

НАУЧНО • ТЕХНИЧЕСКАЯ ИНФОРМАЦИЯ

Серия 2. ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ И СИСТЕМЫ
ЕЖЕМЕСЯЧНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ СБОРНИК

Издается с 1961 г.

№ 10

Москва 2017

ОБЩИЙ РАЗДЕЛ

УДК 001.102 : 316.77

Ф. Н. Ильясов

Информация как мера изменений принимающей системы

Излагается общее определение информации. Развивается понимание информации, предложенное Норбертом Винером, который определил ее как «обозначение содержания, полученного из внешнего мира» субъектом. В этом контексте описывается понимание информация как меры изменений принимающей системы, происходящих вследствие восприятия свойств объектов. Описывается соотношение феномена информации с понятиями знака, смысла, сообщения, символа, сигнала, социальной информации.

Ключевые слова: информация, знак, смысл, сообщение, сигнал, социальная информация

ВВЕДЕНИЕ

Существуют различные понимание информации, ее рассматривают как некий феномен, субстанцию, существующую объективно, независимо от сознания воспринимающего субъекта. Информация также истолковывается как всеобщее свойство материи, ее атрибут. В рамках теории отражения информация есть свойство высокоорганизованной материи, феномен

отражения. Обзор различных подходов см., например, в [1–4].

Сложность истолкования слова «информация» обусловлена тем, что оно употребляется в разных смыслах и используется в качестве синонима таких понятий как знак, смысл, сведения, сообщения, содержание, данные, сигналы, знания и т.д. До настоящего времени нет признанной схемы, концепта, отражающих соотношение указанных терминов, как

нет и общепринятого истолкования самого понятия информации. Наряду с другими авторами, Д. Дусет с коллегами пишет о необходимости формирования универсального междисциплинарного понимания информации и единой теории информации, предлагая разрабатывать эти вопросы в рамках науки об информации (*science of information*¹) [5].

Цель настоящей статьи – попытка изложить наиболее общее понимание информации и описать подходы, позволяющие показать возможное соотношение его с понятиями знака, смысла, сообщения, символа, сигнала. В статье изложены те трактовки указанных понятий, которые, как представляется, позволяют уточнить общее, концептуальное истолкование феномена информации и составляющих ее элементов.

ОБЩЕЕ ИСТОЛКОВАНИЕ ИНФОРМАЦИИ

В рамках различных толкований признается – информация это то, что может восприниматься, приниматься, измеряться, фиксироваться, запоминаться – храниться, обрабатываться, передаваться. Содержание («качество») информации может пониматься, интерпретироваться. Количество информации может измеряться.

Обобщая существующие точки зрения можно заключить, что информация – это результаты измерения и восприятия свойств объектов. Результаты восприятия объектов транслируются посредством знаков, смыслов и сигналов.

Системы, являющиеся источником, носителем информации, можно подразделить на два вида: объект-индуктор² и субъект-индуктор. Система, принимающая информацию, может быть определена как субъект-реципиент³. Объекты-индукторы – это объекты наблюдения, с которыми реципиент не обменивается информацией. С субъектом-индуктором реципиент обменивается информацией.

Информация как мера изменений реципиента

Норберт Винер определил информацию как «обозначение содержания, полученного из внешнего мира» субъектом [6, с. 31]. Из приведенного определения Винера следует, что информация – это феномен субъекта, наблюдателя, т. е. она может существовать только в восприятии субъекта.

Свойства объектов (тел, полей, излучений), находящихся вне поля зрения наблюдателя, можно понимать как потенциальную информацию или информационную потенцию объекта.

Различные исследователи сходятся во мнении, что информация может измеряться. С точки зрения метрологии измерению подлежат величины свойств объектов. В физике величины таких свойств объектов как масса, длина, температура и т.д., измеряются специальными приборами, с использованием специ-

альных шкал, в рамках определенных процедур. Информация не является объектом и не является свойством объекта. Информация есть отражение, фиксация результатов измерения свойств объектов.

В самом широком смысле, информация есть зафиксированная, воспринятая субъектом совокупность величин свойств наблюдаемого («воздействующего») объекта и мера изменений этих величин. Соответственно, единицей измерения, фиксирования информации выступают определенные, наблюдаемые величины свойств объекта и мера, величина изменения этих свойств.

Информация в данном контексте – это то, что получается в результате измерения величин свойств наблюдаемого объекта. Например, «масса объекта = 4 кг», «длина объекта = 25 см».

Совокупность свойств объекта, воспринятая субъектом, позволяет идентифицировать наблюдаемый объект, т. е. отнести его к типу аналогичных объектов, либо идентифицировать его как определенный уникальный объект. Например, субъект получил следующие результаты измерений свойств объекта:

- а) форма: прямоугольный параллелепипед;
- б) размеры: 250×120×65 мм;
- б) выдерживает нагрузку 150 кг на 1 см²;
- в) масса (вес) около 4 кг;

г) преимущественная частота отражаемого светового излучения близка к 400×10^{12} Гц («объект имеет кирпичный цвет»).

Анализ результатов этих измерений позволяет субъекту идентифицировать данный объект как «глиняный кирпич марки «150».

Информация есть совокупность изменений, произошедших в субъекте в результате контакта с объектом. Говоря иначе, к информации относится только та часть свойств объекта, которая оставила след в субъекте. «Получить информацию» – значит измениться.

Объем информации может измеряться количеством изменений, произошедших в субъекте-реципиенте. Можно сказать, информация – это внутренняя характеристика субъекта, в этом смысле она «субъективна», так как представляет собой изменения «внутри субъекта». Например, если человек увидел какой-то предмет, или прочитал какую-то книгу, то информацией является та часть свойств предмета или книги, которая сохранилась в памяти человека, говоря шире, оставила след в его нервной системе.

Знак

Из приведенного понимания Н. Винера следует, что информация – это некоторое «обозначение содержания», оно передается посредством знаков. *Знаки* – это различные искусственные объекты или фрагменты поведения (например, улыбка), отражающие свойства определенных объектов. Знак это созданное для передачи информации условное обозначение определенного объекта, его свойства, величины свойства. Знак является специфической, создаваемой субъектом, формой фиксации, регистрации информации. Совокупность знаков одного порядка называется языком.

Наиболее часто встречающиеся знаки – слова (буквы), числа, знаки арифметических действий и логи-

¹ Выражение «*science of information*» предлагается как понятие более широкое, чем «информатика» («*information science*») [5, с. 201].

² *inductor* (лат.) – побудитель, возбудитель, от *induco* – вожу, навожу, побуждаю.

³ *recipiens* (лат.) – получающий.

ческих отношений. Знаки можно классифицировать по тому, как они отражают свойство или процесс, или наличие свойства или величину совокупности свойств. Например, упомянутый уже знак-слово «кирпич», отражает величины таких свойств объекта как форма, масса, соотношение длины, высоты и ширины, цвет и прочность. Знак «революция» отражает момент быстрого и существенного изменения свойств некоторого процесса. Знак «материя» отражает свойство существования неопределенной группы объектов, имеющих неопределенные физические свойства.

Важным моментом в знаке является тот факт, что знак «разрывает» связь между свойствами объекта и их обозначением. Можно классифицировать знаки по степени их абстракции. Знаки могут обозначать:

1) один конкретный, уникальный объект. Сюда можно отнести имена собственные. Например, имя человека, города, университета;

2) хорошо определенную группу, тип объектов. Например, «яблоки»;

3) плохо определенный тип объектов, имеющих свою специфику. Например, «фрукты»;

4) неопределенный тип объектов, имеющих свою специфику. Например «еда»;

5) неопределенный тип объектов, не имеющих своей специфики. Например «вещь», «материя», «предмет»;

6) сложные понятия, имеющие идеологическую нагрузку – *символы*. Например, флаг, герб, религиозный символ;

7) эмоциональное состояние объекта. Знаки-эмоции передаются синкретной информацией, т. е. посредством звука или изображения, например, звуки смеха, выражения радости или страха.

Чем выше уровень обобщения, абстракции, тем больше смыслов, возможностей истолкования имеет знак.

Смысл

С точки зрения теории информации, *смысл* – это содержание знака, совокупности знаков. Смысл, с одной стороны, есть результат перевода субъектом-индуктором своей информации в знаки, с другой стороны – результат истолкования субъектом-реципиентом этих знаков. Смысл – понятие индивидуальное и субъективное. С точки зрения теории измерения можно сказать, что понимание смысла есть результат измерения знаков. Истолкование одной и той же совокупности знаков может в той или иной мере различаться у разных реципиентов. Истолкование знаков зависит от контекста восприятия и состояния системы-реципиента. Потому смысл знака не может быть однозначным и постоянным.

Юлий Анатольевич Шрейдер определил информацию как «степень изменения системы знаний адресата в результате восприятия данного сообщения» [6]. Однако информация меняет не только систему знаний реципиента, но и его эмоциональную сферу, его отношения, установки, может влиять на физиологические параметры реципиента. К тому же часть невербальной информации не может быть в полной мере отнесена к знанию, а свойства наблюдаемого

объекта не всегда можно классифицировать как сообщение.

Сообщение можно интерпретировать как совокупность знаков, которые передают некоторый смысл. Понимание смысла сообщения индуктором и реципиентом может различаться. Информацией в данном случае являются те изменения в реципиенте, которые были детерминированы воспринимаемым знаком, совокупностью знаков (сообщением). При этом не имеет значения, отражает сообщение объективную реальность или нет.

Истолкование знака, совокупности знаков системой распознавания образов реципиента происходит всякий раз заново, когда реципиент с ней сталкивается. При этом у одного и того же реципиента предыдущее истолкование может отличаться от последующего. Истолкование знака-понятия через описание совокупности его специфических свойств в науке называется эмпирической интерпретацией понятия. Формализация, операционализация содержания знака делает его истолкование более однозначным.

Сигнал

Теория передачи сигналов, теория связи также называется теорией информации, что, как представляется, вносит некую путаницу. Слово «информация» в этом случае используется в узком смысле как совокупность некоторых сигналов, как правило, одного порядка. Например, сигналы, передаваемые в сети Интернет, по телефону, по радио, по ТВ и т.д. Проблемой передачи сигналов занимается математическая теория связи. С физической, технической точки зрения, сигнал представляет собой фрагмент звуковой волны либо излучения, т. е. фрагмент движущейся материи. В теории связи совокупность сигналов рассматривается с точки зрения правильности, целостности передачи сигналов от индуктора к реципиенту.

Сигналы подразделяются на дискретные, вербальные и синкретные, невербальные. Дискретные сигналы, как совокупность некоторых элементов одного порядка, могут вводиться только унимодально и подразумевают аналитическое распознавание образов. Они могут вводиться (и обрабатываться) только последовательно, так как в противном случае утрачивается их содержание. Синкретные сигналы могут существовать только в виде некоторого непрерывного процесса, явления (звук, ощущение, восприятие визуального образа). Синкретные сигналы могут восприниматься параллельно (одновременно) и подразумевают синтетическое, симультантное восприятие. Звуки речи имеют двойственную природу – в этом случае посредством синкретного сигнала передаются дискретные [8, 9].

Сигналы можно подразделить на две группы. Сигналы – фрагменты знака и сигналы-знаки. Посредством сигналов – фрагментов знака, знак обозначается некой совокупностью сигналов, кодируется и транслируется. Каждому знаку ставится в соответствие некая совокупность сигналов, код. У отдельного сигнала – фрагмента знака, нет своего смыслового содержания – ибо это всего лишь неопределенный фрагмент смысла. Одной из сравнительно простых систем сигналов явля-

ется азбука Морзе, где используются два сигнала – длинный и короткий (точка и тире).

Сигналы-знаки имеют смысловое содержание и представляют собой знаки с высоким уровнем обобщения. Известно, что высшие животные обладают системой знаков. У бабочек, например, знаком является специальная химическая формула, которая позволяет реципиенту произвести поло-видовую идентификацию индуктора и вступить с ним половую связь. Животные также имеют сигналы-знаки с высокой степенью абстракции, например, «еда» или «опасность».

Информацией в данном случае являются те изменения в реципиенте, которые были детерминированы совокупностью полученных сигналов.

Информация и время

Время – важная характеристика процессов и объектов. В самом общем понимании время – это некоторая характеристика изменений наблюдаемого объекта. Понятия времени не имеют однозначного истолкования, это в значительной мере обусловлено тем фактом, что время понимается различным образом, выделим два понимания.

1. Дискретное – время как эталон, стандарт длительности (в качестве его принята секунда).

2. Синкретное – время как количество движения, количество изменений, происходящих в объекте.

Важно отметить, что «дискретизация времени» производится в одном, эталонном объекте, а соотносятся эталонные длительности с изменениями в другом, наблюдаемом, объекте, скорость изменений в котором не связана со скоростью и равномерностью изменений (частотой ритмов) в эталонном объекте. Если время понимать как характеристику изменений объекта, то длительность не есть время, так как не является прямой характеристикой количества изменений, происходящих в наблюдаемом объекте. Например, если указать, что возраст некоторого животного составляет 11 лет, то, не имея статистики по продолжительности жизни животных данного вида, нельзя сказать, старое это животное или молодое.

Изменения в объекте есть непрерывный процесс, поэтому время есть синкретная величина, отражающая количество непрерывных изменений объекта. «Время объекта – есть относительная величина, отражающая меру перехода объекта от факта начала к факту конца его существования, показывающая, какую часть своего потенциала существования объект израсходовал на момент наблюдения. Говоря иначе, время объекта – это его (относительный) возраст» [9, с. 7]. Время объекта относительная величина. Например, если Солнце на момент наблюдения израсходовало 46% своей энергии, то возраст Солнца составляет 46%.

Для различения указанных двух толкований понятия времени, предлагается эталоны длительности называть длительностью, а количество изменений объекта – временем.

Время существует только как «внутренняя» характеристика объекта. Большей частью, для измерения синкретных, «внутренних» изменений в наблюдаемом объекте, используется «внешний» дискрет-

ный эталон длительности. Подразумевается, что изменения в наблюдаемом объекте происходят равномерно и тесно коррелируют с эталоном длительности. Однако в значительной части объектов изменения происходят неравномерно. Использование эталона длительности обусловлено тем, что собственный возраст объекта часто сложно измерить. Но, как уже отмечалось, измеряемым свойством (характеристикой) объекта, в случае со временем, является количество изменений, произошедших в объекте.

Информацией о такой характеристике объекта как время, является воспринятый (измеренный) наблюдателем возраст объекта, либо количество изменений объекта за наблюдаемый период. Косвенной («статистической») временной характеристикой объекта является количество эталонных длительностей, зафиксированных с момента появления объекта или за период наблюдения за ним.

Социальная информация

К настоящему времени в науке не сложилось общепринятого определения понятия социальной информации, обзор различных точек зрения см., например, в [10].

Исходя из изложенного понимания, социальную информацию можно определить как изменения, произошедшие в субъекте-реципиенте (индивиде или группе) в результате прямого или опосредованного контакта с субъектом-индуктором. Существенным признаком социальной информации является интерактивность, т.е. к социальной информации относится результат обмена знаками. Цель такого обмена – реализация взаимодействия, координация деятельности, управление сообществом, получение желаемых моральных и материальных ресурсов. Истолкование содержания посланной и полученной информации у индуктора и реципиента может существенно различаться. Обмен знаками может происходить в ходе непосредственного очного взаимодействия, взаимодействия через средства связи, и опосредованно – путем обмена знаками, сохраненными на различных носителях.

СПИСОК ЛИТЕРАТУРЫ

1. Лысак И. В. Информация как общенаучное и философское понятие: основные подходы к определению // Философские проблемы информационных технологий и киберпространства. – 2015. – №2 (10). – С. 9-26.
2. Ушакова А. В. Становление и развитие теории информации // Научный вестник МГТУ ГА. – 2015. – № 215. – С. 112-116.
3. Burgin M. Theory of information. Fundamentality, Diversity and Unification. - Singapore: World Scientific Publishing Co. Pte. Ltd., 2010. – 661 p.
4. Чугунов А. В. Социальная информатика. Учебное пособие. – СПб.: НИУ ИТМО, 2012. – 223 с.
5. Doucette D., Bichler R., Hofkirchner W., Raffl C. Toward a new science of information // Data Science Journal. – 2007. – Vol. 6. – P. 198-205.
6. Винер Н. Кибернетика и общество. – М.: Изд-во иностранной литературы, 1958. – 200 с.

7. Шрейдер Ю.А. Об одной модели семантической теории информации // Проблемы кибернетики. Вып. 13. – М.: Физматгиз, 1965. – С. 234-240.
8. Ильясов Ф.Н. Информационная специализация и функциональная асимметрия мозга // Психологический журнал. – 1987. – Том 8, №6. – С. 44-47.
9. Ильясов Ф.Н. Время как использованный потенциал существования. Web-Институт исследований природы времени. – 2015. – URL: <http://www.chronos.msu.ru/ru/rnameindex/item/ilyasov-f> (дата обращения: 13.08.2017).
10. Варганов В.В. Социальная информация: сущность и функции // Известия российского государственного педагогического университета. Серия: История, социология, политология. – 2009. – № 93. – С. 52-61.

Материал поступил в редакцию 13.08.2017

Сведения об авторе

ИЛЪЯСОВ Фархад Назипович – кандидат философских наук, независимый исследователь, Москва
e-mail: fa08@mail.ru

Т.В. Бусыгина

Электронный научный документ в современном информационно-коммуникационном пространстве

Выполнен обзор электронных форматов, в которых существует научный документ (монография, статья и т.д.) в настоящее время. Показаны особенности (преимущества и недостатки) функционирования и распространения научного документа в электронном виде в интернет-пространстве, на компьютерах, планшетах, специализированных приспособлениях для чтения и других гаджетах. Охарактеризованы новые форматы, предлагаемые для представления научной информации в электронной среде – CDF (Computable Document Format); форматы, описания данных по стандарту RDF (Resource Description Framework), содержащие открытые связанные данные (Linked open data) и обеспечивающие их функционирование в пространстве семантического веба.

Ключевые слова: научный документ, научная статья, научная монография, электронный документ, формат электронного документа, PDF, Portable Document Format, CDF, Computable Document Format, Semantic Web, семантический Веб, связанные данные

Электронный научный документ (монография, статья и т.д.) представляет собой переведенный в электронный вид традиционный научный документ, распространяемый преимущественно в издательском формате PDF (Portable Document Format). Разрабатываются форматы будущего – CDF (Computable Document Format); форматы, описания данных (в частности научных) по стандарту RDF (Resource Description Framework), содержащие открытые связанные данные (Linked open data). Эти форматы используют преимущества научного документа в электронном виде: возможность включать в текст издания мультимедиа, производить вычисления и вносить изменения прямо в файл этого текста, осуществлять переход из текста по гиперссылкам либо в пространство Интернета, либо к другим файлам, расположенным на компьютере локально, а также создавать открытые связанные научные данные и формировать пространство e-Science.

Учреждения Государственной системы научно-технической информации (ГСНТИ) России [1], научно-технические библиотеки в частности, давно и успешно формируют электронные ресурсы (библиографические, реферативные, полнотекстовые базы данных и др.) для информационного обеспечения научных исследований, образовательных процессов [2]. С развитием компьютерных и интернет-технологий все чаще ученые и учащиеся образовательных учреждений разных уровней (аспиранты, студенты, школьники) хотят работать с таким ресурсом, где они могут найти не только описание документа, но и его полный текст. При формировании полнотекстовых ресурсов перед сотрудниками библиотек и других информационных центров стоит вопрос о формате

предоставления пользователю полных текстов документов. Информационные работники должны отслеживать тенденции и явления, происходящие в информационно-коммуникационном пространстве. В их поле зрения должны быть и форматы полнотекстового представления электронных научных документов.

Научные документы являются одной из форм коммуникации, через которую в мировом профессиональном научном сообществе принято доносить порции научной информации другим его членам. Согласно определению, научный документ – это материальный объект, содержащий научно-техническую информацию и предназначенный для ее хранения и использования в общественной практике. Научные документы, как любые объекты могут классифицироваться по различным признакам. В зависимости от способа предоставления информации различают документы: 1) текстовые (книги, журналы, отчеты и др.); 2) графические (чертежи, схемы, диаграммы); 3) аудиовизуальные (звукзаписи, кино- и видеофильмы, компакт-диски); 4) машиночитаемые (например, на микрофотоносителях или электронные – на дискетах или CD) и др. Кроме того, документы подразделяются на первичные (документы, содержащие непосредственные результаты научных исследований и разработок; новые научные сведения, факты, идеи;) и вторичные (полученные в результате аналитико-синтетической и логической переработки сведений или данных, содержащихся в первичных документах (аннотации, рефераты, обзоры)). Как первичные, так и вторичные документы, подразделяются на опубликованные и неопубликуемые. Опубликуемые документы могут выходить в свет с раз-

личной степенью периодичности (журналы, сборники статей). Способы классификации научного документа не ограничиваются перечисленными выше [3]. Разработчики альтметрик¹ предлагают рассматривать в качестве научной продукции (понятие «научный документ» заменяется понятием «научный продукт») более широкий спектр объектов (документов/продуктов): презентации на научных мероприятиях, видеолекции, программные разработки и т.д. [4, 5].

