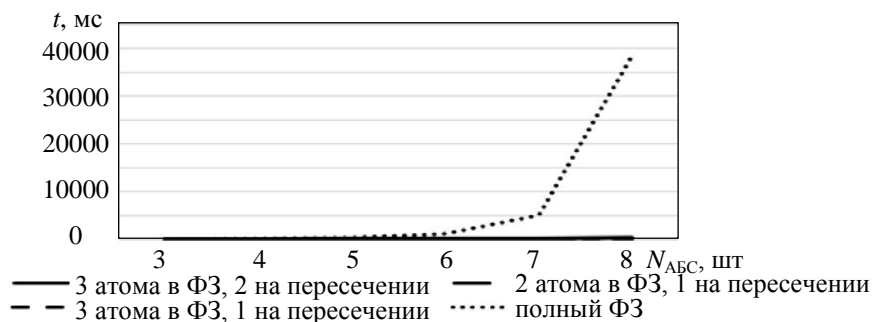


Рис. 10. Эксперимент 2. С поддержанием непротиворечивости фрагмента знаний

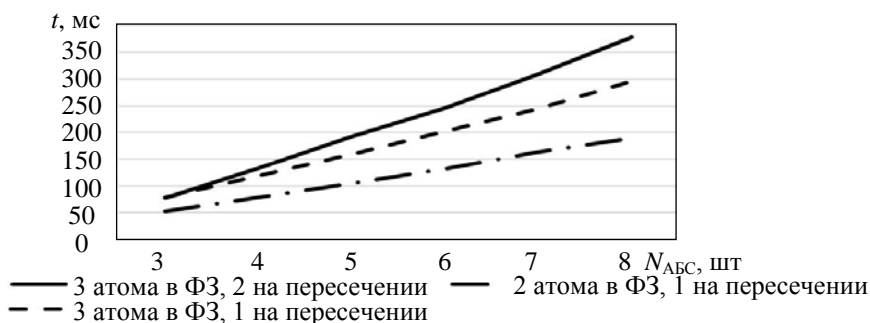


Рис. 11. Эксперимент 2. Без поддержания непротиворечивости фрагмента знаний

Вывод. На основании представленных результатов можно сделать вывод о том, что алгоритм поддержания интернальной непротиворечивости для АБС со звездчатой структурой по времени работы также линейно зависит от количества атомов, над которыми построена АБС. Как и в эксперименте 1, использование интернальной непротиворечивости оптимизирует время поддержания глобальной непротиворечивости.

В целом результаты экспериментов схожи, что объясняется схожестью АБС с цепной и звездчатой структурой.

Заключение

В работе представлены результаты исследований, направленных на поддержание непротиворечивости алгебраической байесовской сети за счет сокращения числа переменных и избыточных условий согласования в задачах линейного программирования, решаемых в алгоритме поддержания интернальной непротиворечивости. Сформулированы алгоритм и теоремы, описывающие сложность его работы. Также проведен эксперимент, демонстрирующий преимущество использования интернальной непротиворечивости по отношению к глобальной, а также различия в его времени работы при различных видах алгебраических байесовских сетей.

Как показало проведенное исследование, время работы алгоритма поддержания интернальной непротиворечивости линейно зависит от числа атомов в алгебраической байесовской сети. Продемонстрирована его неоспоримая эффективность в сравнении с поддержанием непротиворечивости отдельно взятого фрагмента знаний. Кроме того, показано, что при фиксированном числе атомов в алгебраической байесовской сети время работы алгоритма возрастает при увеличении числа атомов как во фрагменте знаний, так и на пересечениях.