С развитием компьютерной техники все виды научных документов могут быть представлены в электронном виде, такой научный документ относится к классу машиночитаемых документов [6, с. 23]. Можно утверждать, что в настоящее время все научные документы изначально создаются в электронном виде, поскольку компьютерная техника прочно вошла в нашу повседневную жизнь, а ученые стали одними из первых активных ее пользователей (собственно говоря, само по себе развитие компьютерной техники является результатом достижений науки). Причем это относится не только к текстовым научным документам: построение графиков, создание рисунков и чертежей, аудио-визуальных документов в наше время осуществляется с использованием компьютерной техники. Издательские организации давно и успешно освоили компьютерную верстку. Электронное издательское дело развивается, и перечисленные виды научных документов могут остаться существовать только в электронном виде, будучи изданы как электронная книга, размещены в научных электронных журналах, в научных репозиториях, электронных библиотеках, полнотекстовых базах данных. Научные документы, существовавшие ранее лишь на бумажных носителях или фото пленке могут быть переведены в цифровой формат [6].

Все виды научных публикаций могут существовать почти во всех форматах, в которых в настоящее время бытует электронная книга [7]:

- простой текст (plain text);
- текстовые с оформлением – HTML, открытый формат электронных книг Electronic Publication (.epub), OPF FlipBook, OpenDocument, SGML, XML, FictionBook (.fb2), TeX, PDF, Microsoft HTMLHelp (.chm), eReader (.lit), PostScript (.ps, .eps), ExeBook, Mobipocket (.prc) и др.;
- графические растровые – TIFF, JPEG, DjVu и т. п.;
- мультимедиа книги – помимо текста и графики содержат файлы в форматах SWF, EXE, аудио и т. п. Ряд программ (FlipBook Maker, Flip PDF позволяют конвертировать электронные книги некоторых форматов (doc, txt, pdf, html) в книги, с эффектом перелистывания, причем перелистывание может сопровождаться аудио эффектом «шуршание страниц»);
- книги в формате java-мидлетов для мобильных устройств, таких как сотовые телефоны с поддерж-

кой java(J2ME), карманных персональных компьютеров (КПК);

- книги в виде приложений под определенную операционную систему, включающие в себя, помимо самого содержания книги, ещё и специализированную программу просмотра. Как правило, содержат средства, препятствующие копированию содержимого книги и его просмотра отдельно от этой программы-оболочки [8].

Существование научного документа в электронном виде локально на компьютере, на электронных носителях, в интернет-пространстве сопряжено как с преимуществами, так и с недостатками.

На сайтах издательств научные документы предоставляются в HTML-формате, где читателю дается возможность воспользоваться целым комплексом сервисов, обеспечить которые стало возможно только в условиях существования электронного документа в рамках интернет-пространства. Например, рисунки и таблицы могут быть представлены в различных форматах (рисунки: графики, карты в виде изображений разного размера – в виде презентации Microsoft PowerPoint, в виде отдельных графических файлов; таблицы – в CSV (Comma-Separated Values), Microsoft Excel и др.). Рисунки, таблицы можно скачать отдельно от текста статьи в удобном для пользователя виде.

Практически всем форматам, в которых существует научный электронный документ, и формату HTML, в частности, присуща функциональная возможность нелинейного считывания информации по заложенным в документе внутритекстовым и внешним гипертекстовым ссылкам без выхода из исходного текста [6, с. 72]. В тексте по внутритекстовым гиперссылкам можно переходить к ссылкам на цитируемую литературу и далее по гиперссылкам на сайты издательств к полным текстам цитируемых статей. В полных текстах статей, представленных в база данных SciFinder (платформа, на которой имеется специализированная химическая база данных Chemical Abstracts), Scopus от химических формул есть гиперссылки на БД химических соединений CAS Registry [9], Reaxys [10], соответственно. Для химиков и биохимиков такая информация представляет большую ценность. В HTML-файлах научных статей по другим специальностям (медики, специалисты в области наук о Земле) в БД Scopus дается возможность воспользоваться другими информационными продуктами издательства Elsevier (PharmaPendium, Pathway Studio, Geofacets).

В наукометрических базах данных Web of Science и Scopus, сайтах издательств, социальных сетях (Mendeley) представление научных документов сопровождается их наукометрическими и альтметрическими характеристиками; предоставляется также возможность автоматического экспорта библиографического описания (с метаданными) в библиографические менеджеры (EndNote, Mendeley, CiteULike и др.) для организации личной электронной библиотеки ученого и возможности формирования списков цитируемой литературы при написании научного электронного документа.

¹ Альтметрия (альтметрика, altmetrics) направление, в рамках которого разрабатываются программные инструменты для оценки отклика на научные публикации и др. виды научной продукции в продуктах web2.0 (социальных сетях, блогах и т.д.).

В связи с развитием цифровой научной документной инфраструктуры, которую составляют их хранилища, возникли агентства (DOI Registry Agency (CrossRef, DataCite)), взявшие на себя функцию сбора информации о цифровых объектах. Информация об объекте (научном электронном документе, в частности) зашифровывается в виде цифрового идентификатора объекта (ЦИО) (digital object identifier (DOI)), представляющего собой уникальную строку букв и цифр, в которой зашифрована информация об издателе и месте локализации документа в виртуальном пространстве. В библиографических базах данных цифровой идентификатор объекта (DOI) в настоящее время становится элементом библиографической записи. Используя этот идентификатор, пользователь может найти документ в Интернете при открытом доступе, либо в определенной базе данных при наличии лицензионного доступа к ней. ЦИО (DOI) может располагаться в самом научном документе (в PDF).

Наиболее широко используемым форматом представления научных документов является Portable Document Format (PDF). PDF с 1 июля 2008 г. – это открытый стандарт ISO 32000 [11].

PDF – межплатформенный формат электронных документов, разработанный фирмой Adobe Systems с использованием ряда возможностей языка PostScript. В первую очередь PDF предназначен для представления полиграфической продукции в электронном виде. Для просмотра существует множество программ, а также официальная бесплатная программа Adobe Reader (MuPDF, Evince, Okular, SumatraPDF, Mozilla Firefox – свободные; Adobe Reader, Foxit Reader, Google Chrome, PDF-XChange Viewer, PDFMaster, STDU Viewer – проприетарные). Значительное количество современного профессионального печатного оборудования имеет аппаратную поддержку формата PDF, что позволяет производить печать документов в этом формате без использования какого-либо программного обеспечения. Традиционным способом создания PDF-документов является виртуальный принтер, т. е. документ формируется в своей специализированной программе – графической программе или текстовом редакторе, САПР и т. д., а затем экспортируется в формат PDF для распространения в электронном виде, передачи в типографию и т. п.

В первое время существования такой формат был крайне непопулярен:

- Программное обеспечение компании Adobe для чтения и создания PDF было платным;
- В PDF отсутствовала поддержка внешних ссылок, что делало его практически бесполезным во всемирной паутине;
- PDF-документы были большего размера по сравнению с обычным текстом, что означало более длительную загрузку на медленных модемах, широко использовавшихся в те времена;
- На слабых машинах отображение PDF-документов осуществлялось с заметными задержками;
- Существовало несколько конкурирующих форматов, таких, как Envoy, Common Ground Digital Paper, Farallon Replica; даже собственный PostScript являлся конкурентом.

После того как Adobe выпустила бесплатную версию Acrobat Reader (позднее переименованную в Adobe Reader) для чтения PDF-документов, популярность этого формата стала возрастать. Формат PDF-файлов несколько раз изменялся и продолжает эволюционировать (появляются возможности: пароли, ссылки, потоки, независимая от устройства цветопередача; интерактивные элементы, обработка событий мыши, мультимедийные типы, Юникод, улучшение представления цвета и графики; цифровые подписи, цветовые пространства ICC и DeviceN, JavaScript; JBIG2, прозрачность, текстовый слой OCR (optical character recognition – оптическое распознавание символов); JPEG 2000, связанное мультимедиа, объектные потоки, перекрестные потоки, слои; внедренное мультимедиа, 3D, XML-формы, AES-шифрование; AES-шифрование 256-битным ключом; XFA 3.0). Существует несколько спецификаций формата, последовательно расширяющих друг друга. Для каждой новой спецификации создаются новые версии программного обеспечения из пакета Adobe Acrobat.

Электронные научные документы (книги и научные статьи, сборники научных статей) в PDF могут быть распечатаны (книга может быть распечатана частично или полностью) и прочтены по бумаге, прочтены с экрана компьютера или специального переносимого устройства, разработанного именно для этих целей – электронная книга как оборудование (девайс). Однако чтение с экрана стационарного компьютера, ноутбука некомфортно, несмотря на то, что программы чтения PDF дают возможность масштабировать документ (изменять расположение текста на экране). Не вполне комфортным является чтение электронного научного документа в PDF и через девайс «электронная книга». Этот формат на электронных устройствах типа «электронная книга» плохо масштабируется. Кроме того, если девайс «электронная книга» поддерживает технологию электронных чернил (e-ink) (электронной бумаги (e-paper)), то текст и графические изображения предстают в градации серого. В разработке уже находится технология многоцветной (полихромной) электронной бумаги, но широкому пользователю она еще малодоступна по цене и имеет целый ряд недостатков (PocketBook Color Lux 801 ~ 25 000 руб.). В частности, цветные дисплеи E-Ink не обладают достаточно насыщенным цветом. В цвете эти электронные документы можно прочитать благодаря устройствам с жидкокристаллическим дисплеем (монитор компьютера, планшета, смартфона).

В формате PDF научная литература наиболее распространена еще и по той причине, что это издательский формат, где книга или статья предстает в готовом для печати виде. В частности, страницы в документе пронумерованы. Это позволяет ссылаться на определенные места в тексте, т. е. цитировать документ. Цитирование в наукометрии рассматривается как акт и способ научной коммуникации [12, с. 4]. Текст в форматах fb2 и др. (но не изображение) и легко форматируется, можно подобрать размер шрифта, удобного для чтения. Но содержимое документа в таком формате нельзя процитировать, поскольку число страниц меняется в зависимости от

выбираемого пользователем кегля. Цитируемый текст или изображение в тексте располагается на разных страницах в зависимости от выбранного кегля.

Для удобства чтения научных документов в некоторых системах стали давать возможность просмотра и скачивания их в разных форматах: PDF (который можно распечатать, процитировать) и в EPUB-формате (PubMed Central), который удобнее масштабировать и читать на устройстве «электронная книга».

На сегодняшний день PDF продолжает оставаться самым распространенным форматом научного документа. Однако в 2011 г. частная компания Wolfram Research (руководитель Стивен Вольфрам) представила новый формат электронных документов, носящий название Computable Document Format (CDF). По мнению разработчиков он является более совершенной альтернативой распространённому формату PDF, его даже называют «убийцей» PDF. Согласно заявлениям компании, предлагаемый ею формат больше подходит для представления таких данных, как инфографика, учебные пособия, журнальные статьи, разного рода доклады и отчёты. В частности, в текст можно включить интерактивные элементы, графики, схемы, диаграммы, содержание которых пользователь может управлять сам. Например, можно перестроить график для других данных. В файл могут быть добавлены видеоролики и другие медиа-файлы, делающие ознакомление с информацией более наглядным и удобным. Для просмотра CDF-файлов разработан специальный плагин, позволяющий загрузить документ в окно браузера. Плагин может быть бесплатно загружен с официального сайта [11]. Для создания файлов CDF предназначена коммерческая компьютерная система Mathematica (Wolfram Mathematica) компании Wolfram Research. Планируется появление онлайн средств для создания CDF-файлов [14, 15].

За другой формат ресурсов и научных документов как разновидности ресурсов ратуют разработчики концепции Семантического Веб (Semantic WEB) [16] и концепции открытых связанных данных [17].

Семантическая паутина (англ. semantic web) – это общедоступная глобальная семантическая сеть, формируемая на базе Всемирной паутины стандартизацией представления информации в виде, пригодном для машинной обработки. Название «Семантическая паутина» было впервые введено сэром Тимом Бернерсом-Ли (Tim Berners-Lee) (изобретателем Всемирной паутины) в сентябре 1998 г. [16].

Т. Бернерс-Ли предложил идею связанных данных на основе четырех принципов, стимулирующих применение базовых принципов Web для доступа к данным:

1. Применение универсальных идентификаторов ресурсов (URI) в качестве имен объектов.
2. Применение HTTP URI для реализации возможности обращения по этим именам.
3. Предоставление полезной информации тому, кто обращается по URI, с помощью стандартов (Resource Description Format (RDF), Protocol and RDF Query Language (SPARQL)).
4. Включение ссылок на другие URI, позволяющих найти дополнительную информацию.

Связанные данные (Linked Data) – это наборы данных, опубликованные в RDF-формате с использованием унифицированного идентификатора ресурсов URI (Uniform Resource Identifier), для идентификации элементов, которые они содержат. Связанные открытые данные (Linked Open Data) – это связанные наборы данных, опубликованные в RDF-формате и пригодные для использования и переиздания пользователями в своих целях, без каких-либо ограничений в виде авторских прав, патентов и других механизмов контроля.

Научный документ, база данных и другие научные ресурсы должны быть организованы таким образом, чтобы они были источником связанных открытых данных. Это предполагает качественно иную организацию и функционирование информации (научной, в частности) на основе современных информационных компьютерных технологий. Открытые связанные научные данные создают пространство e-Science, где преобладающую роль играет научная информация и её обработка [16-18].

Традиционная книга, прежде чем приобрести современную форму, прошла большой путь становления (от записей на глиняных дощечках, бересте, на папирусе, ткани, коже т.д.). Печатная книга к настоящему времени сформировалась: она создается в определенных форматах (размеры), характеризующихся эргономичностью, имеет необходимый набор сведений для исчерпывающего библиографического описания (сведения об авторе(ах), издатель, месте издания, годе издания, систематизации (УДК, ББК), ISBN). Эстетические потребности человека при создании традиционной книги реализуются в оформлении (обложка, иллюстрации).

В профессиональном научном сообществе сформировались традиции оформления на бумажном носителе научной статьи, тезисов, диссертаций и т. п. Что касается научной электронной книги (научной монографии), научной статьи в электронном виде, то человечество ищет формы их существования. Пока электронная книга, статья в электронном виде представляет собой переведенную в электронный вид традиционную книгу или статью. Даже устройство «электронная книга» имитирует традиционную книгу, оно наиболее удобно в использовании при размерах традиционной книги. Однако разработчики предлагают форматы, которые использовали бы преимущества существования научного документа в электронном виде: возможность включать в книгу мультимедиа, производить вычисления и вносить изменения прямо в файл книги (статьи) (CDF), возможность перехода из книги по гиперссылкам, либо в пространство Интернета, либо к другим файлам, расположенным на компьютере локально, возможность создавать открытые связанные научные данные и формировать пространство e-Science. Электронному документу, научному документу в электронном виде, вероятно, предстоит пройти еще долгий путь развития.

СПИСОК ЛИТЕРАТУРЫ

1. Арский Ю.М. и др. Информационная система России: прошлое, настоящее и будущее // Информационные ресурсы России. – 2006. – № 2. – С. 1-5.
2. Бусыгина Т.В. и др. Эволюция библиографической деятельности в процессе создания информационных ресурсов в ГПНТБ СО РАН // Библиосфера. – 2015. – № 4. – С. 49-56.
3. Гореликова Г.А. Основы научных исследований : учебное пособие. – Кемерово, 2003. – 52 с.
4. Мазов Н.А., Гуреев В.Н. Альтернативные подходы к оценке научных результатов // Вестник Российской академии наук. – 2015. – Т. 85, № 2. – Р. 115-122.
5. Бусыгина Т.В. Альтметрия как комплекс новых инструментов для оценки продуктов научной деятельности // Идеи и идеалы. – 2016. – Т. 2, № 2 (28). – С. 79-87.
6. Электронные документы : создание и использование в публичных библиотеках: справочник / науч. ред. проф. Р.С. Гиляревский, проф. Г.Ф. Гордукалова. – СПб. : Профессия, 2007. – 664 с.
7. Бусыгина Т.В. Как прочитать электронную книгу? // Книга в медиaprостранстве. – Новосибирск, 2011. – С. 289-295. – (Труды ГПНТБ СО РАН ; вып. 2).
8. Вуль В.А. Электронные издания: учеб. пособие. – СПб.: БХВ - Петербург, 2003. – 560 с.
9. Круковская Н.В., Ефременкова В.М. 100-летний юбилей Chemical abstracts Service: факты и цифры // Научно-техническая информация. Сер. 1. – 2007. – № 12. – С. 24-29.
10. Reaxys: страница сайта / Elsevier: сайт. – URL: <https://www.elsevier.com/solutions/reaxys> (дата обращения 08.02.2017).
11. Формат PDF стал международным стандартом // Lenta.ru: новостное интернет-издание. М. 1999 г. – URL: <https://lenta.ru/news/2007/12/05/pdf/> (дата обращения: 21.07.2016).
12. Бредихин С.В., Кузнецов А.Ю., Щербакова Н.Г. Анализ цитирования в библиометрии. – Новосибирск: ИВМиМГСОРАН, НЭИКОН, – 2013. – 344 с.
13. Wolfram Research : вебсайт – URL: <http://www.wolfram.com> (дата обращения 08.02.2017).
14. Формат вычисляемых документов // Википедия. Свободная энциклопедия. – URL: <https://ru.wikipedia.org/> (дата обращения 08.02.2017).
15. Формат вычисляемых документов. Документы оживают благодаря возможностям вычислений // Wolfram. Computation meets knowledge – сайт. URL: <https://www.wolfram.com/cdf/faq/> (дата обращения 08.02.2017).
16. Berners-Lee T., Hendler J., Lassila J. The Semantic Web // Scientific American. – 2001. – Vol. 284, № 5. – Р. 35-43.
17. Berners-Lee T. Design Issues: Linked Data // W3C – website. URL: <https://www.w3.org/Design-Issues/LinkedData.html> (access date 08.08.2016).
18. Еркимбаев А.О., Зицерман В.Ю., Кобзев Г.А., Серебряков В.А., Теймуразов К.Б. Технология научных публикаций в среде «открытых связанных данных» // Научно-техническая информация. Сер. 1. – 2013. – №12. – С. 1-11; Erkimbayev A.O., Zitserman V.Yu., Kobzev G.A., Serebrjakov V.A., Teymurazov K.V. Publishing Scientific Data as Linked Open Data // Scientific and Technical Information Processing. – 2015. – Vol. 42, № 4. – Р. 253-263.

Материал поступил в редакцию 21.08.17.

Сведения об авторе

БУСЫГИНА Татьяна Владимировна – кандидат биологических наук, заведующий Отделом научной библиографии ГПНТБ СО РАН, г. Новосибирск.
e-mail: busig@gpntbsib.ru

О ДСМ-рассуждениях, применимых к объединениям подмножеств баз фактов. Часть I

Рассматривается усиление ДСМ-метода автоматизированной поддержки научных исследований посредством введения тернарного отношения причинности такого, что в позитивных фактах могут содержаться негативные причины, а в негативных фактах – позитивные причины. В связи с этим эффектом рассматриваются объединения подмножеств баз фактов, которым соответствуют комбинации пар предикатов сходства, представляющие бинарные и тернарные отношения причинности, соответственно, для формализации правил ДСМ-рассуждений.

Ключевые слова: ДСМ-рассуждение, правила индуктивного вывода, правила вывода по аналогии, стратегии ДСМ-рассуждений, объединения подмножеств баз фактов, непротиворечивость объединения стратегий ДСМ-рассуждений

§1. ВВЕДЕНИЕ

Одним из вариантов ДСМ-метода автоматизированной поддержки научных исследований (ДСМ-метода АПНИ) является обобщенный ДСМ-метод автоматического порождения гипотез¹, использующий тернарное отношение причинности, представленное предикатом $T(V, X, W)$: V есть причина эффекта W при отсутствии тормозов из множества X .

Дело в том, что ДСМ-метод АПНИ содержит в качестве начальной процедуры индукцию, посредством которой порождается отношение тернарной причинности. В [1] Д.С. Милль, характеризуя свои индуктивные каноны (правила индуктивного вывода), полагал, что они порождают утверждения о причинах эффектов, если не имеется **препятствий** (по нашей терминологии – **тормозов**) для этих причин. В [2, 3] была создана интеллектуальная система, реализующая стандартный ДСМ-метод АПНИ [4], для исследования проблемы возможного развития сахарного диабета у больных хроническим панкреатитом. В [5] был сформулирован тест для распознавания применимости, так называемого, обобщенного ДСМ-метода порождающего гипотезы о тернарных отношениях причинности. Этот тест оказался применим к базам фактов, содержащих истории болезней с хроническим панкреатитом [2, 3]. Обнаруженное обстоятельство послужило поводом для создания новой версии обобщенного ДСМ-метода [6] такого, что в

нем **явно** определены предикаты $Ob^+(V, Z)$ и $Ob^-(V, Z)$, тормозящие (+)- и (–)-причины, соответственно. Кроме того, новая версия ДСМ-метода допускает возможность самих гипотез о причинах быть тормозами.

Экспериментальное исследование массивов данных, представленных расширяемыми базами фактов $B\Phi_i, i = 0, 1, 2$ и $B\Phi_0 \subset B\Phi_1 \subset B\Phi_2$, показало, что применение ДСМ-рассуждений в интеллектуальной системе (ИС-ДСМ, [2, 3]) для обнаружения эмпирических закономерностей (ЭЗК), которыми являются условия развития сахарного диабета у больных хроническим панкреатитом, оказалось неадекватным. Было порождено много фактических противоречий (т. е., гипотез с типом оценки $V = 0$ – «фактически противоречиво») и ранее обнаруженные эмпирические закономерности [2, 3] новой версией ДСМ-метода АПНИ [6] не были установлены.

Неадекватность применения указанной версии ДСМ-метода АПНИ имеет объяснение. В определении предикатов сходства $M_{ag}^+(V, X, Y)$ содержится предикат $Ob^+(Z, V)$ – « Z – тормоз (+)-причины V », что означает ложность формулы $J_{(1,n)}(X \Rightarrow_1 Y)$, если $Z \subset X$, где Z – тормоз (+)-причины V , где $J_{(1,s)}(V \Rightarrow_2 Y)$. Таким образом, если $Z \subset X$, то $\neg J_{(1,n)}(X \Rightarrow_1 Y)$ для всех X . В противном случае, необходимым условием истинности $J_{(1,n)}(X \Rightarrow_1 Y)$ является $\neg(Z \subset X)$ – отсутствие тормоза Z (+)-причины V в объекте X . Заметим, что предикаты

¹ Первые версии ДСМ-метода предназначались для автоматического порождения гипотез, а современные версии являются средствами автоматизированной поддержки исследований.

$X \Rightarrow_1 Y$ и $V \Rightarrow_2 Y$ означают, соответственно, что «X обладает множеством свойств Y» и «V есть причина Y».

Исследовательский массив данных, представленный в $B\Phi_p$, где $p = 0, 1, 2, [2, 3]$ обладает интересной особенностью, такой, что каждая $B\Phi_p$ является объединением подмножеств БФ $B\Phi_p = B\Phi_{p,g} \cup B\Phi_{p,b}$, где $B\Phi_{p,g}, B\Phi_{p,b}$ есть подмножества фактов $B\Phi_p$ ($B\Phi_{p,g} \subset B\Phi_p; B\Phi_{p,b} \subset B\Phi_p$) такие, что в $B\Phi_{p,g}$ применим тест на наличие тормозов причин противоположного знака [5], а $B\Phi_{p,b}$ есть подмножество $B\Phi_p$ такое, что этот тест не применим (это означает, что могут быть применимы правила индуктивного вывода с запретом на контрпримеры [7]).

Таким образом,

$$B\Phi_p = (B\Phi_{p,g}^+ \cup B\Phi_{p,b}^+) \cup (B\Phi_{p,g}^- \cup B\Phi_{p,b}^-),$$

$$B\Phi_p^+ = B\Phi_{p,g}^+ \cup B\Phi_{p,b}^+,$$

$$B\Phi_p^- = B\Phi_{p,g}^- \cup B\Phi_{p,b}^-, \text{ где } p = 0, 1, 2^2.$$

$$B\Phi^+ = \{ \langle X, Y \rangle \mid J_{\langle 1,0 \rangle}(X \Rightarrow_1 Y) \},$$

$$B\Phi^- = \{ \langle X, Y \rangle \mid J_{\langle -1,0 \rangle}(X \Rightarrow_1 Y) \}.$$

$B\Phi^+$ и $B\Phi^-$ взаимно-однозначно соответствуют их описания вида $J_{\langle 1,0 \rangle}(C \Rightarrow_1 Q), J_{\langle -1,0 \rangle}(C \Rightarrow_1 Q)$, где C – константа, представляющая объект, а Q – константа, представляющая эффект.

Множество описаний, соответствующих

$$B\Phi = B\Phi^+ \cup B\Phi^-, B\Phi_p^+, B\Phi_p^-, B\Phi_p = B\Phi_p^+ \cup B\Phi_p^-$$

обозначим посредством

$$\Omega, \Omega^+, \Omega^-, \Omega^+(p), \Omega^-(p), \Omega(p).$$

Результаты применения ДСМ-рассуждений к последовательности вложенных БФ, которыми являются $B\Phi_0, B\Phi_1, B\Phi_2$ [2, 3] схематически можно представить на рис. 1–4, в которых X_i и X_j представляют объекты для (+)-факта и (-)-факта, V^- и V^+ – (-)- и (+)-причины, а Z_0^- и Z_0^+ – тормоза (-)-причины и (+)-причины, соответственно. Рис. 1 и рис. 2 изображают (+)- и (-)-примеры, а рис. 3 и рис. 4 изображают $B\Phi^+ = B\Phi_g^+ \cup B\Phi_b^+, B\Phi^- = B\Phi_g^- \cup B\Phi_b^-$.

Замечание 1-1. Объединение подмножеств базы фактов $B\Phi = B\Phi^+ \cup B\Phi^-$ в соответствии с применимостью обобщенного ДСМ-метода посредством теста [5], распознающего применимость $M_{ag}^\sigma(V, \mathcal{X}, Y)$ или $M_{ab,n}^\sigma(V, Y)$, или их неприменимость, где «ab» – имя M^σ -предиката с запретом на контрпримеры, специфицируется заданием ДСМ-стратегии, пред-

ставленной в решетках правил правдоподобного вывода [8, 9]. В связи с этим корректным обозначением баз фактов будет $B\Phi(x, y)$, где (x, y) имя стратегии $Str_{x,y}$ [2, 8, 9]. Это уточнение вызвано тем фактом, что определение предиката сходства обобщенного ДСМ-метода $M_{ag}^\sigma(V, \mathcal{X}, Y)$ содержит представления гипотез о (\pm) -причинах, порожденных стратегиями $Str_{x,y} (\sigma \in \{+, -\})$.

(+) – пример $J_{\langle 1,n \rangle}(X \Rightarrow_1 Y)$

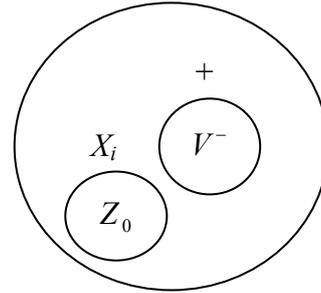


Рис. 1: V^- – (-)-причина; Z_0^- – тормоз V^-

(-) – пример $J_{\langle -1,n \rangle}(X \Rightarrow_1 Y)$

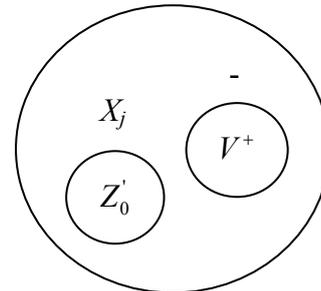


Рис. 2: V^+ – (+)-причина; Z_0^+ – тормоз V^+

$B\Phi_g^+$	g
$B\Phi_b^+$	b

Рис. 3

$B\Phi_g^-$	g
$B\Phi_b^-$	b

Рис. 4

Заметим также, что аксиомы каузальной полноты [9, 10] и тесты распознавания применимости $M_{ag}^\sigma(V, \mathcal{X}, Y)$ должны быть релятивизованы относительно заданных $Str_{x,y}$.

Замечание 2-1. На рис.3 и рис. 4 представлено общее объединение БФ (без рассмотрения пустоты или непустоты подмножеств $B\Phi_g^\sigma$ и $B\Phi_b^\sigma$, где $\sigma \in \{+, -\}$). Ниже перечислим формально возможные случаи, представив их в табл. 1 такой, что непустота $B\Phi_g^\sigma$ и

² В исследовании [2, 3] имеются три расширяемые БФ, где $p=0,1,2$, а Q есть развитие сахарного диабета панкреатогенного генеза.

$B\Phi_b^\sigma$ выражается – посредством 1, а пустота – посредством 0. Очевидно, что **реально** возможны случаи, такие, что, по крайней мере, должно быть непусто одно из подмножеств $B\Phi^+$ (т. е., $B\Phi_g^+$ или $B\Phi_b^+$) и одно из подмножеств $B\Phi^-$ (т. е., $B\Phi_g^-$ или $B\Phi_b^-$).

Таблица 1

	$B\Phi_g^+$	$B\Phi_b^+$	$B\Phi_g^-$	$B\Phi_b^-$
(1)	1	1	1	1
(2)	1	1	1	0
(3)	1	1	0	1
(4)	1	1	0	0
(5)	1	0	1	1
(6)	1	0	1	0
(7)	1	0	0	1
(8)	1	0	0	0
(9)	0	1	1	1
(10)	0	1	1	0
(11)	0	1	0	1
(12)	0	1	0	0
(13)	0	0	1	1
(14)	0	0	1	0
(15)	0	0	0	1
(16)	0	0	0	0

Случаи (6) и (11) выражают тот факт, что

$$B\Phi^+ = B\Phi_g^+, B\Phi^- = B\Phi_g^-,$$

и, соответственно, $B\Phi^+ = B\Phi_b^+, B\Phi^- = B\Phi_b^-$.

Реально же невозможны случаи (4), (8), (12), (13), (14), (15), (16).

§2. ОБОБЩЕННЫЙ ДСМ-МЕТОД АВТОМАТИЗИРОВАННОЙ ПОДДЕРЖКИ НАУЧНЫХ ИССЛЕДОВАНИЙ

В соответствии с обнаруженными характеристиками предметной области, представленной в $B\Phi_0, B\Phi_1, B\Phi_2$, где $B\Phi_0 \subset B\Phi_1 \subset B\Phi_2$ сформулируем новую версию обобщенного ДСМ-метода АПНИ.

В [11] было отмечено существенное разделение (и спецификация) знаний, которыми являются декларативные, процедурные и концептуальные знания.

Посредством декларативных знаний представляются факты из баз фактов и гипотезы из баз знаний интеллектуальных систем (ИС-ДСМ), реализующих ДСМ-метод АПНИ. Концептуальные же знания используются для формулирования принципов ДСМ-метода АПНИ и средств обнаружения эмпирических закономерностей.

JSM-язык JL [2]: X, Z, V (быть может, с нижними индексами) – переменные для объектов и подобъектов;

C, C_1, C_2, \dots – константы (множеств элементов), являющиеся значениями переменных для объектов и подобъектов;

Y, U, W (быть может, с нижними индексами) – переменные для эффектов (множества свойств);

Q, Q_1, Q_2, \dots – константы (множества свойств), являющиеся значениями переменных Y, U, W и т.д.;

n, m, l, k, r, s (быть может, с нижними индексами) – переменные, значениями которых являются натуральные числа ($n \in N, N$ – множество натуральных чисел);

$\mathfrak{X}, \mathfrak{X}_1, \mathfrak{X}_2, \dots$ – переменные для множеств подобъектов;

$\bar{\mathfrak{X}}, \bar{\mathfrak{X}}_1, \bar{\mathfrak{X}}_2, \dots$ – константы, являющиеся значениями переменных \mathfrak{X} ;

$-, \cap, \cup$ – операции алгебры множеств;

$+$ – операция сложения;

$=$ – предикаты равенства для приведенных выше сортов переменных;

\leq, \geq – предикаты для числовых переменных;

\subseteq – предикаты включения для множеств (подобъектов и объектов и множеств свойств);

$X \Rightarrow_1 Y$ – предикат «объект X имеет множество свойств Y »;

$V \Rightarrow_2 W$ – предикат « V есть причина W »;

$T(V, \mathfrak{X}, W)$ – предикат « V есть причина эффекта W при отсутствии тормозов из множества \mathfrak{X} »;

$Z \in \mathfrak{X}$ – предикат принадлежности элемента Z множеству \mathfrak{X} ;

$\neg, \&, \vee, \rightarrow$ – логические связи двузначной логики;

$J_{\bar{\nu}} - J$ - операторы В. Россера – А. Тюркетта [12],

где $\bar{\nu} = \langle \nu, n \rangle, \nu \in \{1, -1, 0\}$, или $\bar{\nu} = (\tau, n)$;

$1, -1, 0, \tau$ – типы истинностных значений $\bar{\nu}$: «фактическая истина (1)», «фактическая ложь» (-1), «фактическое противоречие» (0), «неопределенность» (τ); n – число применений правил правдоподобного вывода (п.п.в.); $n \in N$, а (τ, n) – множество возможных истинностных значений

$$(\tau, n) = \{\langle 1, n+1 \rangle, \langle -1, n+1 \rangle, \langle 0, n+1 \rangle\} \cup (\tau, n+1);$$

\forall, \exists – кванторы всеобщности и существования (соответственно, для приведенных выше трех сортов переменных).

Термы и формулы JSM-языка JL определяются стандартно, но с существенным добавлением формул «переменной длины» с кванторными приставками «переменной длины» и термами «переменной длины».

Это расширение языка логики предикатов 1-го порядка обусловлено тем, что при поиске эмпирических зависимостей в БФ необходимо установить сходство или различие фактов (примеров) на конечном, но заранее **неопределенном** множестве. Число таких примеров k , следовательно, является переменной величиной (k называется параметром эмпирической индукции). Это обстоятельство вынуждает расширить язык логики предикатов 1-го порядка посредством добавления формул и термов «переменной длины», что означает использование кванторов по кортежам (Часть I, Гл. 3 – Д.П. Скворцов «О некоторых способах построения логических языков с кванторами по кортежам», [4, с. 214-232]). В силу такого расширения JL является языком слабой логики предикатов 2-го порядка [13], в котором выразимо транзитивное замыкание.

В [14] было установлено, что предикаты, используемые для формализации правдоподобных рассуждений ДСМ-метода для конечных моделей, выразимы в логике предикатов 1-го порядка. Однако выразительные средства **ЖЛ** удобны для компактной формализации определяемых ниже предикатов для правил правдоподобного вывода.

Подформулами и термами переменной длины в **ЖЛ**, соответственно, являются $J_{\bar{v}}(X \Rightarrow_1 Y_1) \& \dots \& J_{\bar{v}}(X_k \Rightarrow_1 Y_k), (X = X_1 \vee \dots \vee X = X_k)$ и $T_1 \cap \dots \cap T_k = T$, к которым применим квантерный префикс $\exists k \exists X_1 \dots \exists X_k \exists Y_1 \dots \exists Y_k$.

Семантика **ЖЛ** использует исходные непустые множества $U^{(1)}$ и $U^{(2)}$ такие, что $2^{U^{(1)}}$ и $2^{U^{(2)}}$ являются областью определения (domain) переменных X (для объектов и подобъектов) и Y (для множеств свойств); а областью определения переменной \mathfrak{X} является $2^{2^{U^{(1)}}}$.

Для представления объектов (подобъектов) и множеств свойств рассматриваются, соответственно, две булевы алгебры $\mathcal{B}_1 = \langle 2^{U^{(1)}}, \emptyset, U^{(1)}, -, \cap, \cup \rangle$ и $\mathcal{B}_2 = \langle 2^{U^{(2)}}, \emptyset, U^{(2)}, -, \cap, \cup \rangle$.

Истинностные значения высказываний **ЖЛ** подразделяются на **фактические** и **логические** [10]. Фактические истинностные значения $\bar{\nu} = \langle \nu, n \rangle$, где $\nu \in \{1, -1, 0\}$, $n = 0, 1, 2, \dots$, имеют высказывания из баз фактов $J_{\bar{\nu}}(C \Rightarrow_1 Q)$, где $\bar{\nu} = \langle \nu, 0 \rangle$ и гипотезы $J_{\bar{\nu}}(C' \Rightarrow_2 Q)$, $J_{\bar{\nu}}(C \Rightarrow_1 Q)$, где $\bar{\nu} = \langle \nu, n \rangle$, $n > 0$; а также высказывания $J_{(\tau, n)}(C \Rightarrow_1 Q)$ и $J_{(\tau, n)}(C' \Rightarrow_2 Q)$ с типом истинностного значения τ – «неопределенность» [10].

Множество фактических истинностных значений

$$V_{in} = \{ \langle \nu, n \rangle \mid (\nu \in \{1, -1, 0\}) \& \\ \& (n \in N) \} \cup \{ \langle \tau, n \rangle \mid n \in N \}.$$

Множество логических истинностных значений $V_{ex} = \{t, f\}$ двузначной логики (t – истина, f – ложь) используется для определения операторов $J_{\bar{\nu}}\varphi = \begin{cases} t, & \text{если } v[\varphi] = \bar{\nu} \\ f, & \text{если } v[\varphi] \neq \bar{\nu} \end{cases}$, где $v[\varphi]$ – функция оценки высказывания φ .

Предикаты $X \Rightarrow_i Y$ и $T(V, \mathfrak{X}, Y)$, где $i = 1, 2$, определяются, соответственно, отображениями $\Rightarrow_i: 2^{U^{(1)}} \times 2^{U^{(2)}} \rightarrow V_{in}$, $T: 2^{U^{(1)}} \times 2^{2^{U^{(1)}}} \times 2^{U^{(2)}} \rightarrow V_{in}$.

§3. ОПРЕДЕЛЕНИЕ ПРЕДИКАТОВ ОБОБЩЕННОГО ДСМ-МЕТОДА АВТОМАТИЗИРОВАННОЙ ПОДДЕРЖКИ НАУЧНЫХ ИССЛЕДОВАНИЙ

ДСМ-рассуждение является синтезом трех познавательных процедур – индукции, аналогии и абдукции [10, 15].

Индуктивные выводы формализуются посредством М-предикатов сходства, которые определяются посредством экзистенциальных условий (ЭУ), условий сходства (СХ), эмпирических зависимостей, представляющих отношение каузальности (ЭЗ) и нижней границей числа сходных (\pm)-примеров k , где $k \geq 2$.

В [6] была предложена новая версия обобщенного ДСМ-метода АПНИ с исходными предикатами сходства $M_{ag}^\sigma(V, \mathfrak{X}, Y)$, где $\sigma \in \{+, -\}$, такими, что $M_{ag}^+(V, \mathfrak{X}, Y)$ определяются посредством предиката $Ob^+(Z, V)$ – « Z есть тормоз (+)-причины V , вызывающей эффект Y ». Предикат $M_{ag}^+(V, \mathfrak{X}, Y)$ выполним, если не существует тормоза Z такого, что $Z \in \mathfrak{X}$ и $Z \subset X$, а посредством $J_{(1, n)}(X \Rightarrow_1 Y)$ выразимы примеры представляющие эффект Y , вызываемый причиной V , $J_{(1, s)}(V \Rightarrow_2 Y)$ и $V \subset X$. Двойственным образом в [6] определялся $M_{ag}^-(V, \mathfrak{X}, Y)$ и предикат $Ob^-(Z, V)$ для тормозов (-)-причины V .

Формулируемые ниже определения предикатов сходства обобщенного ДСМ-метода используют идею применения предикатов $Ob^\sigma(Z, V)$, выражающих тормоза Z (σ)-причин, где $\sigma \in \{+, -\}$. Однако в $M_{ag}^+(V, \mathfrak{X}, Y)$ используется предикат $Ob^-(Z, Z_0)$, тормозящий (-)-причину Z_0 , содержащуюся в (+)-примерах $J_{(1, 0)}(X \Rightarrow_1 Y)$, где $Z_0 \subset X$ и $Z \subset X$; аналогично в $M_{ag}^-(V, \mathfrak{X}, Y)$ используется предикат $Ob^+(Z, Z_0)$, тормозящий (+)-причину Z_0 , содержащуюся в (-)-примерах $J_{(-1, 0)}(X \Rightarrow_1 Y)$, где $Z_0 \subset X$ и $Z \subset X$ ³.

Сформулируем ниже определение предиката $Ob^-(Z, V_0)$: Z тормоз (-)-причины V_0 . Предварительно определим предикат $V_0 \Rightarrow_2 Y: V_0$ – минимальная причина Y .

$$J_{(-1, m)}(V_0 \Rightarrow_2 Y) \Rightarrow J_{(-1, m)}(V_0 \Rightarrow_2 Y) \& \\ \& \forall l \forall V_1 ((V_1 \subset V_0) \rightarrow \neg J_{(-1, l)}(V_1 \Rightarrow_2 Y))$$
⁴.

$$Ob^-(Z, V_0) \Rightarrow \forall X \forall m \forall m_1 ((V_0 \subset X) \& \\ \& J_{(-1, m)}(V_0 \Rightarrow_2 Y) \& (V \subset X) \& J_{(1, m)}(V \Rightarrow_2 Y) \& \\ \& \neg(Z \subset X)) \rightarrow \neg J_{(1, m_1)}(X \Rightarrow_1 Y)) \& \\ \& \forall X \forall m \forall m_2 ((V_0 \subset X) \& J_{(-1, m)}(V_0 \Rightarrow_2 Y) \&$$

³ Напомним, что в [6] $M_{ag}^+(V, \mathfrak{X}, Y)$ выполним, если тормоза Z из \mathfrak{X} не содержатся в X , где $J_{(1, 0)}(X \Rightarrow_1 Y)$, что выражено в $Ob^+(Z, V)$; аналогично определен $M_{ag}^-(V, \mathfrak{X}, Y)$ и $Ob^-(Z, V)$.

⁴ \Rightarrow – равенство по определению.

$$\begin{aligned} & \&(Z \subset X) \& V \subset X) \rightarrow J_{(1,m_2)}(X \Rightarrow {}_1Y) \& \\ & \& \neg \exists V_1 \exists I(J_{(-1,l)}(V_1 \Rightarrow {}_2Y) \& (V_1 \subset X) \& \\ & \& \neg(V_1 = V_0)). \end{aligned}$$

Заметим, что подформула $\neg \exists V_1 \exists I(J_{(-1,l)}(V_1 \Rightarrow {}_2Y) \& (V_1 \subset X) \& \neg(V_1 = V_0))$ выражает тот факт, что V_0 – единственная отрицательная причина Y такая, что $V_0 \subset X$.