Цель машинного обучения – не только сформировать алгебраические байесовские сети, наиболее полно сохраняющие информацию, содержащуюся в выборке (либо в наименьшей степени противоречивую, несовпадающую), но и такую сеть, дальнейшая обработка которой требует достижимых временных затрат, а лучше – наименьших. Результаты исследования полностью соответствуют этой цели. На основании полученных данных о времени поддержания интернальной непротиворечивости и времени поддержания непротиворечивости фрагмента знаний над тем же числом атомов можно выдвинуть требования к алгебраической байесовской сети, позволяющие уменьшить вычислительную сложность: сеть, построенная в результате машинного обучения, должна быть ациклической и максимально разреженной, состоящей из определенным образом скомбинированных линейных и звездчатых структур. Кроме того, результаты работы планируется использовать при интеграции алгебраических байесовских сетей в модели оценки защищенности пользователей информационных систем от социоинженерных атак [21, 22].

Литература

References

1. Тулупьев А.Л., Николенко С.И., Сироткин А.В. Байесовские сети: логико-вероятностный подход. СПб: Наука, 2006. 607 с.
2. Тулупьев А.Л., Сироткин А.В., Николенко С.И. Байесовские сети доверия: логико-вероятностный вывод в ациклических направленных графах. СПб: СПбГУ, 2009. 400 с.
3. Тулупьев А.Л. Алгебраические байесовские сети: локальный логико-вероятностный вывод: Учеб. пособие. СПб: СПбГУ-Анатолия, 2007. 80 с.
4. Augustin T., Seising R. Weichselberger's contribution to imprecise probabilities and statistical inference // *International Journal of Approximate Reasoning*. 2018. V. 98. P. 132–145. doi: 10.1016/j.ijar.2018.04.009
5. Quost B., Destercke S. Classification by pairwise coupling of imprecise probabilities // *Pattern Recognition*. 2018. V. 77. P. 412–425. doi: 10.1016/j.patcog.2017.10.019
6. Abellan J., Mantas C.J., Castellano J.G., Moral-Garcia S. Increasing diversity in random forest learning algorithm via imprecise probabilities // *Expert Systems with Applications*. 2018. V. 97. P. 228–243. doi: 10.1016/j.eswa.2017.12.029
7. Zhang J., Shields M.D. On the quantification and efficient propagation of imprecise probabilities resulting from small datasets // *Mechanical Systems and Signal Processing*. 2018. V. 98. P. 465–483. doi: 10.1016/j.ymsp.2017.04.042
8. Romanov A.V., Levenets D.G., Zolotin A.A., Tulup'ev A.L. Incremental synthesis of the tertiary structure of algebraic Bayesian networks // *Proc. 19th Int. Conf. on Soft Computing and Measurements, SCM 2016*. St. Petersburg, Russia, 2016. P. 28–30. doi: 10.1109/SCM.2016.7519673
9. Kharitonov N.A., Tulup'ev A.L., Zolotin A.A. Software implementation of reconciliation algorithms in algebraic Bayesian networks // *Proc. 20th Int. Conf. on Soft Computing and Measurements, SCM 2017*. St. Petersburg, Russia, 2017. P. 8–10.
10. Zolotin A.A., Tulup'ev A.L. Sensitivity statistical estimates for local a posteriori inference matrix-vector equations in algebraic Bayesian networks over quantum propositions // *Vestnik St. Petersburg University: Mathematics*. 2018. V. 51. N 1. P. 42–48. doi: 10.3103/s1063454118010168
11. Kang H.G., Lee S.H. et al. Development of a Bayesian belief network model for software reliability quantification of digital protection systems in nuclear power plants // *Annals of Nuclear Energy*. 2018. V. 120. P. 62–73. doi: 10.1016/j.anucene.2018.04.045
12. Dal F.N., Quinn C., Morari F. A Bayesian belief network framework to predict SOC dynamics of alternative management scenarios // *Soil and Tillage Research*. 2018. V. 179. P. 114–124. doi: 10.1016/j.still.2018.01.002