Определение $Ob^-(V_0, Z)$ может быть ослаблено устранением этой подформулы, если в определение $M_{ag}^+(V, \mathfrak{X}, Y)$ включить эмпирическую зависимость $(\exists Z)_3^+$, формулируемую ниже:

$$\begin{aligned} & (\exists Z)_3^+ \forall V_2 \forall X \forall m_0 \forall n_0 \exists Z((J_{(-1,m_0)}(V_2 \Rightarrow {}_2Y) \& \\ & \& (V_2 \subset X) \& J_{(1,n_0)}(X \Rightarrow {}_1Y)) \rightarrow (Ob^-(V_2, Z) \& \\ & \& (\bigvee_{i=1}^{k_1} (X = X_i))). \end{aligned}$$

Очевидно, что при $(\exists Z)_3^+$ в $Ob^-(V_2, Z)$ формула, выражающая единственность отрицательной причины V_0 , рассмотренная выше, не входит.

Замечание 1-3. Расширим ДСМ-язык JL термом $\{Z \mid Ob^-(Z, V_0)\}$ – множество таких Z , что они являются тормозами отрицательной причины V_0 . Соответственно, вводятся равенства $\mathfrak{X} = \{Z \mid Ob^-(Z, V_0)\}$ и $\mathfrak{X} = \emptyset$ для переменной \mathfrak{X} .

Аналогично ниже определяется терм $\{Z \mid Ob^+(Z, Z_0)\}$ для тормозов Z (+)-причин Z_0 .

Определим теперь экзистенциальные условия $(\exists Y)_1^+$ и $(\exists Y)_2^+$ предиката позитивного сходства $M_{a,g}^+(V, \mathfrak{X}, Y)$ для версии обобщенного ДСМ-метода, отличного от рассмотренного в [6].

$$\begin{aligned} & (\exists Y)_1^+ = J_{(1,n)}(X_1 \Rightarrow {}_1Y) \& \dots \& J_{(1,n)}(X_{k_1} \Rightarrow {}_1Y) \& \\ & \& (V \cup Z_0 \cup Z_1 \subseteq \bigcap_{i=1}^{k_1} X_i) \& J_{(1,s)}(V \Rightarrow {}_2Y) \& \\ & \& J_{(-1,r)}(Z_0 \Rightarrow {}_2Y) \& \\ & \& \neg(Z_0 = V) \& (\mathfrak{X} = \{Z \mid Ob^-(Z, Z_0)\}) \& \\ & \& \neg(\mathfrak{X} = \emptyset) \& (Z_1 \in \mathfrak{X}) \& Ob^-(Z_1, Z_0) \end{aligned}$$

$$\begin{aligned} & (\exists Y)_2^+ = \neg J_{(1,n)}(X_{k+1} \Rightarrow {}_1Y) \& \dots \& \neg J_{(1,n)}(X_{k_2} \Rightarrow {}_1Y) \& \\ & \& J_{(-1,r)}(Z_0 \Rightarrow {}_2Y) \& (V \cup Z_0 \subseteq \bigcap_{j=k+1}^{k_2} X_j) \& \\ & \& \neg \exists Z Ob^-(Z, Z_0) \& \neg(Z_0 = V), \end{aligned}$$

где оператор $J_{(1,n)} \Psi = \bigvee_{i=1}^n J_{(1,i)} \Psi$.

Позитивные эмпирические зависимости $(\exists Z)_1^+$ и $(\exists Z)_2^+$ определим следующим образом:

$$\begin{aligned} & (\exists Z)_1^+ = \forall X \forall Z \forall V_0 \forall q \forall p ((V \subset X) \& (V_0 \subset X) \& \\ & \& Ob^-(Z, V_0) \& J_{(-1,q)}(V_0 \Rightarrow {}_2Y) \& J_{(1,s)}(V \Rightarrow {}_2Y) \& \\ & \& (Z \subset X)) \rightarrow (J_{(1,p)}(X \Rightarrow {}_1Y) \& \bigvee_{i=1}^{k_1} (X = X_i)), \end{aligned}$$

$$\begin{aligned} & (\exists Z)_2^+ = \forall X \forall Z \forall V_0 \forall l \forall q (((V \subset X) \& (V_0 \subset X) \& \\ & \& Ob^-(Z, V_0) \& J_{(-1,q)}(V_0 \Rightarrow {}_2Y) \& J_{(1,s)}(V \Rightarrow {}_2Y) \& \\ & \& \neg(Z \subset X)) \rightarrow (\neg J_{(1,l)}(X \Rightarrow {}_1Y) \& (\bigvee_{j=k_1+1}^{k_2} (X = X_j))). \end{aligned}$$

Посредством $(\exists Y)_1^+, (\exists Y)_2^+, (\exists Z)_1^+, (\exists Z)_2^+, (\exists Z)_3^+$, а также нижних границ параметров k_1, k_2 таких, что $k_1 \geq 2$ и $k_2 \geq 1$ определим (+)-предикат сходства следующим образом:

$$\begin{aligned} & M_{a,g}^+(V, \mathfrak{X}, Y) = \\ & = \exists X_1 \dots \exists X_{k_2} \exists Z_0 \exists Z_1 \exists r \exists k_1 \exists k_2 \exists s \exists n \\ & ((\exists Y)_1^+ \& (\exists Y)_2^+ \& (\exists Z)_1^+ \& (\exists Z)_2^+ \& (\exists Z)_3^+ \& \\ & \& (k_1 \geq 2) \& (k_2 \geq 1)). \end{aligned}$$

Таким образом, предикат сходства $M_{ag}^+(V, \mathfrak{X}, Y)$ обобщенного ДСМ-метода, определенный выше, содержит средство распознавания тормозов (-)-причин V_0 , которым является предикат $Ob^-(Z, V_0)$, где Z включается в носителей эффекта Y .

Тогда как в [6] $M_{a,g}^+(V, \mathfrak{X}, Y)$ определяется посредством $Ob^+(Z, V)$, выполнимость которого запрещает наличие тормозов Z искомой причины Y , где Z не включается в носителей эффекта Y , а в данной статье тормоза определяются посредством $Ob^-(Z, Z_0)$.

Аналогично $M_{ag}^+(V, \mathfrak{X}, Y)$ определим предикат сходства для (-)-примеров.

$$\begin{aligned} & (\exists Y)_1^- = J_{(-1,n)}(X_1 \Rightarrow {}_1Y) \& \dots \& J_{(-1,n)}(X_{k_1} \Rightarrow {}_1Y) \& \\ & \& (V \cup Z_0 \cup Z_1 \subseteq \bigcap_{i=1}^{k_1} X_i) \& J_{(-1,s)}(V \Rightarrow {}_2Y) \& \\ & \& J_{(1,r)}(Z_0 \Rightarrow {}_2Y) \& \neg(Z_0 = V) \& \\ & \& (\mathfrak{X} = \{Z \mid Ob^+(Z_0, Z)\}) \& \neg(\mathfrak{X} = \emptyset) \& \\ & \& (Z_1 \in \mathfrak{X}) \& Ob^-(Z_1, Z_0). \end{aligned}$$

$$\begin{aligned} & (\exists Y)_2^- = \neg J_{(-1,n)}(X_{k+1} \Rightarrow {}_1Y) \& \dots \& \neg J_{(-1,n)}(X_{k_2} \Rightarrow {}_1Y) \& \\ & \& J_{(1,r)}(Z_0 \Rightarrow {}_2Y) \& (V \cup Z_0 \subseteq \bigcap_{j=k+1}^{k_2} X_j) \& \\ & \& \neg \exists Z Ob^+(Z, Z_0) \& \neg(Z_0 = V), \end{aligned}$$

где

$$\begin{aligned} Ob^+(Z, V_0) = & \forall X \forall m \forall m_1 ((V_0 \subset X) \& J_{(1,m)}(V_0 \Rightarrow_2 Y) \\ & \& J_{(-1,m)}(V_0 \Rightarrow_2 Y) \& \\ & (V \subset X) \& \neg(Z \subset X) \rightarrow \neg J_{(-1,m_1)}(X \Rightarrow_1 Y)) \& \\ & \& \forall X \forall m \forall m_2 ((V_0 \subset X) \& J_{(1,m)}(V_0 \Rightarrow_2 Y) \& (Z \subset X) \& \\ & (V \subset X)) \rightarrow J_{(-1,m_2)}(X \Rightarrow_1 Y)) \& , \\ & \& \neg \exists V_1 \exists l (J_{(1,l)}(V_1 \Rightarrow_2 Y) \& (V_1 \subset X) \& \neg(V_1 = V_0)), \end{aligned}$$

где

$$\begin{aligned} J_{(1,m)}(V_0 \Rightarrow_2 Y) = \\ J_{(1,m)}(V_0 \Rightarrow_2 Y) \& \forall V_2 ((V_2 \subset V_0) \rightarrow \neg J_{(1,l)}(V_2 \Rightarrow_2 Y)). \end{aligned}$$

$$\begin{aligned} (\exists Z)_1^- = & \forall X \forall Z \forall V_0 \exists q \exists p ((V \subset X) \& (V_0 \subset X) \& \\ & \& Ob^+(Z, V_0) \& J_{(1,q)}(V_0 \Rightarrow_2 Y) \& (Z \subset X)) \rightarrow \\ & (J_{(1,p)}(X \Rightarrow_1 Y) \& (\bigvee_{i=1}^{k_1} (X = X_i))). \end{aligned}$$

$$\begin{aligned} (\exists Z)_3^- = & \forall V_1 \forall X \exists m_0 \exists n_0 \exists Z ((J_{(1,m_0)}(V_1 \Rightarrow_2 Y) \& (V_1 \subset X) \& \\ & \& J_{(-1,n_0)}(X \Rightarrow_1 Y)) \rightarrow (Ob^+(Z, V_1) \& \\ & (\bigvee_{i=1}^{k_1} (X = X_i))). \end{aligned}$$

Определим (-)-предикат сходства следующим образом:

$$\begin{aligned} M_{ag}^-(V, \mathfrak{X}, Y) = \\ \exists X_1 \dots \exists X_{k_2} \exists Z_0 \exists Z_1 \exists r \exists k_1 \exists k_2 \exists s \exists n ((\exists V)_1^- \\ \& (\exists V)_2^- \& (\exists Z)_1^- \& (\exists Z)_2^- \& (\exists Z)_3^- \& \\ (\mathfrak{K}_1 \geq 2) \& (\mathfrak{K} \geq 1)). \end{aligned}$$

Замечание 2-3. Определения предикатов $M_{ag}^+(V, \mathfrak{X}, Y)$ и $M_{ag}^-(V, \mathfrak{X}, Y)$ используют формулы

$$J_{(1,s)}(V \Rightarrow_2 Y), J_{(-1,r)}(Z_0 \Rightarrow_2 Y) (\text{в } M_{a,g}^+(V, \mathfrak{X}, Y))$$

$$\text{и } J_{(-1,s)}(V \Rightarrow_2 Y), J_{(1,r)}(Z_0 \Rightarrow_2 Y) (\text{в } M_{a,g}^-(V, \mathfrak{X}, Y)),$$

соответственно. Эти формулы, представляющие гипотезы о (\pm)-причинах, порождены применением правил правдоподобного (индуктивного) вывода *н.н.в.*- $1^{(\pm)}$ [8, 9]. А *н.н.в.*- $1^{(\pm)}$ специфицируются дополнительными условиями, условиями, усиливающими исходные предикаты сходства $M_{a,n}^\sigma(V, W)$, где $\sigma \in \{+, -\}$. Таким образом, *н.н.в.*- $1^{(\sigma)}$ $(I)_{x,y}^\sigma$, где $x \in \bar{I}^+, y \in \bar{I}^-$, а \bar{I}^+, \bar{I}^- – множества имен указанных усилений M^σ – предикатов [8], формализуются посредством комбинаций предикатов $M_{x,n}^+(V, W), M_{y,n}^-(V, W)$ и логической связки отрицания « \neg » двузначной логики. Таковыми являются

$$\begin{aligned} M_{x,n}^+(V, W) \& \neg M_{y,n}^-(V, W), \\ \neg M_{x,n}^+(V, W) \& M_{y,n}^-(V, W), \end{aligned}$$

$$\begin{aligned} M_{x,n}^+(V, W) \& M_{y,n}^-(V, W), \\ \neg M_{x,n}^+(V, W) \& \neg M_{y,n}^-(V, W), \end{aligned}$$

соответственно, для гипотез с типами истинностных значений $1, -1, 0, \tau$ [10].

Эти *н.н.в.*- $1^{(\sigma)}$ упорядочены ($\sigma \in \{+, -, 0, \tau\}$) и образуют алгебраические решетки [8, 16].

В [8] были определены решетки интенционалов и экстенционалов M^σ -предикатов ($\sigma \in \{+, -\}$), а для *н.н.в.*- $1^{(\sigma)}$ были определены прямые произведения соответствующих решеток.

Приведем обозначения имен M^σ -предикатов, формализующих индуктивные методы (каноны) Д.С. Милля [1] и их усиления: a, ad_0, ad_2 – имена методов сходства, различия и сходства – различия, соответственно; ab, ad_0b, ad_2b – имена методов сходства с запретом на контрпримеры, различия с запретом на контрпримеры и сходства – различия с запретом на контрпримеры. Добавление σ означает, что формулируется (+)- или (-)-предикат с соответствующим именем: $a^\sigma, (ab)^\sigma$ и т.д., где $\sigma \in \{+, -\}$.

На рис. 5 представлены решетки интенционалов M^σ -предикатов $IntL^\sigma$, а на рис. 6 и рис. 7 их подрешетки $IntL_2^\sigma$ и $IntL_1^\sigma$.

Интенционалы M^σ -предикатов заданы их определениями [8], аналогично определяются и интенционалы для $\neg M_{x,n}^+(V, W)$ и $\neg M_{y,n}^-(V, W)$, которым соответствуют решетки $Int(\neg L^\sigma), Int(\neg L_2^\sigma)$ и $Int(\neg L_1^\sigma)$, дуальные исходным решеткам.

Приведем примеры правил индуктивного вывода – *н.н.в.*- $1^{(\sigma)}$, где

$$\begin{aligned} \sigma \in \{+, -\}: (I)_{x,y}^+ \frac{J_{(\tau,n)}(V \Rightarrow_2 W), M_{x,n}^+(V, W) \& \neg M_{y,n}^-(V, W)}{J_{\langle 1, n+1 \rangle}(V \Rightarrow_2 W)}, \\ (I)_{x,y}^- \frac{J_{(\tau,n)}(V \Rightarrow_2 W), \neg M_{x,n}^+(V, W) \& M_{y,n}^-(V, W)}{J_{\langle -1, n \rangle}(V \Rightarrow_2 W)}. \end{aligned}$$

Аналогично формулируются п.п.в.-1 $(I)_{x,y}^\sigma$, где $\sigma \in \{0, \tau\}$, а пара $\langle x, y \rangle$ – имя комбинаций M^σ -предикатов, образующий посылки *н.н.в.*- $1^{(\sigma)}$, где $\sigma \in \{1, -1, 0, \tau\}$.

Выше было сказано, что правилам индуктивного вывода (*н.н.в.*- $1^{(\sigma)}$) $(I)_{x,y}^\sigma$ соответствуют прямые произведения решеток M^σ -предикатов и их отрицаний $\neg M^\sigma$ [8, 9]: $IntL^+ \times Int(\neg L^-)$ (следствие $J_{\langle 1, n+1 \rangle}(V \Rightarrow_2 W)$), $Int(\neg L^+) \times IntL^-$ (следствие $J_{\langle -1, n+1 \rangle}(V \Rightarrow_2 W)$), $IntL^+ \times IntL^-$ (следствие $J_{\langle 0, n+1 \rangle}(V \Rightarrow_2 W)$), $Int(\neg L^+) \times Int(\neg L^-)$ (следствие $I_{\langle \tau, n+1 \rangle}(V \Rightarrow_2 W)$).

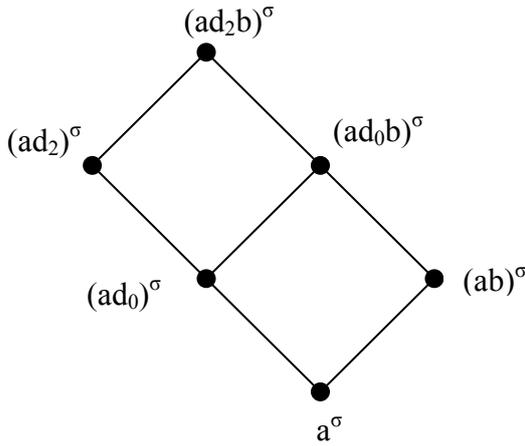


Рис. 5

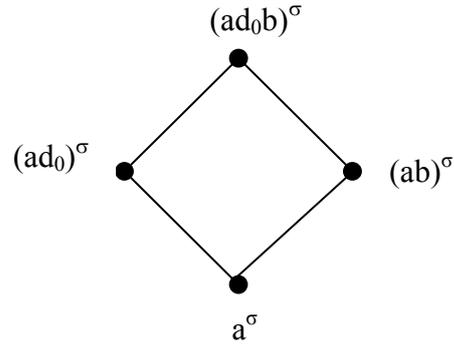


Рис. 6

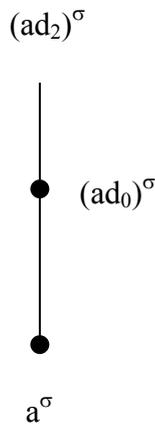


Рис. 7

Аналогично имеют место прямые произведения подрешеток $IntL_i^\sigma$ и $Int(\neg L_i^\sigma)$, где $\sigma \in \{+, -, 0, \tau\}$, а $i = 1, 2$ (рис. 7 и рис. 6).

Рассмотрим случай $B\Phi_b^+ \cup B\Phi_b^-$ (рис. 3 и рис. 4) такой, что $B\Phi_b^+ \cup B\Phi_b^- \subseteq B\Phi$, а $B\Phi_b^\sigma$ - область применения п.п.в.-1 с условием запрета на контрпримеры [8, 10]:

$$(b)^+ \forall X \forall Y (((V \subset X) \& (W \subseteq Y)) \rightarrow (J_{(1,n)}(X \Rightarrow_2 Y) \vee J_{(\tau,n)}(X \Rightarrow_2 Y))),$$

$$(b)^- \forall X \forall Y (((V \subset X) \& (W \subseteq Y)) \rightarrow (J_{(-1,n)}(X \Rightarrow_1 Y) \vee J_{(\tau,n)}(X \Rightarrow_1 Y))).$$

Таким образом, в определениях предикатов $M_{ab,n}^+(V,W)$ и $M_{ab,n}^-(V,W)$ содержатся условия $(b)^+$ и $(b)^-$, что означает выполнимость $M_{ab,n}^+(V,W)$ и $M_{ab,n}^-(V,W)$ на $B\Phi_b^+$ и $B\Phi_b^-$, соответственно. Поэтому $B\Phi_b^+$ и $B\Phi_b^-$ определимы следующим образом:

$$B\Phi_b^+ = \{ \langle X, Y \rangle \mid J_{(1,0)}(X \Rightarrow_1 Y) \& \neg \exists V \exists W \exists m (J_{(-1,m)}(V \Rightarrow_2 W) \& (V \subset X) \& (W \subseteq Y)) \},$$

$$B\Phi_b^- = \{ \langle X, Y \rangle \mid J_{(-1,0)}(X \Rightarrow_1 Y) \& \neg \exists V \exists W \exists m (J_{(1,m)}(V \Rightarrow_2 W) \& (V \subset X) \& (W \subseteq Y)) \}.$$

Замечание 3-3. Существенно отметить, что гипотезы о (\pm) -причинах $J_{(-1,m)}(V \Rightarrow_2 W)$ и $J_{(1,m)}(V \Rightarrow_2 W)$ порождены стратегиями ДСМ-рассуждений $Str_{x,y}$ [9], которым соответствуют решетки интенционалов $IntL_i^\sigma, Int(\neg L_i^\sigma), IntL_i^\sigma, Int(\neg L_i^\sigma)$, где $i = 1, 2$ (рис. 5, 6, 7) и их прямые произведения, соответствующие правилам индуктивного вывода п.п.в.-1 [8, 9].

Таким образом, информативным обозначением предиката $V \Rightarrow_2 W$ будет $V \Rightarrow_{2(x,y)} W$, которое является релятивизацией ДСМ-рассуждений относительно заданного их множества стратегий \overline{Str} [9].

Характеризация гипотез о (\pm) -причинах основана на обратимости п.п.в. (Гл. 5. О дедуктивной имитации некоторых вариантов ДСМ-метода автоматического порождения гипотез [4, с. 240-286]):

$$\forall V \forall W ((J_{(\tau,n)}(V \Rightarrow_2 W) \& M_{x,n}^+(V,W) \& \neg M_{y,n}^-(V,W)) \leftrightarrow J_{(1,n+1)}(V,W)),$$

где « \leftrightarrow » логическая связка эквивалентности (аналогичные утверждения имеют место для трех других комбинаций M^σ - предикатов).