13. Liu H., Kim J., Shlizerman E. Functional connectomics from neural dynamics: probabilistic graphical models for neuronal network of *Caenorhabditis elegans* // *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2018. V. 373. N 1758. doi: 10.1098/rstb.2017.0377
14. Kang Z., Yang J. A probabilistic graphical model for the classification of mobile LiDAR point clouds // *ISPRS Journal of Photogrammetry and Remote Sensing*. 2018. V. 143. P. 108–123. doi: 10.1016/j.isprsjprs.2018.04.018
15. Tulup'ev A.L. Probabilistic estimates consistency in conjuncts
1. Tulup'ev A.L., Nikolenko S.I., Sirotkin A.V. *Bayesian Networks: Logical-Probabilistic Approach*. St. Petersburg, Nauka Publ., 2006, 607 p. (in Russian)
2. Tulup'ev A.L., Sirotkin A.V., Nikolenko S.I. *Bayesian Belief Networks: Logical-Probabilistic Inference in the Acyclic Directed Graph*. St. Petersburg, SPbSU Publ., 2009, 400 p. (in Russian)
3. Tulup'ev A.L. *Algebraic Bayesian Networks: a Local Logic-Probabilistic Inference*. Tutorial. St. Petersburg, SPbGU Publ., Anatoliya Publ., 2007, 80 p. (in Russian)
4. Augustin T., Seising R. Weichselberger's contribution to imprecise probabilities and statistical inference. *International Journal of Approximate Reasoning*, 2018, vol. 98, pp. 132–145. doi: 10.1016/j.ijar.2018.04.009
5. Quost B., Destercke S. Classification by pairwise coupling of imprecise probabilities. *Pattern Recognition*, 2018, vol. 77, pp. 412–425. doi: 10.1016/j.patcog.2017.10.019
6. Abellan J., Mantas C.J., Castellano J.G., Moral-Garcia S. Increasing diversity in random forest learning algorithm via imprecise probabilities. *Expert Systems with Applications*, 2018, vol. 97, pp. 228–243. doi: 10.1016/j.eswa.2017.12.029
7. Zhang J., Shields M.D. On the quantification and efficient propagation of imprecise probabilities resulting from small datasets. *Mechanical Systems and Signal Processing*, 2018, vol. 98, pp. 465–483. doi: 10.1016/j.ymsp.2017.04.042
8. Romanov A.V., Levenets D.G., Zolotin A.A., Tulup'ev A.L. Incremental synthesis of the tertiary structure of algebraic Bayesian networks. *Proc. 19th Int. Conf. on Soft Computing and Measurements, SCM 2016*. St. Petersburg, Russia, 2016, pp. 28–30. doi: 10.1109/SCM.2016.7519673
9. Kharitonov N.A., Tulup'ev A.L., Zolotin A.A. Software implementation of reconciliation algorithms in algebraic Bayesian networks. *Proc. 20th Int. Conf. on Soft Computing and Measurements, SCM 2017*. St. Petersburg, Russia, 2017, pp. 8–10.
10. Zolotin A.A., Tulup'ev A.L. Sensitivity statistical estimates for local a posteriori inference matrix-vector equations in algebraic Bayesian networks over quantum propositions. *Vestnik St. Petersburg University: Mathematics*, 2018, vol. 51, no. 1, pp. 42–48. doi: 10.3103/s1063454118010168
11. Kang H.G., Lee S.H. et al. Development of a Bayesian belief network model for software reliability quantification of digital protection systems in nuclear power plants. *Annals of Nuclear Energy*, 2018, vol. 120, pp. 62–73. doi: 10.1016/j.anucene.2018.04.045
12. Dal F.N., Quinn C., Morari F. A Bayesian belief network framework to predict SOC dynamics of alternative management scenarios. *Soil and Tillage Research*, 2018, vol. 179, pp. 114–124. doi: 10.1016/j.still.2018.01.002
13. Liu H., Kim J., Shlizerman E. Functional connectomics from neural dynamics: probabilistic graphical models for neuronal network of *Caenorhabditis elegans*. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2018, vol. 373, no. 1758. doi: 10.1098/rstb.2017.0377
14. Kang Z., Yang J. A probabilistic graphical model for the classification of mobile LiDAR point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018, vol. 143, pp. 108–123. doi: 10.1016/j.isprsjprs.2018.04.018
15. Tulup'ev A.L. Probabilistic estimates consistency in conjuncts