Определим теперь $B\Phi_g^+$ и $B\Phi_g^-$ следующим образом:

$$B\Phi_g^+ = \{ \langle X, Y \rangle \mid J_{(1,0)}(X \Rightarrow_1 Y) \& \exists V \exists W \exists n (J_{(-1,n)}(V \Rightarrow_2 W) \& (V \subset X) \& (W \subseteq Y)) \},$$

$$B\Phi_g^- = \{ \langle X, Y \rangle \mid J_{(-1,0)}(X \Rightarrow_1 Y) \& \exists V \exists W \exists n (J_{(1,n)}(V \Rightarrow_2 W) \& (V \subset X) \& (W \subseteq Y)) \}.$$

Замечание 3-2 следует применить к определениям $B\Phi_g^\sigma$ и $B\Phi_b^\sigma$, где $\sigma \in \{+, -\}$, специфицировав эти определения относительно $Str_{x,y}$ из заданного множества стратегий ДСМ-рассуждений [9]. Таким образом, информативно ввести обозначения $B\Phi_b^\sigma(x, y)$ и $B\Phi_g^\sigma(x, y)$. Следовательно, имеются два семейства

$$B\Phi^\sigma(\sigma \in \{+, -\}) : F_g^\sigma = \{B\Phi_g^\sigma(x, y) | (x \in \bar{I}^+) \& (y \in \bar{I}^-)\},$$

где \bar{I}^σ – множество имен M^σ -предикатов (\bar{I}^σ – носитель (carrier) решеток $IntL^\sigma = \langle \bar{I}^\sigma, \circ, \wedge \rangle$) [8].

Аналогично определим семейства

$$B\Phi^\sigma(\sigma \in \{+, -\}) F_b^\sigma = \{B\Phi_b^\sigma(x, y) | (x \in \bar{I}^+) \& (y \in \bar{I}^-)\}.$$

Введенная спецификация баз фактов, использующая предикат $V \Rightarrow_{2(x,y)} W$ и множество стратегий \overline{Str} [9], где $Str_{x,y} \in \overline{Str}$, порождает семейство объединений массивов данных (он представлен в БФ) $B\Phi(x, y) = B\Phi^+(x, y) \cup B\Phi^-(x, y)$.

Применение же теста [5], распознающего применимость предикатов $M_{ag}^\sigma(V, \mathfrak{X}, Y)$, где $\sigma \in \{+, -\}$, порождает уточненное объединение подмножеств $B\Phi(x, y)$, где $B\Phi^+(x, y) = B\Phi_b^+(x, y) \cup B\Phi_g^+(x, y)$,

$$B\Phi_b^+(x, y) \cap B\Phi_g^+(x, y) = \Lambda$$

$$\text{и } B\Phi^-(x, y) = B\Phi_b^-(x, y) \cup B\Phi_g^-(x, y)$$

$$\text{и } B\Phi_b^-(x, y) \cap B\Phi_g^-(x, y) = \Lambda.$$

Это объединение имеет место и для описаний Ω , соответствующих

$$B\Phi : \Omega(x, y) = \Omega^+(x, y) \cup \Omega^-(x, y), \Omega(x, y)(p) = \Omega^+(x, y)(p) \cup \Omega^-(x, y)(p),$$

где $p = 0, 1, \dots, s$ параметр, обозначающий номер расширений БФ, где $B\Phi_0 \subset \dots \subset B\Phi_s$ ⁵.

Для множества стратегий \overline{Str} имеется столько экземпляров табл. 1, сколько элементов имеет \overline{Str} , соответствующее заданным прямым произведениям решеток M^σ -предикатов и их отрицаний $IntL^\sigma, Int(-L^\sigma)$, где $\sigma \in \{+, -\}$ [8].

Рассмотренное объединение подмножеств БФ порождается тестами (T_1^+) и (T_1^-) из [5]:

$$(T_1^+) \exists m \exists V \exists X \exists Y (J_{(1,0)}(X \Rightarrow_1 Y) \& J_{(-1,m)}(V \Rightarrow_2 Y) \& (V \subset X)),$$

$$(T_1^-) \exists m \exists V \exists X \exists Y (J_{(-1,0)}(X \Rightarrow_1 Y) \& J_{(1,m)}(V \Rightarrow_2 Y) \& (V \subset X)).$$

Выполнимость (T_1^+) и (T_1^-) в БФ означает существование (+)-фактов и (-)-фактов таких, что в (+)-факте содержится (-)-причина V^- (Рис. 1), а в (-)-факте содержится (+)-причина V^+ (рис. 2). Это означает, что в (+)-факте существует тормоз (-)-причины Z_0 , а в (-)-факте существует тормоз (+)-причины Z_0' . В предикатах $M_{ag}^+(V, \mathfrak{X}, Y)$ и $M_{ag}^-(V, \mathfrak{X}, Y)$ **распознавание** тормозов Z_0 и Z_0' реализуется посредством выполнимости предикатов $Ob^-(Z_0, V)$ и $Ob^+(Z_0', V)$, соответственно.

Если для (+)-примера выполняется (T_1^+) , то существует $J_{(-1,m)}(V \Rightarrow_2 Y)$ и $V \subset X$, следовательно, неверно, что имеет место $(b)^-$ – условие запрета на контрпримеры:

$$\forall X \forall Y ((V \subset X) \& (W \subseteq Y)) \rightarrow \rightarrow (J_{(-1,n)}(X \Rightarrow_1 Y) \vee J_{(\tau,n)}(X \Rightarrow_1 Y))$$

для $M_{ag}^+(V, \mathfrak{X}, Y)$. Аналогично для $M_{ag}^-(V, \mathfrak{X}, Y)$, если выполняется (T_1^-) , то неверно, что имеет место условие $(b)^+$.

Имеет место **Утверждение 1-3**.

$$(1) \forall V \forall \mathfrak{X} \forall Y (M_{ag}^+(V, \mathfrak{X}, Y) \rightarrow \neg M_{ab,n}^-(V, Y)),$$

$$(2) \forall V \forall \mathfrak{X} \forall Y (M_{ag}^-(V, \mathfrak{X}, Y) \rightarrow \neg M_{ab,n}^+(V, Y)).$$

Доказательство. Пусть $\langle C', \bar{\mathfrak{X}}, Q \rangle$, где $C', \bar{\mathfrak{X}}, Q$ – константы, выполняет $M_{ag}^+(V, \mathfrak{X}, Y)$. Тогда из истинности $M_{ag}^+(C', \bar{\mathfrak{X}}, Q)$ следует, что существуют $J_{(-1,m)}(C'' \Rightarrow_2 Q)$ и $J_{(1,n)}(C \Rightarrow_1 Q)$ такие, что $C'' \subset C$.

$$M_{ab,n}^-(V, Y) = (M_{a,n}^-(V, Y) \& \forall X ((V \subset X) \rightarrow \rightarrow (J_{(-1,n)}(X \Rightarrow_1 Y) \vee J_{(\tau,n)}(X \Rightarrow_1 Y)))) \leftrightarrow M_{a,n}^-(V, Y) \& \forall X (\neg(V \subset X) \vee J_{(-1,n)}(X \Rightarrow_1 Y) \vee J_{(\tau,n)}(X \Rightarrow_1 Y)),$$

$$\neg M_{a,n}^-(V, Y) \vee \neg \forall X (\neg(V \subset X) \vee J_{(-1,n)}(X \Rightarrow_1 Y) \vee J_{(\tau,n)}(X \Rightarrow_1 Y)) \vee \vee J_{(\tau,n)}(X \Rightarrow_1 Y) \leftrightarrow \neg M_{a,n}^-(V, Y) \vee \exists X ((V \subset X) \& \neg (J_{(-1,n)}(X \Rightarrow_1 Y) \vee J_{(\tau,n)}(X \Rightarrow_1 Y))) \leftrightarrow \neg M_{a,n}^-(V, Y) \vee$$

$$\exists X ((V \subset X) \& (J_{(1,n)}(X \Rightarrow_1 Y) \vee J_{(0,n)}(X \Rightarrow_1 Y))),$$

так как

$$\forall X \forall Y (J_{(1,n)}(X \Rightarrow_1 Y) \vee J_{(-1,n)}(X \Rightarrow_1 Y) \vee \vee J_{(0,n)}(X \Rightarrow_1 Y) \vee J_{(\tau,n)}(X \Rightarrow_1 Y))$$

есть общезначимая формула («закон исключенного пятого»), то

⁵ В приводимом исследовании [2, 3] $p = 1, 2, 3$.

$$\neg(J_{(-1,n)}(X \Rightarrow_1 Y) \vee J_{(\tau,n)}(X \Rightarrow_1 Y)) \leftrightarrow \\ \leftrightarrow J_{(1,n)}(X \Rightarrow_1 Y) \vee J_{(0,n)}(X \Rightarrow_1 Y).$$

Следовательно,

$$\neg M_{ab,n}^-(V, Y) \leftrightarrow \neg M_{a,n}^-(V, Y) \vee \exists X((V \subset X) \& \\ \& J_{(1,n)}(X \Rightarrow_1 Y)) \vee \exists X((V \subset X) \& J_{(0,n)}(X \Rightarrow_1 Y)).$$

В силу определения $M_{ag}^+(V, \mathfrak{X}, Y)$ истинна конъюнкция $J_{(1,n)}(C \Rightarrow_1 Q) \& (C' \subset C)$, где C' такой, что $J_{(1,s)}(C' \Rightarrow_2 Q)$ истинно, а, следовательно, истинно, $\exists X((C' \subset X) \& J_{(1,n)}(X \Rightarrow_1 Q))$, где Q – значение Y , соответствующее $\langle C', \bar{\mathfrak{X}}, Q \rangle$, выполняющей $M_{ag}^+(V, \mathfrak{X}, Y)$.

Таким образом, истинно $\neg M_{ab,n}^-(C', Q)$, если истинно $M_{ag}^+(C', \bar{\mathfrak{X}}, Q)$ для любых $C', \bar{\mathfrak{X}}, Q$, что доказывает Утверждение 1-3. Аналогично доказывается случай (2) для $M_{ag}^-(V, \mathfrak{X}, Y)$.

Следствие Утверждения 1-3. Пусть $M_{y,n}^-(V, Y)$ предикат такой, что у есть $(ab)^-, (ad_0b)^-$ или $(ad_2b)^-$, т.е. у содержит условие $(b)^-$ запрета на контрпримеры, пусть также аналогичное имеет место и для $M_{y,n}^+(V, Y)$, тогда имеет место

$$(1) \forall V \forall \mathfrak{X} \forall Y (M_{ag}^+(V, \mathfrak{X}, Y) \rightarrow \neg M_{y,n}^-(V, Y)),$$

$$(2) \forall V \forall \mathfrak{X} \forall Y (M_{ag}^-(V, \mathfrak{X}, Y) \rightarrow \neg M_{x,n}^+(V, Y)).$$

(1) получим в силу того, что $M_{y,n}^\sigma(V, Y) = M_{ab,n}^\sigma(V, Y) \& \varphi$, где φ есть $(d_0)^\sigma$ или $(d_2)^\sigma$, а $\sigma \in \{+, -\}$ ⁶.

Следствием Утверждения 1 является тот факт, что в $M_{ag}^+(V, \mathfrak{X}, Y)$ реализуются (\pm) -гипотезы, представимые предикатами $V \Rightarrow_{2(x,y)} Y$ такими, что

$$x \in \{(ab)^+, (ad_0b)^+, (ad_2b)^+\},$$

$$а у \in \{(ab)^-, (ad_0b)^-, (ad_2b)^-\}$$

для соответствующих стратегий ДСМ-рассуждений $Str_{x,y}$ [9]. Указанное ограничение применимо к БФ для $Str_{x,y} : B\Phi(x, y) = B\Phi_b^+(x, y) \cup B\Phi_g^+(x, y) \cup B\Phi_b^-(x, y) \cup B\Phi_g^-(x, y)$, где $Str_{x,y} \in \overline{Str}$ [9].

Сформулированные выше тесты (T_1^+) и (T_1^-) не предполагают, что в каждом (+)-факте и каждом (-)-факте содержатся (+)- и (-)-причины, соответственно. Следовательно, (T_1^+) и (T_1^-) совместимы с аксиомами (\exists^+) :

$$(\exists^+) \exists X \exists Y \exists V \exists n (J_{(1,0)}(X \Rightarrow_1 Y) \&$$

$$J_{(1,n)}(V \Rightarrow_2 Y) \& (V \subset X)),$$

$$(\exists^-) \exists X \exists Y \exists V \exists n (J_{(-1,0)}(X \Rightarrow_1 Y) \& J_{(-1,n)}(V \Rightarrow_2 Y) \& \\ (V \subset X)).$$

Это означает, что в $B\Phi(x, y)$ имеется недостаточное множество фактов для порождения (\pm) -причин, объясняющих **все факты** посредством абдукции [9, 10]. Очевидно, что для фиксированной $Str_{x,y}$ $V \Rightarrow_2 Y$ заменяется на $V \Rightarrow_{2(x,y)} Y$.

Пусть в идеальном случае для $B\Phi(x, y)$ истинны аксиомы каузальной полноты $AKIP_{(x,y)}^{(+)}$ и $AKIP_{(-x,y)}^{(-)}$ [9]:

$$AKIP_{(x,y)}^{(+)} \forall X \forall Y \exists V (J_{(1,0)}(X \Rightarrow_1 Y) \rightarrow \exists n (J_{(1,n)}(V \Rightarrow_2 Y) \& \\ (V \subset X) \& \neg(V = \emptyset)),$$

$$AKIP_{(-x,y)}^{(-)} \forall X \forall Y \exists V (J_{(-1,0)}(X \Rightarrow_1 Y) \rightarrow \exists n (J_{(-1,n)}(V \Rightarrow_2 Y) \& \\ (V \subset X) \& \neg(V = \emptyset))).$$

Тогда будем говорить, что $B\Phi(x, y)$ **каузально полна** относительно стратегии ДСМ-рассуждений $Str_{x,y}$.

В случае каузальной полноты $B\Phi(x, y)$ будем применять усиленные тесты (T_2^+) и (T_2^-) [6]:

$$(T_2^+) \exists m \exists n \exists V \exists Z \exists X \exists Y (J_{(1,0)}(X \Rightarrow_1 Y) \& J_{(-1,m)}(V \Rightarrow_2 Y) \& \\ \& (V \subset X) \& J_{(1,n)}(Z \Rightarrow_2 Y) \& (Z \subset X)),$$

$$(T_2^-) \exists m \exists n \exists V \exists Z \exists X \exists Y (J_{(-1,0)}(X \Rightarrow_1 Y) \& J_{(1,m)}(V \Rightarrow_2 Y) \& \\ \& (V \subset X) \& J_{(-1,n)}(Z \Rightarrow_2 Y) \& (Z \subset X)).$$

Рассмотрим фрагменты $B\Phi^+$ и $B\Phi^-$, имеющие объяснения посредством (+)- и (-)-гипотез о причинах:

$$\tilde{B}\Phi^+ = \{ \langle X, Y \rangle \mid \exists n \exists V (J_{(1,0)}(X \Rightarrow_1 Y) \& \\ \& J_{(1,n)}(V \Rightarrow_2 Y) \& (V \subset X)) \},$$

$$\tilde{B}\Phi^- = \{ \langle X, Y \rangle \mid \exists n \exists V (J_{(-1,0)}(X \Rightarrow_1 Y) \& \\ \& J_{(-1,n)}(V \Rightarrow_2 Y) \& (V \subset X)) \}.$$

Определим функции $\rho^+(p)$ и $\rho^-(p)$, где $p = 0, 1, \dots, s$ для $B\Phi_0 \subset \dots \subset B\Phi_s$:

$$\rho^+(p) = \frac{|\tilde{B}\Phi^+(p)|}{|B\Phi^+(p)|}, \quad \rho^-(p) = \frac{|\tilde{B}\Phi^-(p)|}{|B\Phi^-(p)|},$$

где $||$ – число элементов множества, $B\Phi^+(p) =$

$$\{ \langle X, Y \rangle \mid J_{(1,0)}(X \Rightarrow_1 Y) \}, \text{ а } \rho^+(p) \text{ выражают степе-}$$

нь абдуктивного объяснения элементов (фактов) из $B\Phi^+(p)$ для соответствующего расширения $B\Phi^+(0)$ [17]. Аналогично определяются $B\Phi^-(p)$ и $\rho^-(p)$.

⁶ Другое утверждение о связи $M_{ag}^\sigma(V, \mathfrak{X}, Y)$ и M_{ab}^- предикатом имеется в [6].

Очевидно, что $0 \leq \rho^\sigma(p) \leq 1$; если $\rho^\sigma(p) = 1$, то $AKP^{(\sigma)}$ истинно для $B\Phi(p)$. Если же $0 < \rho^\sigma(p) < 1$, то истинны (\exists^σ) , что означает каузальную неполноту $B\Phi(p)$. Таким образом, проверка тестов (T_i^σ) и вычисление функций $\rho^\sigma(p)$, где $i = 1, 2$, а $\sigma \in \{+, -\}$ устанавливают **типологию** предметной области, представленной в БФ. Очевидно, что эта типология зависит от используемой стратегии ДСМ-рассуждений $Str_{x,y}$ для соответствующих $V \Rightarrow_{2(x,y)} Y$ и аксиом каузальной полноты $AKP_{(x,y)}^{(+)}$ и $AKP_{(-x,y)}^{(-)}$, так как $\rho^\sigma(p)$ $AKP^{(\sigma)}$ и (\exists^σ) формулируются для $B\Phi(x,y)$ и для их расширений $B\Phi(p)$, где $p = 0, 1, \dots, s$, а $B\Phi(0) \subset \dots \subset B\Phi(s)$, а $\Omega(0) \subset \dots \subset \Omega(s)$.

Для $B\Phi_g = B\Phi_g^+ \cup B\Phi_g^-$ формулируются аксиомы $AKP^{(\sigma)}$ и (\exists^σ) для каузальной полноты и неполноты, соответственно [6], если $B\Phi = B\Phi_g$. Таким образом, получаем:

$$AKP_T^{(+)} \forall X \forall Y \exists V \exists \mathfrak{X} (J_{(1,0)}(X \Rightarrow_1 Y) \rightarrow \exists n (J_{(1,n)} T(V, \mathfrak{X}, Y) \& (V \subset X))),$$

$$AKP_T^{(-)} \forall X \forall Y \exists V \exists \mathfrak{X} (J_{(-1,0)}(X \Rightarrow_1 Y) \rightarrow \exists n (J_{(-1,n)} T(V, \mathfrak{X}, Y) \& (V \subset X))),$$

$$(\exists^+)_T \exists X \exists Y \exists V \exists \mathfrak{X} \exists n (J_{(1,0)}(X \Rightarrow_1 Y) \& J_{(1,n)} T(V, \mathfrak{X}, Y) \& (V \subset X)),$$

$$(\exists^-)_T \exists X \exists Y \exists V \exists \mathfrak{X} \exists n (J_{(-1,0)}(X \Rightarrow_1 Y) \& \mathfrak{X} \& J_{(-1,n)} T(V, \mathfrak{X}, Y) \& (V \subset X)).$$

Заметим, что $T(V, \mathfrak{X}, Y)$ определены для $Str_{x,y}$, поэтому информативными будут обозначения $AKP_{T,(x,y)}^{(+)}$, $AKP_{T,(-x,y)}^{(-)}$ и $(\exists^+)_T,(x,y)$, $(\exists^-)_T,(-x,y)$.

В [9] в соответствии с целью ДСМ-метода АПНИ, которой является обнаружение в последовательностях расширяемых баз фактов эмпирических закономерностей, были изменены предикаты для аналогии $\Pi_n^\sigma(X, Y)$, где $\sigma \in \{+, -, 0, \tau\}$. Согласно допущению ДЗ [9] эффект Y имеет в качестве причины один подбъект V такой, что $V \subset X$, где X – объект, имеющий эффект Y , что выразимо предикатом $X \Rightarrow_1 Y$. В общем определении вывода по аналогии в ДСМ-рассуждениях использовалось множество гипотез с причинами V_1, \dots, V_k с эффектами Y_1, \dots, Y_k , соответственно, таких, что $\bigcup_{i=1}^k Y_i = Y$, где Y – предсказываемый эффект объекта X (то есть, $X \Rightarrow_1 Y$).

Определим теперь предикат аналогии, использующий тернарные предикаты каузальности

$$T(V, \mathfrak{X}, Y): \Pi_n^+(X, Y) = \exists V \exists \mathfrak{X} ((J_{(1,n)} T(V, \mathfrak{X}, Y) \& \& (V \subset X) \& \forall U ((U \subseteq Y) \& \neg(U = \emptyset)) \rightarrow \rightarrow \neg \exists V_0 \exists \mathfrak{X}_0 ((J_{(-1,n)} T(V_0, \mathfrak{X}_0, U) \vee \vee J_{(0,n)} T(V_0, \mathfrak{X}_0, U)) \& (V_0 \subset X))).$$

Аналогично определим

$$\Pi_n^-(X, Y) = \exists V \exists \mathfrak{X} ((J_{(-1,n)} T(V, \mathfrak{X}, Y) \& (V \subset X) \& \forall U ((U \subseteq Y) \& \neg(U = \emptyset)) \rightarrow \rightarrow \neg \exists V_0 \exists \mathfrak{X}_0 ((J_{(1,n)} T(V_0, \mathfrak{X}_0, U) \vee J_{(0,n)} T(V_0, \mathfrak{X}_0, U)) \& (V_0 \subset X))).$$

Определим также предикаты $\Pi_n^0(X, Y)$ и $\Pi_n^\tau(X, Y)$ следующим образом:

$$\Pi_n^0(X, Y) = \exists V_1 \exists \mathfrak{X}_1 \exists V_2 \exists \mathfrak{X}_2 (J_{(1,n)} T(V_1, \mathfrak{X}_1, Y) \& \& J_{(-1,n)} T(V_2, \mathfrak{X}_2, Y) \& (V_1 \subset X) \& (V_2 \subset X)) \vee \vee \exists V \exists \mathfrak{X} (J_{(0,n)} T(V, \mathfrak{X}, Y) \& (V \subset X)),$$

$$\Pi_n^\tau(X, Y) = \neg(\Pi_n^+(X, Y) \vee \Pi_n^-(X, Y) \vee \Pi_n^0(X, Y)).$$

В [2, 3] исследуемым эффектом является «развитие сахарного диабета панкреатогенного генеза», а это означает, что значением Y в предикатах $X \Rightarrow_1 Y$, $X \Rightarrow_2 Y$ и $T(V, \mathfrak{X}, Y)$ является множество, состоящее из одного элемента. В силу этого $\Pi_n^+(X, Y)$ и $\Pi_n^-(X, Y)$ упрощаются, а именно:

$$\Pi_n^+(X, Y) = \exists V \exists \mathfrak{X} ((J_{(1,n)} T(V, \mathfrak{X}, Y) \& \& (V \subset X)) \& \neg \exists V_0 \exists \mathfrak{X}_0 ((J_{(-1,n)} T(V_0, \mathfrak{X}_0, Y) \vee \vee J_{(0,n)} T(V_0, \mathfrak{X}_0, Y)) \& (V_0 \subset X))),$$

аналогично определяется $\Pi_n^-(X, Y)$.

Для предикатов $\Pi_n^\sigma(X, Y)$, определяемых посредством предикатов $T(V, \mathfrak{X}, Y)$ введем обозначения $\Pi_{n,g}^\sigma(X, Y)$.

Рассмотрим объединение подмножеств БФ для случая (6) $B\Phi = B\Phi_g^+ \cup B\Phi_g^-$ представленного в табл. 1. Этому разбиению БФ соответствуют правила индуктивного вывода (п.п.в.-1):

$$(I)_{g,1}^+ \frac{J_{(\tau,n)} T(V, \mathfrak{X}_1, Y), M_{ag}^+(V, \mathfrak{X}_1, Y) \& \neg M_{ag}^-(V, \mathfrak{X}_2, Y)}{J_{(1,n+1)} T(V, \mathfrak{X}_1, Y)},$$

$$(I)_{g,1}^- \frac{J_{(\tau,n)} T(V, \mathfrak{X}_1, Y), \neg M_{ag}^+(V, \mathfrak{X}_1, Y) \& M_{ag}^-(V, \mathfrak{X}_2, Y)}{J_{(-1,n+1)} T(V, \mathfrak{X}_2, Y)},$$

$$(I)_{g,1}^0 \frac{J_{(\tau,n)} T(V, \mathfrak{X}_1, Y), M_{ag}^+(V, \mathfrak{X}_1, Y) \& M_{ag}^-(V, \mathfrak{X}_2, Y)}{J_{(0,n+1)} T(V, \mathfrak{X}_1 \cup \mathfrak{X}_2, Y)},$$

$$(I)_{g,1}^\tau \frac{J_{(\tau,n)} T(V, \mathfrak{X}_1, Y), \neg M_{ag}^+(V, \mathfrak{X}_1, Y) \& \neg M_{ag}^-(V, \mathfrak{X}_2, Y)}{J_{(\tau,n+1)} T(V, \mathfrak{X}_1 \cup \mathfrak{X}_2, Y)}.$$

Областью применения $n.n.v.-1_{g,1}$ является $B\Phi = B\Phi_g^+ \cup B\Phi_g^-$, что представлено случаем (6) в табл. 1.

Рассмотрим объединение подмножеств БФ для случая (7). $B\Phi = B\Phi_g^+ \cup B\Phi_b^-$, представленного в табл. 1. Этому разбиению БФ соответствуют правила индуктивного вывода (п.п.в.-1):

$$(I)_{g,2}^+ \frac{J_{(\tau,n)}T(V, \mathfrak{X}, Y), M_{ag}^+(V, \mathfrak{X}, Y) \& \neg M_{y,n}^-(V, Y)}{J_{(1,n+1)}T(V, \mathfrak{X}, Y)},$$

$$(I)_{g,2}^- \frac{J_{(\tau,n)}(V \Rightarrow_2 Y), \neg M_{ag}^+(V, \mathfrak{X}, Y) \& M_{y,n}^-(V, Y)}{J_{(-1,n+1)}(V \Rightarrow_2 Y)},$$

$$(I)_{g,2}^0 \frac{J_{(\tau,n)}T(V, \mathfrak{X}, Y), J_{(\tau,n)}(V \Rightarrow_2 Y), M_{ag}^+(V, \mathfrak{X}, Y) \& M_{y,n}^-(V, Y)}{J_{(0,n+1)}T(V, \mathfrak{X}, Y), J_{(0,n+1)}(V \Rightarrow_2 Y)},$$

$$(I)_{g,2}^\tau \frac{J_{(\tau,n)}T(V, \mathfrak{X}, Y), J_{(\tau,n)}(V \Rightarrow_2 Y), \neg M_{ag}^+(V, \mathfrak{X}, Y) \& \neg M_{y,n}^-(V, Y)}{J_{(\tau,n+1)}T(V, \mathfrak{X}, Y), J_{(\tau,n+1)}(V \Rightarrow_2 Y)}.$$

Областью применения $n.n.v.-1_{g,2}$ является $B\Phi = B\Phi_g^+ \cup B\Phi_b^-$, что представлено в табл. 1 случаем (7).

Рассмотрим объединение подмножеств БФ для случая (10) $B\Phi = B\Phi_b^+ \cup B\Phi_g^-$ представленного в табл. 1. Этому разбиению БФ соответствуют правила индуктивного вывода (п.п.в.-1):

$$(I)_{g,3}^+ \frac{J_{(\tau,n)}(V \Rightarrow_2 Y), M_{x,n}^+(V, Y) \& \neg M_{ag}^-(V, \mathfrak{X}, Y)}{J_{(1,n+1)}(V \Rightarrow_2 Y)},$$

$$(I)_{g,3}^- \frac{J_{(\tau,n)}T(V, \mathfrak{X}, Y), \neg M_{x,n}^+(V, Y) \& M_{ag}^-(V, \mathfrak{X}, Y)}{J_{(-1,n+1)}T(V, \mathfrak{X}, Y)},$$

$$(I)_{g,3}^0 \frac{J_{(\tau,n)}T(V, \mathfrak{X}, Y), J_{(\tau,n)}(V \Rightarrow_2 Y), M_{x,n}^+(V, Y) \& M_{ag}^-(V, \mathfrak{X}, Y)}{J_{(0,n+1)}T(V, \mathfrak{X}, Y), J_{(0,n+1)}(V \Rightarrow_2 Y)},$$

$$(I)_{g,3}^\tau \frac{J_{(\tau,n)}T(V, \mathfrak{X}, Y), J_{(\tau,n)}(V \Rightarrow_2 Y), \neg M_{x,n}^+(V, Y) \& \neg M_{ag}^-(V, \mathfrak{X}, Y)}{J_{(\tau,n+1)}T(V, \mathfrak{X}, Y), J_{(\tau,n+1)}(V \Rightarrow_2 Y)}.$$

Областью применения $n.n.v.-1_{g,3}$ является $B\Phi = B\Phi_b^+ \cup B\Phi_g^-$, что представлено в табл. 1 случаем (10).

Рассмотрим объединение подмножеств БФ для случая (11) $B\Phi = B\Phi_b^+ \cup B\Phi_b^-$, представленного в табл. 1. Этому разбиению БФ соответствуют п.п.в.-1:

$$(I)_{(x,y),4}^+ \frac{J_{(\tau,n)}(V \Rightarrow_2 Y), M_{x,n}^+(V, Y) \& \neg M_{y,n}^-(V, Y)}{J_{(1,n+1)}(V \Rightarrow_2 Y)},$$

$$(I)_{(x,y),4}^- \frac{J_{(\tau,n)}(V \Rightarrow_2 Y), \neg M_{x,n}^+(V, Y) \& M_{y,n}^-(V, Y)}{J_{(-1,n+1)}(V \Rightarrow_2 Y)},$$

$$(I)_{(x,y),4}^0 \frac{J_{(\tau,n)}(V \Rightarrow_2 Y), M_{x,n}^+(V, Y) \& M_{y,n}^-(V, Y)}{J_{(0,n+1)}(V \Rightarrow_2 Y)},$$

$$(I)_{(x,y),4}^\tau \frac{J_{(\tau,n)}(V \Rightarrow_2 Y), \neg M_{x,n}^+(V, Y) \& \neg M_{y,n}^-(V, Y)}{J_{(\tau,n+1)}(V \Rightarrow_2 Y)}.$$

Областью применения $n.n.v.-1_{(x,y),4}$ является $B\Phi = B\Phi_b^+ \cup B\Phi_b^-$, что представлено случаем (11) в табл. 1.

Заметим, что x и y могут не иметь условий $(b)^+$ и $(b)^-$, соответственно, а $(I)_{x,y}^\sigma$ соответствуют решеткам, элементами которых являются $\langle \neg x, y \rangle$ (для $\sigma = +$), $\langle x, \neg y \rangle$ (для $\sigma = -$), $\langle x, y \rangle$ (для $\sigma = 0$), $\langle \neg x, \neg y \rangle$ (для $\sigma = \tau$).

В соответствии с типами правил индуктивного вывода $n.n.v.\bar{r}^{(\nu)}$, где $\nu \in \{1, -1, 0, \tau\}$, а $j = 1, 2, 3$, определим правила вывода по аналогии $n.n.v.-2_{\bar{r}}^{(\nu)}$ для предикатов $M_{ag}^\sigma(V, \mathfrak{X}, Y)$, где $\sigma \in \{+, -\}$, $\bar{r} = \langle \lambda, j, \langle x, y \rangle \rangle$, а $\lambda = g, j = 1, 2, 3$. Случаи для $\text{Str}_{x,y}$ с $a, ab, ad_0, ad_0b, ad_2, ad_2b$ и $j = 4$ ранее были рассмотрены при определении предикатов $\Pi_n^\sigma(Z, W)$, где $\sigma \in \{+, -\}$ [8, 9].

1. $\Pi_{n,(g,1)}^+(Z, W) = \exists V \exists \mathfrak{X} ((J_{(1,n)}T(V, \mathfrak{X}, W) \& (V \subset Z)) \& \neg \exists V_0 \exists \mathfrak{X}_0 ((J_{(-1,n)}T(V_0, \mathfrak{X}_0, W) \vee J_{(0,n)}T(V_0, \mathfrak{X}_0, W)) \& (V_0 \subset Z)))$,

$\Pi_{n,(g,1)}^+(Z, W)$ и $\Pi_{n,(g,1)}^\sigma(Z, W)$, определяемые ниже, где $\sigma \in \{-, 0, \tau\}$, определены для $B\Phi = B\Phi_g^+ \cup B\Phi_g^-$;

$\Pi_{n,(g,1)}^-(Z, W) = \exists V \exists \mathfrak{X} ((J_{(-1,n)}T(V, \mathfrak{X}, W) \& (V \subset Z)) \& \neg \exists V_0 \exists \mathfrak{X}_0 ((J_{(1,n)}T(V_0, \mathfrak{X}_0, W) \vee J_{(0,n)}T(V_0, \mathfrak{X}_0, W)) \& (V_0 \subset Z)))$,

$\Pi_{n,(g,1)}^0(Z, W) = \exists V_1 \exists \mathfrak{X}_1 \exists V_2 \exists \mathfrak{X}_2 (J_{(1,n)}T(V_1, \mathfrak{X}_1, W) \& (J_{(-1,n)}T(V_2, \mathfrak{X}_2, W) \& (V_1 \subset Z) \& (V_2 \subset Z))) \vee \exists V \exists \mathfrak{X} (J_{(0,n)}T(V, \mathfrak{X}, W) \& (V \subset Z))$.

2. $\Pi_{n,(g,2)}^+(Z, W) = \exists V \exists \mathfrak{X} ((J_{(1,n)}T(V, \mathfrak{X}, W) \& (V \subset Z)) \& \neg \exists V_0 ((J_{(-1,n)}(V_0 \Rightarrow_2 W) \vee J_{(0,n)}(V_0 \Rightarrow_2 W)) \& (V_0 \subset Z))$,

$\Pi_{n,(g,2)}^-(Z, W) = \exists V (J_{(-1,n)}(V \Rightarrow_2 W) \& (V \subset Z) \& \neg \exists V_0 \exists \mathfrak{X} (J_{(1,n)}T(V_0, \mathfrak{X}, W) \vee J_{(0,n)}T(V_0, \mathfrak{X}, W)) \& (V_0 \subset Z))$,

$\Pi_{n,(g,2)}^0(Z, W) = \exists V \exists V_0 \exists \mathfrak{X} ((J_{(1,n)}T(V, \mathfrak{X}, W) \& J_{(-1,n)}(V_0 \Rightarrow_2 W) \& (V_0 \subset Z) \& (V \subset Z)) \vee \exists V_1 \exists \mathfrak{X}_1 \exists V_2 ((J_{(0,n)}T(V_1, \mathfrak{X}_1, W) \& (V_1 \subset Z)) \vee (J_{(0,n)}(V_2 \Rightarrow_2 W) \& (V_2 \subset Z)))$

$\Pi_{n,(g,2)}^\sigma(Z, W)$ определены для $B\Phi = B\Phi_g^+ \cup B\Phi_b^-$, где $\sigma \in \{+, -, 0, \tau\}$.

3. $\Pi_{n,(g,3)}^+(Z,W) = \exists V((J_{(1,n)}(V \Rightarrow_2 W) \& (V \subset Z)) \& \neg \exists V_0 \exists \mathfrak{X}((J_{(-1,n)}T(V_0, \mathfrak{X}, W) \vee J_{(0,n)}T(V_0, \mathfrak{X}, W)) \& (V_0 \subset Z)))$,

$\Pi_{n,(g,3)}^-(Z,W) = \exists V \exists \mathfrak{X}(J_{(-1,n)}T(V, \mathfrak{X}, W) \& (V \subset Z)) \& \neg \exists V_0((J_{(1,n)}(V_0 \Rightarrow_2 W) \vee J_{(0,n)}(V_0 \Rightarrow_2 W)) \& (V_0 \subset Z))$,

$\Pi_{n,(g,3)}^0(Z,W) = \exists V \exists V_0 \exists \mathfrak{X}(J_{(1,n)}(V \Rightarrow_2 W) \& (J_{(-1,n)}T(V_0, \mathfrak{X}, W) \& (V \subset Z)) \& (V_0 \subset Z)) \vee \exists V_1 \exists V_2 \exists \mathfrak{X}_2((J_{(0,n)}(V_1 \Rightarrow_2 W) \& (V_1 \subset Z)) \vee (J_{(0,n)}T(V_2, \mathfrak{X}_2, W) \& (V_2 \subset Z)))$,

$\Pi_{n,(g,3)}^\sigma(Z,W)$ определены для $B\Phi = B\Phi_b^+ \cup B\Phi_g^-$, где $\sigma \in \{+, -, 0, \tau\}$.

$\Pi_{n,\bar{r}}^\sigma(Z,W) = \neg(\Pi_{n,\bar{r}}^+(Z,W) \vee \Pi_{n,\bar{r}}^-(Z,W) \vee \Pi_{n,\bar{r}}^0(Z,W))$, где $\bar{r} = \langle \lambda, j, \langle x, y \rangle \rangle$, $\lambda = g, b$, а $j = 1, 2, 3, 4$. Случай $j = 4$ был рассмотрен ранее.

Все определенные $\Pi_{n,\bar{r}}^\sigma(Z,W)$ используются в *н.п.в.* $-1_{\bar{r}}$, где $\bar{r} = \langle \lambda, j, \langle x, y \rangle \rangle$ для $Str_{x,y}$ из заданного множества стратегий ДСМ-рассуждений \overline{Str} .

Замечание 4-3. В табл. 1 имеются шестнадцать возможных объединений подмножеств БФ, характеризующихся посредством тестов на применимость предикатов $M_{ag}^\sigma(V, \mathfrak{X}, Y)$, где $\sigma \in \{+, -\}$ и Утверждения 1-3.

Случаи (4), (8), (12) – (16) представляют «невозможные миры», удовлетворяющие условию **неприменимости** ДСМ-рассуждений, которое есть отрицание условия их **применимости**.

Условие применимости ДСМ-рассуждений:

(AP) $\exists V \exists W M_{a,n}^+(V, W) \& \exists V \exists W M_{a,n}^-(V, W)$ ⁷.

Таким образом, условием неприменимости ДСМ-рассуждений является $\neg(AP)$.

$\neg(AP)$ представлено в табл. 1 случаями (4), (8), (12) – (16).

Случаи (1) – (3), (5) – (7) и (9) – (11) удовлетворяют условию (AP). Однако они подразделяются на два типа «возможных миров» – (α) и (β), содержащиеся в табл. 1.

Случаями (α) являются: (6), (7), (10), (11). Случаями (β) являются (1), (2), (3), (5), (9).

Случай (α)

(6) $B\Phi = B\Phi_g^+ \cup B\Phi_g^-$,

(7) $B\Phi = B\Phi_g^+ \cup B\Phi_b^-$,

(10) $B\Phi = B\Phi_b^+ \cup B\Phi_g^-$,

(11) $B\Phi = B\Phi_b^+ \cup B\Phi_b^-$.

Случай (β)

(1) $B\Phi = B\Phi_g^+ \cup B\Phi_b^+ \cup B\Phi_g^- \cup B\Phi_b^-$,

(2) $B\Phi = B\Phi_g^+ \cup B\Phi_b^+ \cup B\Phi_g^-$,

(3) $B\Phi = B\Phi_g^+ \cup B\Phi_b^+ \cup B\Phi_b^-$,

(5) $B\Phi = B\Phi_g^+ \cup B\Phi_g^- \cup B\Phi_b^-$,

(9) $B\Phi = B\Phi_b^+ \cup B\Phi_g^- \cup B\Phi_b^-$.

Очевидно, что в случае (6) применяются *н.п.в.* $-1_{g,1}^{(\sigma)}$; в случае (7) применяются *н.п.в.* $-1_{g,2}^{(\sigma)}$; а в случаях (10) и (11) – соответственно, *н.п.в.* $-1_{g,3}^{(\sigma)}$ и *н.п.в.* $-1_{g,4}^{(\sigma)}$, где $\sigma \in \{+, -, 0, \tau\}$.

Таким образом, в случаях (α) *н.п.в.* $-1_{g,j}^{(\sigma)}$, где $j = 1, 2, 3, 4$, эти правила применяются к **бинарным** объединениям подмножеств БФ; а в случаях (β) эти правила применяются к объединениям подмножеств БФ, состоящим из трех или четырех компонент. Следовательно, в случаях (β) требуется объединение результатов ДСМ-рассуждений и распознавание непротиворечивости таких объединений множеств гипотез о причинах и гипотез о предсказаниях исследуемых эффектов.

Сформулируем теперь правила правдоподобного вывода по аналогии (п.п.в.-2), соответствующие объединениям подмножеств БФ типа (α) и *н.п.в.* $-1_{g,j}$ и *н.п.в.* -1_b , где $j = 1, 2, 3$.

Обозначим посредством \bar{r} кортеж $\langle \lambda, j, \langle x, y \rangle \rangle$, где $\lambda = b, g$; $j = 1, 2, 3, 4$, а $\langle x, y \rangle$ имя стратегии ДСМ-рассуждений $Str_{x,y}$.

Для $\sigma = +$ имеется $\langle x, \neg y \rangle$, для $\sigma = -$ имеется $\langle \neg x, y \rangle$, а для $\sigma = 0, \tau$, соответственно, $-\langle x, y \rangle$ и $\langle \neg x, \neg y \rangle$. Тогда представим п.п.в.-2 следующим образом:

$$(II)_r^+ \frac{J_{(\tau,n)}(X \Rightarrow_1 Y), \Pi_{n,\bar{r}}^+(X, Y)}{J_{(1,n+1)}(X \Rightarrow_1 Y)},$$

$$(II)_r^- \frac{J_{(\tau,n)}(X \Rightarrow_1 Y), \Pi_{n,\bar{r}}^-(X, Y)}{J_{(-1,n+1)}(X \Rightarrow_1 Y)},$$

$$(II)_r^0 \frac{J_{(\tau,n)}(X \Rightarrow_1 Y), \Pi_{n,\bar{r}}^0(X, Y)}{J_{(0,n+1)}(X \Rightarrow_1 Y)},$$

$$(II)_r^\tau \frac{J_{(\tau,n)}(X \Rightarrow_1 Y), \Pi_{n,\bar{r}}^\tau(X, Y)}{J_{(\tau,n+1)}(X \Rightarrow_1 Y)}.$$

Замечание 5-3. Объединения подмножеств БФ типа (α) или (β) специфицируются относительно $Str_{x,y}$, которые образованы правилами индуктивного вывода (п.п.в.-1) и правилами вывода по ана-

⁷ $M_{a,n}^+(V, W), M_{a,n}^-(V, W)$ – предикаты минимального сходства [8].

логии (п.п.в.-2). Таким образом, имеются БФ $\langle x, y \rangle$, где $Str_{x,y} \in \overline{Str}$, $(I)_{x,y}^\sigma$, а пара $\langle x, y \rangle$ – имя стратегии $Str_{x,y}$.