15. Тулупьев А.Л. Непротиворечивость оценок вероятностей в алгебраических байесовских сетях // Вестник Санкт-Петербургского университета. Серия 10. Прикладная математика. Информатика. Процессы управления. 2009. № 3. С. 143–150.
16. Тулупьев А.Л. Байесовские сети: логико-вероятностный вывод в циклах. СПб: СПбГУ, 2008. 140 с.
17. Сироткин А.В. Проверка и поддержание непротиворечивости алгебраических байесовских сетей: вычислительная сложность алгоритмов // Труды СПИИРАН. 2010. Т. 15. С. 162–192.
18. Dalkiran E., Ghalami L. On linear programming relaxations for solving polynomial programming problems // *Computers and Operations Research*. 2018. V. 99. P. 67–77. doi: 10.1016/j.cor.2018.06.010
19. Kolev L., Skalna I. Exact solution to a parametric linear programming problem // *Numerical Algorithms*. 2018. V. 78. N 4. P. 1183–1194. doi: 10.1007/s11075-017-0418-6
20. Feng J., Che A. Novel integer linear programming models for the facility layout problem with fixed-size rectangular departments // *Computers and Operations Research*. 2018. V. 95. P. 163–171. doi: 10.1016/j.cor.2018.03.013
21. Абрамов М.В. Автоматизация анализа социальных сетей для оценивания защищённости от социоинженерных атак // Автоматизация процессов управления. 2018. № 1(51). С. 34–40.
22. Азаров А.А., Тулупьева Т.В., Суворова А.В., Тулупьев А.Л., Абрамов М.В., Юсупов Р.М. Социоинженерные атаки. Проблемы анализа. СПб: Наука, 2016. 352 с.
- and disjuncts ideals. *Vestnik of the St. Petersburg University: Seriya 10: Prikladnaya Matematika. Informatika. Protsessy Upravleniya*, 2009, no. 3, pp. 143–150. (in Russian)
16. Tulup'ev A.L. *Bayesian Networks: Logical-Probabilistic Output in Cycles*. St. Petersburg, SPbSU Publ., 2008, 140 p. (in Russian)
17. Sirotkin A.V. Algebraic bayesian networks reconciliation: computational complexity. *Trudy SPIIRAN*, 2010, vol. 15, pp. 162–192.
18. Dalkiran E., Ghalami L. On linear programming relaxations for solving polynomial programming problems. *Computers and Operations Research*, 2018, vol. 99, pp. 67–77. doi: 10.1016/j.cor.2018.06.010
19. Kolev L., Skalna I. Exact solution to a parametric linear programming problem. *Numerical Algorithms*, 2018, vol. 78, no. 4, pp. 1183–1194. doi: 10.1007/s11075-017-0418-6
20. Feng J., Che A. Novel integer linear programming models for the facility layout problem with fixed-size rectangular departments. *Computers and Operations Research*, 2018, vol. 95, pp. 163–171. doi: 10.1016/j.cor.2018.03.013
21. Abramov M.V. Automation of the social networks websites content analysis in the problems of forecasting the protection of the information systems users from social engineering attacks. *Automation of Control Processes*, 2018, no. 1, pp. 34–40. (in Russian)
22. Azarov A.A., Tulup'eva T.V., Suvorova A.V., Tulup'ev A.L., Abramov M.V., Yusupov R.M. *Social Engineering Attacks. Analysis Problems*. St. Petersburg, Nauka Publ., 2016, 352 p. (in Russian)

Автор

Харитонов Никита Алексеевич – младший научный сотрудник, Санкт-Петербургский институт информатики и автоматизации Российской академии наук (СПИИРАН), Санкт-Петербург, 199178, Российская Федерация, Scopus ID: 57195327488, ORCID ID: 0000-0001-7531-941X, nikita.kharitonov95@yandex.ru

Author

Nikita A. Kharitonov – junior scientific researcher, Saint Petersburg Institute for Informatics and Automation RAS (SPIIRAS), Saint Petersburg, 199178, Russian Federation, Scopus ID: 57195327488, ORCID ID: 0000-0001-7531-941X, nikita.kharitonov95@yandex.ru