Замечание 6-3. Термин «стратегия ДСМ-рассуждения $Str_{x,y}$ » неоднократно использовался в данной статье без формулирования его определения, данного в [9]. Приведем это определение Df.5. Стратегией ДСМ-рассуждения будем называть последовательное применение *п.п.в.* $-1^{(\sigma)}$ и *п.п.в.* $-2^{(\sigma)}$, которые состоят из правил $(I)^\sigma$, $(II)^\sigma$, где $\sigma \in \{+, -, 0, \tau\}$. Каждая стратегия имеет сигнатуру $\langle x, y \rangle$ для четырех прямых произведений решеток, соответствующих типам истинностных значений 1, -1, 0, τ . Стратегии обозначаются посредством $Str_{x,y}$, а их множество – посредством \overline{Str} .

Напомним введенную ранее терминологию и соответствующие обозначения.

Посредством БФ(p), где $p = 0, 1, \dots, s$, обозначаются базы фактов; $\Omega(p)$ – их описания (формулы вида $J_{\langle v, 0 \rangle}(C \Rightarrow_1 Q)$ и $J_{\langle \tau, 0 \rangle}(C \Rightarrow_1 Q)$, где $v \in \{1, -1, 0\}$). Результатом применения п.п.в.-1 к БФ(p) являются множества формул вида $J_{\langle \nu, n \rangle}(C' \Rightarrow_2 Q)$ и $J_{\langle \tau, n \rangle}(C' \Rightarrow_2 Q)$, где $v \in \{1, -1, 0\}$, $n > 0$, образующие множества $\Delta(p)$. Результатами применения п.п.в.-2 к БФ(p) являются множества формул вида $J_{\langle \nu, n \rangle}(C \Rightarrow_1 Q)$ и $J_{\langle \tau, n \rangle}(C \Rightarrow_1 Q)$.

Применение п.п.в.-2 к $B\Phi^\tau(p)$ и $\Delta(p)$, где $B\Phi^\tau(p) \subset B\Phi(p)$, порождает множества $\Omega_{2l}(p)$, где l – число тактов применения п.п.в.-1 \rightarrow п.п.в.-2 (такт есть их последовательное применение [10]).

Итеративная последовательность применения п.п.в.-1 и п.п.в.-2 (т.е., последовательность n тактов) оканчивается стабилизацией – новые гипотезы (для п.п.в.-1 и п.п.в.-2) не порождаются.

Пусть $n + 1$ – число тактов ДСМ-рассуждения, тогда последовательность n тактов, оканчивающихся стабилизацией представим следующим образом:

$$\begin{aligned} & (п.п.в. - 1(\Omega_0) \rightarrow п.п.в. - 2(\Delta_1))_1 \rightarrow (п.п.в. - 1(\Omega_2) \rightarrow \\ & \rightarrow п.п.в. - 2(\Delta_3))_2 \rightarrow \dots \rightarrow (п.п.в. - 1(\Omega_{2l}) \rightarrow \\ & \rightarrow п.п.в. - 2(\Delta_{2l+1}))_l \rightarrow \dots \rightarrow (п.п.в. - 1(\Omega_{2n}) \rightarrow \\ & \rightarrow п.п.в. - 2(\Delta_{2n+1}))_n, \Delta_{2n-1} = \Delta_{2n+1}, \end{aligned}$$

где

$$\begin{aligned} & п.п.в. - 1(\Omega_0) = \Delta_1, п.п.в. - 2(\Delta_1) = \Omega_2; \\ & п.п.в. - 1(\Omega_2) = \Delta_3, п.п.в. - 2(\Delta_3) = \Omega_4; \\ & п.п.в. - 1(\Omega_4) = \Delta_5, п.п.в. - 2(\Delta_5) = \Omega_6; \dots; \\ & п.п.в. - 1(\Omega_{2n-2}) = \Delta_{2n-1}, п.п.в. - 2(\Delta_{2n-1}) = \Omega_{2n}; \\ & п.п.в. - 1(\Omega_{2n}) = \Delta_{2n+1}, п.п.в. - 2(\Delta_{2n+1}) = \Omega_{2n+2}. \end{aligned}$$

Вместо п.п.в.-1 и п.п.в.-2 введем, соответственно, обозначения $(I)_{x,y}$ и $(II)_{x,y}$,

где

$$(I)_{x,y} = \{(I)_{x,y}^+, (I)_{x,y}^-, (I)_{x,y}^0, (I)_{x,y}^\tau\},$$

$$(II)_{x,y} = \{(II)_{x,y}^+, (II)_{x,y}^-, (II)_{x,y}^0, (II)_{x,y}^\tau\},$$

$$(II)_{x,y}((I)_{x,y}(\Omega_{2l}(p))) = \Omega_{2l+2}(p), \text{ где } l = 0, 1, \dots, n+1,$$

$$\text{а } p = 0, 1, \dots, s; (I)_{x,y}(\Omega_{2l}(p)) = \Delta_{2l+1}(p);$$

$$(I)_{x,y}(\Omega_{2l}(p)) = \Delta_{2l+1}(p), l = 0, 1, \dots, n+1;$$

$$(I)_{x,y}(\Omega_{2n}(p)) = \Delta_{2n+1}(p),$$

$$(I)_{x,y}(\Omega_{2(n-1)}(p)) = \Delta_{2n-1}(p). \text{ Условие стабилизации ДСМ-рассуждений: } \Delta_{2n+1}(p) = \Delta_{2n-1}(p).$$

В [8, 9] были определены оператор $O_{x,y}(\Omega)$ и его расширение $\bar{O}_{x,y}(\Omega) = O_{x,y}(\Omega) \cup \Omega$, представляющие взаимодействие **индукции (п.п.в.-1)** и **аналогии (п.п.в.-2)**, которое завершается применением **абдукции**, посредством функций $\rho^\sigma(p)$, где $\sigma \in \{+, -\}$, $p = 0, 1, \dots, s$.

Получим следующие определения операторов $O_{x,y}$ и $\bar{O}_{x,y}$:

$$\mathbf{Df.6-2.} \quad O_{x,y}(\Omega_{2l}(p)) = (II)_{x,y}((I)_{x,y}(\Omega_{2l}(p))) =,$$

$$= \Omega_{2l+2}(p) \bar{O}_{x,y}(\Omega_{2l}(p)) = \Omega_{2l+2}(p) \cup \Omega_{2l}(p),$$

$$\text{где } l = 0, 1, \dots, n+1; p = 0, 1, \dots, s,$$

$$\text{а } B\Phi_{x,y}(0) \subset \dots \subset B\Phi_{x,y}(s).$$

Очевидно следующее.

Утверждение 2-3. 1. $\Omega \subset \bar{O}_{x,y}(\Omega)$ (рефлексивность $\bar{O}_{x,y}$),

$$2. \bar{O}_{x,y}(\bar{O}_{x,y}(\Omega)) = \bar{O}_{x,y}(\Omega) \text{ (идемпотентность).}$$

Df.7-3. ДСМ-рассуждение, реализующее стратегию $Str_{x,y}$, является последовательностью применений ДСМ-оператора $\bar{O}_{x,y}(\Omega(p))$ и функций степени

$$\text{абдуктивного принятия гипотез } \rho^+(p), \rho^-(p),$$

$$\text{где } \Omega(0) \subset \dots \subset \Omega(s),$$

$$\text{а } p = 0, 1, \dots, s: \bar{O}_{x,y}(\Omega(0)), \rho^+(0), \rho^-(0), \dots,$$

$$\bar{O}_{x,y}(\Omega(s)), \rho^+(s), \rho^-(s).$$

Df.8-3. Будем говорить, что ДСМ-рассуждение **допустимо**, если $\rho^\sigma(s) \geq \bar{\rho}^\sigma$, где $\sigma \in \{+, -\}$, а $\bar{\rho}^\sigma$ – заданные пороги (например, 0, 8).

Заметим, что множеству стратегий \overline{Str} соответствует **семейство** ДСМ-рассуждений, образованное ДСМ-рассуждениями для $Str_{x,y}$ из \overline{Str} . Этим ДСМ-рассуждениям в свою очередь соответствуют прямые произведения решеток $IntL^\sigma$ и $Int(-L^\sigma)$, где $\sigma \in \{+, -\}$ [8, 9].

Рассмотрим снова случаи (β) «возможных миров» (объединений подмножеств БФ), соответствующих фиксированной стратегии $Str_{x,y}$. Наиболее общим случаем является (1)

$$B\Phi(x,y) = B\Phi_g^+(x,y) \cup B\Phi_b^+(x,y) \cup B\Phi_g^-(x,y) \cup B\Phi_b^-(x,y),$$

где $\bar{r} = \langle \lambda, j, \langle x, y \rangle \rangle$, а к $B\Phi_{\bar{r}}(x,y)$ применимы $(I)_{\bar{r}}^\sigma$ и $(II)_{\bar{r}}^\sigma$, где $\sigma \in \{+, -, 0, \tau\}$, которые соответствуют четырем вариантам п.п.в.: (1-1) с $\bar{r}_1 = \langle g, 1, \langle x, y \rangle \rangle$, (1-2) с $\bar{r}_2 = \langle g, 2, \langle x, y \rangle \rangle$, (1-3) с $\bar{r}_3 = \langle g, 2, \langle x, y \rangle \rangle$, (1-4) с $\bar{r}_4 = \langle b, 4, \langle x, y \rangle \rangle$.

П.п.в.-1 и п.п.в.-2 для (1-1), (1-2), (1-3) и (1-4) применимы, соответственно, к

$$B\Phi_{\bar{r}_1}(x,y) = B\Phi_g^+(x,y) \cup B\Phi_b^-(x,y),$$

$$B\Phi_{\bar{r}_2}(x,y) = B\Phi_g^+(x,y) \cup B\Phi_b^-(x,y),$$

$$B\Phi_{\bar{r}_3}(x,y) = B\Phi_b^+(x,y) \cup B\Phi_g^-(x,y),$$

$$B\Phi_{\bar{r}_4}(x,y) = B\Phi_b^+(x,y) \cup B\Phi_g^-(x,y).$$

Для вариантов объединений подмножеств БФ(x,y) будем применять ДСМ-операторы $\bar{O}_{\bar{r}}(\Omega(p))$, тогда допустимое ДСМ-рассуждение представимо последовательностью применений $\bar{O}_{\bar{r}}(\Omega(p))$ и $\rho^+(p), \rho^-(p)$, где $p = 0, 1, \dots, s$, а $\rho^\sigma(s) \geq \bar{\rho}^\sigma$.

Для случаев БФ типа (β) естественно определить результат применения $Str_{x,y}$ к БФ посредством объединения результатов для подмножеств БФ, образующих их объединения такие, что эти объединения будут **непротиворечивыми**.

Предварительно определим контрарные пары, для ДСМ-рассуждений. 1. Пусть $\nu, \mu \in \{1, -1, 0\}$ и $\nu \neq \mu$, тогда контрарными парами будут:

$$1. J_{\bar{\nu}}(C' \Rightarrow_2 Q), J_{\bar{\mu}}(C' \Rightarrow_2 Q) \text{ и } J_{\bar{\nu}}(C \Rightarrow_1 Q), J_{\bar{\mu}}(C \Rightarrow_1 Q);$$

$$2. J_{\bar{\nu}}T(C', \bar{x}, Q), J_{\bar{\mu}}(C' \Rightarrow_2 Q), \text{ где } \bar{x} - \text{константа};$$

$$3. J_{\bar{\nu}}T(C', \bar{x}, Q), J_{\bar{\mu}}T(C', \bar{x}, Q), \text{ где } \bar{x} - \text{константа}.$$

Дф.9-3. Пусть $\bar{O}_{\bar{r}}(\Omega(p)) = \tilde{\Omega}_{\bar{r}}(p)$

и $(I)_{\bar{r}}(\Omega(p)) = \tilde{\Delta}_{\bar{r}}(p)$,

тогда $\tilde{\Omega}_{\bar{r}_i}(p) \cup \tilde{\Omega}_{\bar{r}_j}(p)$ и $\tilde{\Delta}_{\bar{r}_i}(p) \cup \tilde{\Delta}_{\bar{r}_j}(p)$ будем называть **корректными** объединениями, если они, соответственно, не содержат контрарных пар.

Для обозначения корректных объединений множеств гипотез введем метапредикат **Consis** и будем писать $\text{Consis}(\tilde{\Omega}_{\bar{r}_i}(p) \cup \tilde{\Omega}_{\bar{r}_j}(p))$ и $\text{Consis}(\tilde{\Delta}_{\bar{r}_i}(p) \cup \tilde{\Delta}_{\bar{r}_j}(p))$.

Некорректные объединения будем обозначать посредством $\neg \text{Consis}(\tilde{\Omega}_{\bar{r}_i}(p) \cup \tilde{\Omega}_{\bar{r}_j}(p))$ и $\neg \text{Consis}(\tilde{\Delta}_{\bar{r}_i}(p) \cup \tilde{\Delta}_{\bar{r}_j}(p))$.

Таким образом, корректный результат применения $Str_{x,y}$ к $B\Phi(x,y)$ для варианта (1-1) с \bar{r} ,

где $j=1,2,3,4$ будет представлен посредством $\text{Consis}(\tilde{\Omega}_{\bar{r}_1}(p) \cup \tilde{\Omega}_{\bar{r}_2}(p) \cup \tilde{\Omega}_{\bar{r}_3}(p) \cup \tilde{\Omega}_{\bar{r}_4}(p))$ и $\text{Consis}(\tilde{\Delta}_{\bar{r}_1}(p) \cup \tilde{\Delta}_{\bar{r}_2}(p) \cup \tilde{\Delta}_{\bar{r}_3}(p) \cup \tilde{\Delta}_{\bar{r}_4}(p))$.

Аналогично представимы корректные объединения результатов применения $Str_{x,y}$ для других случаев (β) .

СПИСОК ЛИТЕРАТУРЫ

1. Милль Д.С. Система логики силлогистической и индуктивной. Издание пятое. – М.: ЛЕНАНД, 2011; Mill J.S. System of Logic Ratiocinative and Inductive, Being a Connected View of Principles of Evidence and The Methods of Scientific Investigation. – London: Parker, Son and Bowin, 1843.
2. Шестерникова О.П., Агафонов М.А., Винокурова Л.В., Панкратова Е.С., Финн В.К. Интеллектуальная система прогнозирования развития сахарного диабета у больных хроническим панкреатитом // Искусственный интеллект и принятие решений. – 2015. – №4. – С.12-50; Shesternikova O.P., Agafonov M.A., Vinokurova L.V., Pankratova E.S., Finn V.K. Intelligent System for Diabetes Prediction in Patients With Chronic Pancreatitis // Scientific and Technical Information Processing. – 2016. – Vol. 43, № 5-6. – P. 315-345.
3. Шестерникова О.П. О применении интеллектуальной системы прогнозирования развития диабета у больных хроническим панкреатитом // Искусственный интеллект и принятие решений. – 2016. – №3. – С. 62-71.
4. ДСМ-метод автоматического порождения гипотез: логические и эпистемологические основания / под общей ред. О.М. Аншакова. – М.: Книжный дом «ЛИБРОКОМ», 2009.
5. Финн В.К., Михеенкова М.А. Некоторые проблемы обобщенного ДСМ-метода автоматизированного порождения гипотез // Семиотика и информатика. Выпуск 33. – М.: МАИК НАУКА, 1993. – С. 136-163.
6. Финн В.К., Шестерникова О.П. О новом варианте обобщенного ДСМ-метода автоматизированной поддержки научных исследований // Искусственный интеллект и принятие решений. – 2016. – №1. – С. 57-63.
7. Финн В.К. Индуктивные методы Д.С. Милля в системах искусственного интеллекта. Часть I. // Искусственный интеллект и принятие решений. – 2010. – №3. – С. 3-21; Finn V.K. J.S. Mill's Inductive Methods in Artificial Intelligence Systems. Part I // Scientific and Technical Information processing. – 2011. – Vol. 38, № 6. – P. 385-402.
8. Финн В.К. Дистрибутивные решетки индуктивных процедур // Научно-техническая информация. Сер. 2. – 2014. – №11. – С. 1-29; Finn V.K. Distributive Lattices of Inductive JSM Procedures. – Automatic Documentation and Mathematical Linguistics – 2014. – Vol. 48, №6. – P. 263-295.

9. Финн В.К. О классе ДСМ-рассуждений, использующих изоморфизм правил индуктивного вывода // Искусственный интеллект и принятие решений. – 2016. – №3. – С. 48-61.
10. Финн В.К. Эпистемологические основания ДСМ-метода автоматического порождения гипотез // Научно-техническая информация. Часть I. – 2013. – №9. – С.1-29; Часть II. – 2013. – №12. – С. 1-26; Finn V.K. Epistemic foundations of the JSM method automatic hypothesis generation // Autom. Doc. Math. Linguist. – 2014. – Vol. 48, №2. – P. 96-148.
11. Финн В.К. Об интеллектуальном анализе данных // Новости искусственного интеллекта. – 2004. – №3. – С. 3-18.
12. Rosser J.B., Turquette A.R. Many – Valued Logics // North – Holland Publishing Company, Amsterdam. – 1958.
13. Барвайс Д. Введение в логику первого порядка. Справочная книга по математической логике. Часть I. Теория моделей. – М.: Наука, 1982. – С. 51-52; Handbook of mathematical logic / ed. J. Barwise. – North – Holland Publishing Company, Amsterdam, New York, Oxford. 1977.
14. Виноградов Д.В. Формализация правдоподобных рассуждений в логике предикатов 1-го порядка // Научно-техническая информация. Сер. 2. – 2000. – №11. – С. 17-20.
15. Финн В.К. Искусственный интеллект: методология, применения, философия. М.: КРАСАНД. – 2011.
16. Гретцер Г. Общая теория решеток. – М.: Мир, 1982; Grätzer G. General Lattice Theory. – Berlin: Academic – Verlag, 1978.
17. Финн В.К. Об определении эмпирических закономерностей посредством ДСМ-метода автоматического порождения гипотез // Искусственный интеллект и принятие решений. – 2010. – №4. – С. 41-48; Finn V.K. On the Definition of Empirical Regularities by the JSM Method for the Automatic Generation of Hypotheses // Scientific and Technical Information Processing. – 2012. – Vol. 39, №5. – P. 261-267.

Материал поступил в редакцию 07.07.17.

Сведения об авторах

ШЕСТЕРНИКОВА Ольга Павловна – начальник отдела разработки ООО «Прогтек», Москва.
e-mail: oshesternikova@gmail.com

ФИНН Виктор Константинович – доктор технических наук, профессор, заслуженный деятель науки РФ, главный научный сотрудник Института проблем информатики ФИЦ ИУ РАН; руководитель Отделения интеллектуальных систем в гуманитарной сфере Российского государственного гуманитарного университета, Москва.
e-mail: ira.finn@gmail.com

Д.Л. Ольшанский

О влиянии структур данных на производительность порождения гипотез в ДСМ-методе

Описывается влияние структур данных на производительность алгоритмов порождения гипотез в решателе ДСМ-системы. Впервые рассматриваются неоднородные комбинации структур данных. Именно такая комбинация дает существенный выигрыш по производительности на порядок. Влияние структур данных настолько существенно, что простейший алгоритм CbO (Close by One) с этой комбинацией оказывается существенно быстрее более совершенного FcbO (Fast Close by One) при традиционном выборе битовых векторов.

Ключевые слова: анализ формальных понятий, ДСМ, генерация гипотез, структуры данных

ВВЕДЕНИЕ

Поиск связи между структурой объекта и его свойствами – важная задача в области машинного обучения. В ДСМ-методе эта связь интерпретируется как причинно-следственная, характерные черты структуры объекта считаются причинами особенностей его поведения. Детальное описание интеллектуальных ДСМ-систем можно найти в работах В.К. Финна [1–3]. Первым шагом ДСМ-системы является *индукция*, порождение гипотез о причинах наличия (или отсутствия) целевых свойств по набору известных объектов. Следующий шаг — *аналогия*, применяет полученные гипотезы для определения наличия или отсутствия свойств у тех объектов, для которых это неизвестно.

Связь ДСМ-метода и анализа формальных понятий (АФП) установил С.О. Кузнецов [4, 5]. Это позволило применять алгоритмы порождения формальных понятий для генерации гипотез в ДСМ-системах. Заметим, что индукция является наиболее вычислительно затратной частью ДСМ-метода, и имеет экспоненциальную сложность в худшем случае. Наиболее эффективными на сегодняшний день являются алгоритмы семейства Close by One, во многом получившие развитие именно благодаря исследователям [6–10] в области анализа формальных понятий. Напротив, коллективы, изучающие ДСМ-метод заостряли вопрос в основном на применении алгоритмов порождения гипотез. В настоящей статье мы нарушим традиции и продемонстрируем важные особенности именно реализации алгоритмов.

Существенным вопросом, зачастую игнорируемым исследователями в области анализа формальных понятий, являются применяемые в алгоритмах структуры данных. С.О. Кузнецов одним из первых объяснил важность описаний структур данных в обзоре алгоритмов [11]. Новых работ, затрагивающих влияние структур данных на скорость работы алгоритмов, на

удивление немного, работа [12] является едва ли не единственной за последние 10 лет. В ней рассматриваются три алгоритма: NextClosure, CbO и Lindig и пять структур данных – сортированные массивы, сортированные связные списки, двоичные деревья, битовые массивы и хеш-таблицы. Ключевым недостатком этой работы является применение только однородных комбинаций структур данных и лишь классических алгоритмов, имеющих более современные аналоги.

В настоящей статье рассматриваются три классических типа структур данных: сортированные массивы, двоичные деревья и битовые массивы. Существенно новый шаг – применение неоднородных комбинаций структур данных различных типов для экстенционала и интенционала. Ключевым результатом является высокая эффективность именно одной из неоднородных комбинаций структур данных. В итоге, пять комбинаций структур данных были испытаны на трех алгоритмах, включая современные FcbO и InClose3. Подобная гибкость в реализации стала возможна благодаря применению языка C++11, имеющего большой потенциал в создании мощных эффективных абстракций и параметризации алгоритмов на основе типов (механизм шаблонов). Практически было бы невозможно реализовать такой объем комбинаций на языке более низкого уровня, например, С, используемого другими исследователями [13–15].

ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ

Будем придерживаться терминологии, употребляемой при анализе формальных понятий, так как они применяются в работах описывающих приведенные алгоритмы. Введем основные понятия.

Определение 1. Формальный контекст представлен тройкой $(X; Y; I)$, где X и Y – конечные множества, а I – отношение между X и Y , $I \subseteq X \times Y$. Элементы множества X называют объектами, а элементы мно-

жества Y – атрибутами. Запись $x \downarrow y$ означает, что $(x, y) \in I$, т.е. объект x обладает атрибутом y .

Таким образом, формальный контекст можно представить в виде прямоугольной таблицы, где колонкам соответствуют атрибуты, а строкам – объекты. Отношение I обозначим крестом в соответствующих позициях, где объекту принадлежит соответствующий атрибут (таблица).

Формальный контекст

	1	2	3	4	5	6	7
1	X	X	X	X	X		
2		X	X		X	X	
3		X	X	X		X	
4		X				X	X
5	X		X		X		X
6			X	X			X
7	X		X	X			X

Определение 2. Для любого подмножества A из X объектов контекста $(X; Y; I)$ определим оператор \downarrow , порождающий множество A^\downarrow общих атрибутов для каждого объекта из A :

$$A^\downarrow = \{ y \in Y \mid x \downarrow y, x \in A \}.$$

Аналогично, для любого множества атрибутов V из Y , введем оператор \uparrow порождающий множество объектов V^\uparrow , имеющих все атрибуты из V .

$$V^\uparrow = \{ x \in X \mid x \downarrow y, y \in V \}.$$

В совокупности эти операторы являются связью Галуа в $(X; Y; I)$ и позволяют дать определение формального понятия.

Пусть V — множество атрибутов, тогда V^\uparrow — множество объектов, и, следовательно, $V^{\uparrow\downarrow}$ — снова множество атрибутов. Следовательно у нас есть оператор: $V \subseteq Y \Rightarrow V^{\uparrow\downarrow} \subseteq Y$. Аналогично для множества объектов A , получим $A \subseteq X \Rightarrow A^{\downarrow\uparrow} \subseteq X$.

Определение 3. Формальное понятие в контексте $(X; Y; I)$ это пара (A, B) , где $A \subseteq X$, $B \subseteq Y$, и $A = B^\uparrow$, $B = A^\downarrow$. A называется экстенсионалом (объемом) понятия, B – интенсионалом (содержанием).

АЛГОРИТМЫ И РЕАЛИЗАЦИЯ

Современные версии Close by One (CbO) - FCbO и InClose3 общепризнанно являются одними из самых быстрых для порождения всех понятий формального контекста [15]. Рассмотрим подробно алгоритм CbO, его псевдокод представлен на рис. 1.

Алгоритм CbO [9] использует следующий тест на каноничность: полученный из V интенционал D является каноническим, если V и D согласуются по всем атрибутам до текущего атрибута j . Если же в D есть атрибут, идущий до j , которого нет в V , то понятие считается неканоническим, т.е. понятие каноническое, если:

$$V \cap Y_j = D \cap Y_j,$$

где Y_j – множество атрибутов до j , не включая j :

$$Y_j = \{ y \in Y \mid y < j \}.$$

```

1 CbO((A,B), y):
2   Output(A, B)
3   for j=y to n-1:
4     if j ∉ B:
5       C = A ∩ {j}↓
6       D = C↑
7       if B ∩ Yj == D ∩ Yj:
8         CbO((C,D), j+1)

Старт рекурсии - CbO(X, X↑)

```

Рис. 1. Псевдокод алгоритма CbO («Замыкай по одному»)

В исследовании были воссозданы CbO, InClose3 [15] и FCbO [14] в рамках одной реализации. Основной из примененных оптимизаций является сортировка атрибутов по их весу, т.е. числу объектов, имеющих данный атрибут. Несложно заметить, что, переставляя наиболее редкие атрибуты в начало контекста, мы снижаем вероятность получения неканонического интенционала (объекты с меньшим номером атрибутов – более редки) после замыкания.

Большой сложностью в реализации является вопрос выделения и освобождения памяти для множеств, в особенности в версиях алгоритмов, применяющих наследование записей об ошибках, где нет четких диапазонов существования отдельных переменных. Большая часть проблем решается посредством механизма «перемещения» (move semantics) в C++11, что позволяет передать множество в рекурсивный вызов, переложив на него ответственность за освобождение этой памяти. Вызывающая же процедура оставляет себе «нулевое», пустое множество.

ЭКСПЕРИМЕНТЫ

Испытания с описанной выше реализацией проводились в Amazon EC2 (Elastic Compute Cloud) на облачной виртуальной машине со следующими характеристиками: 32-х ядерный процессор Intel E5-2680 v2, 60 Гб оперативной памяти, дисковые носители – RAID1 на основе SSD (Solid State Drive). Применение SSD позволило параллельно запускать множество тестов (кроме тех, где уже используются все ядра процессора) не опасаясь искажений результата из-за низкой отзывчивости (latency) магнитных дисков при одновременной записи в несколько файлов. Удобство облачной инфраструктуры заключается в возможности на ограниченное время получить для подобных испытаний достаточно мощные вычислительные ресурсы.

Ни в одном из испытаний пределы доступной памяти не исчерпывались. Время на чтение данных и вывод гипотез было включено в общий результат, что делает условия испытаний максимально реалистичными (в отличие от более ранних результатов в [16]). Время выполнения программы и объем используемой памяти измерялись с помощью утилиты time. Для

большей точности измерений время выполнения также измерялось внутри самой программы с помощью библиотеки chrono из стандарта C++11, где применяется таймер ОС с максимальным доступным разрешением.

Для получения более достоверных результатов, для каждой комбинации параметров программы брался наименьший результат из трех запусков. Этот подход обоснован тем, что большинство шумов в измерениях имеют аддитивный характер – операционная система (ОС) реагирует на внешние события и, при неудачном стечении обстоятельств, отнимает небольшую долю процессорного времени у запущенной программы. Далее, с целью максимизации процессорного времени, выделяемого ОС задачам генерации гипотез, все тесты запускались с максимально возможным приоритетом. Такой эффект достигался с помощью утилиты nice со значением nicenesses -20, что является самым высшим приоритетом на ОС Linux и установление таких значений требует прав администратора.

В качестве первого блока испытаний тестировались сгенерированные случайным образом наборы данных, с применением широко распространенного генератора Mersenne Twister из стандартной библиотеки языка C++11.

Второй блок – реальные выборки из репозитория данных для машинного обучения UCI [UCI]. Как и в первом блоке, испытывались различные комбинации структур данных для экстенционала и интенционала.

Также на данных второго блока для сравнения были заново получены результаты от последней версии оригинальной программы FCbO на C из статьи [14]. За прошедшие годы технология компиляторов шагнула вперед и та же программа, скомпилированная заново показала результаты примерно на 20% лучше, чем версия, собранная в 2012 г.

Синтетические выборки

В блоке синтетических выборок используются данные полученные генератором случайных чисел с различными статистическими свойствами. Испытания описываются парой структур данных со следующими условными обозначениями: BIT – битовый вектор фиксированной длины (bit set), LIN – сортированный динамический массив (linear), TRE – двоичное дерево, из стандартной библиотеки C++11 (tree set). Чтобы не рассматривать все комбинации, тесты были ограничены «ортогональными» парами структур данных, предполагалось, что удачный выбор типа данных для экстенционала, в общем, не зависит от выбора интенционала и – наоборот. Всего имеется пять таких комбинаций: BIT-BIT, LIN-BIT, TRE-BIT, BIT-LIN, BIT-TRE. Результаты для этих пар приведены на рис. 2-4, каждый рисунок показывает зависимость от отдельного параметра выборки. На предварительных испытаниях BIT-TRE оказалась наиболее медленной и в целях экономии времени была исключена из результатов.

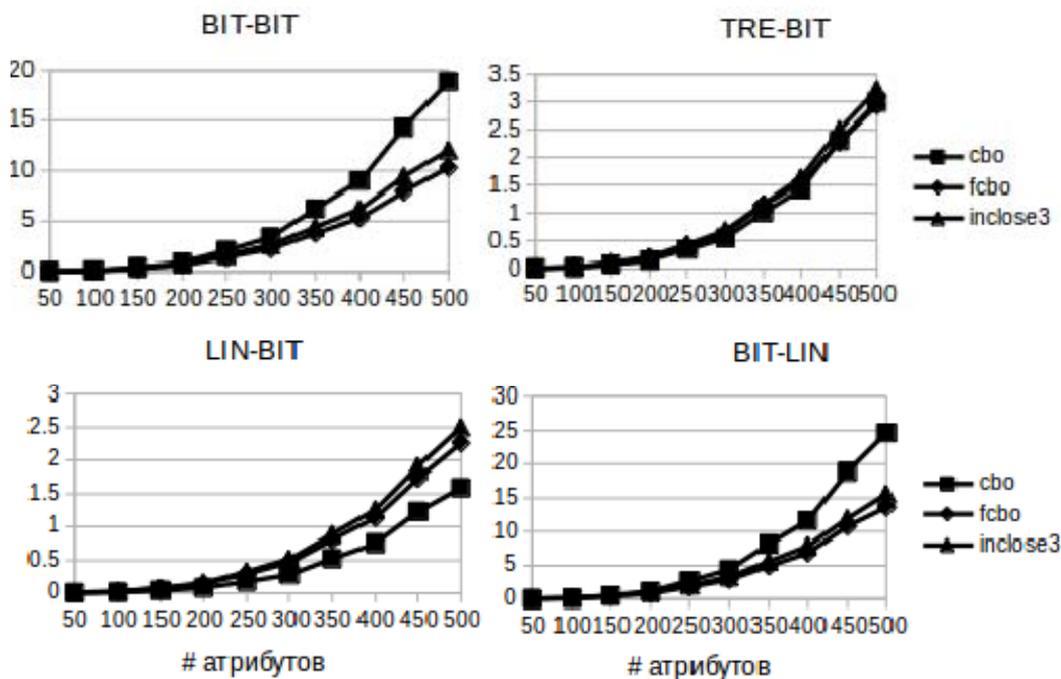


Рис. 2. Синтетические выборки, время выполнения в секундах в зависимости от числа атрибутов.

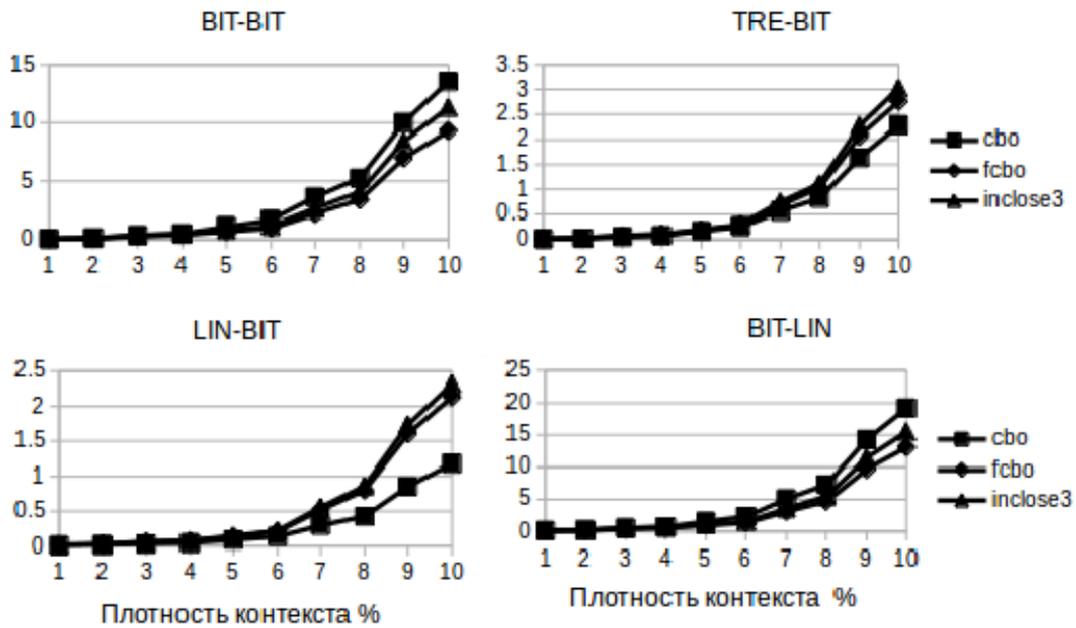


Рис. 3. Синтетические выборки, время выполнения в секундах в зависимости от плотности контекста.

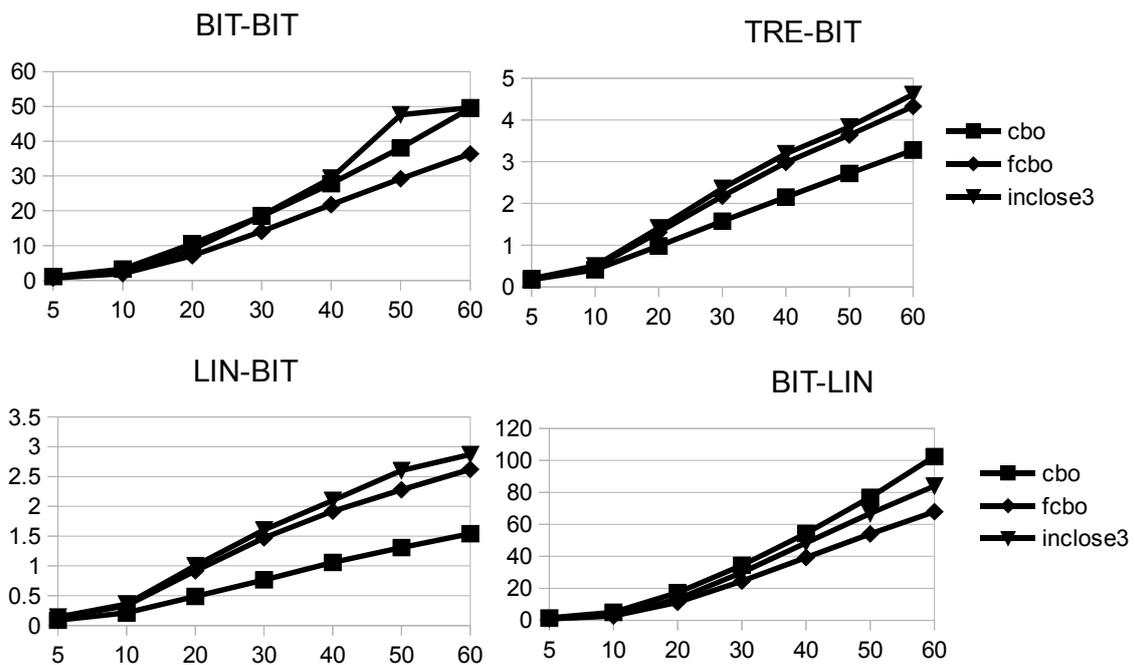


Рис. 4. Синтетические выборки, время выполнения в секундах в зависимости от числа объектов (тысячи).

Реальные выборки

Рассматриваются две комбинации структур данных на двух реальных выборках из репозитория UCI [17], пара BIT-BIT была выбрана поскольку применяется в программах других исследователей. Для приведения контекста к булевому виду использовался кодировщик FCABedrock [18], который заменяет непрерывные значения на дискретную шкалу и выполняет ряд других преобразований. Adult содержит

615 атрибутов, 32561 объектов, плотность контекста – 2,27%, полное число гипотез – 1719641. Mushroom – 120 атрибутов, 8124 объектов, плотность контекста – 19.16%, полное число гипотез – 238710. Эти выборки были взяты за контрольные, поскольку именно они приводятся в большинстве проводимых испытаний для алгоритмов генерации формальных понятий, а также имеют достаточно большой размер и отличающиеся показатели плотности. На рис. 5 и 6 приведены средние скорости генерации – гипотез/с.

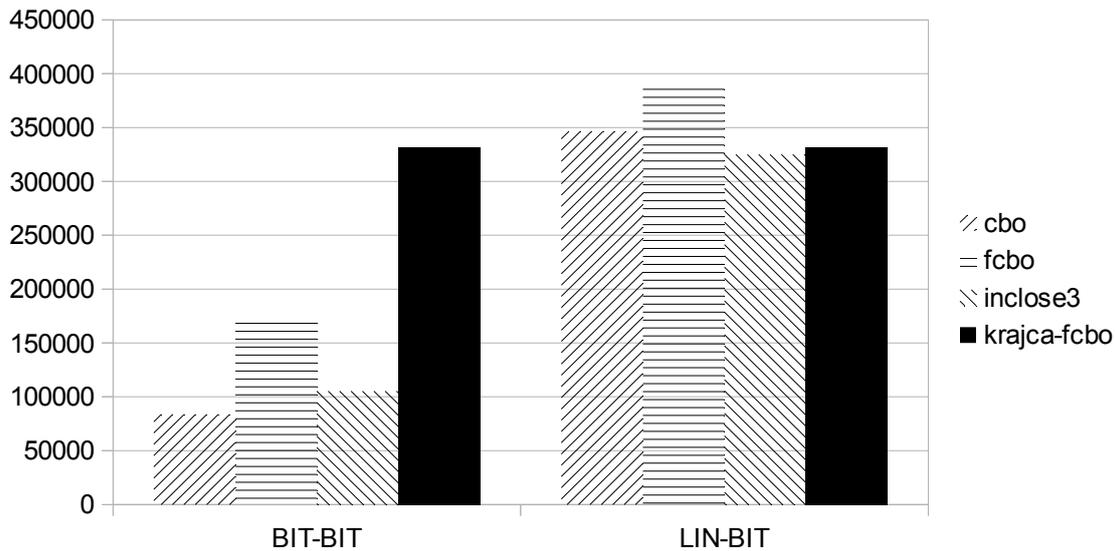


Рис. 5. Производительность последовательных алгоритмов на выборке MUSHROOM в гипотезах/с, и пиковые значения потребления памяти.

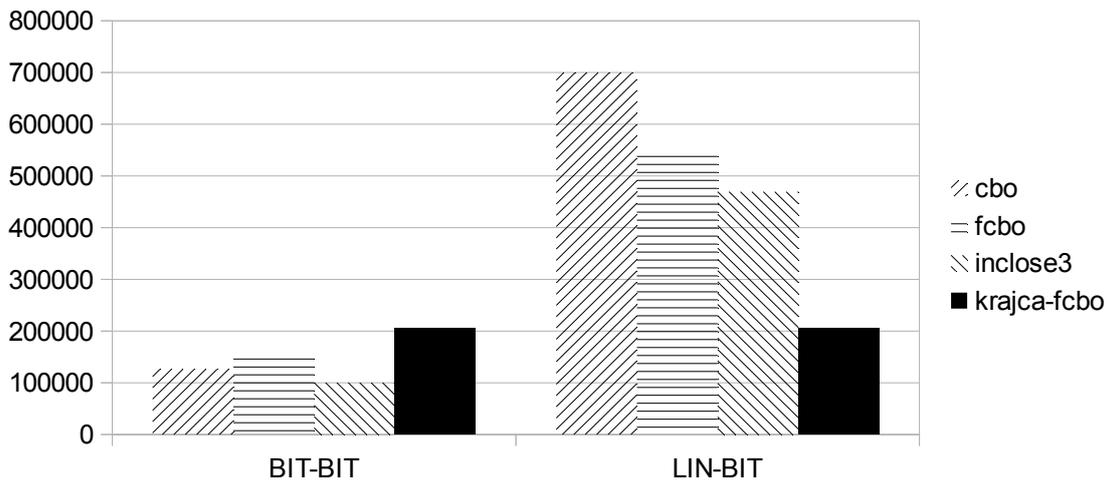


Рис. 6. Производительность последовательных алгоритмов на выборке ADULT, в гипотезах/с.

АНАЛИЗ РЕЗУЛЬТАТОВ

Принципиально новым результатом исследований является ускорение на порядок всех алгоритмов, полученное для следующей комбинации структур данных: сортированный динамический массив для экстенционала и фиксированный битовый – для интенционала. Имеется в виду, что динамический массив выгоден для разреженных множеств, а битовый массив для плотных. Очевидно, что именно эти условия соблюдаются во всех тестах. Тем не менее, выборки имеют разную структуру и плотность, а, следовательно, мы имеем дело с некоторой общей закономерностью.

Чтобы изучить причины низкой плотности экстенционалов в производных от СбО алгоритмах, получим оценку на плотность, исходя из предположения независимости всех атрибутов. А затем

оценим влияние на полученную оценку некоторых видов зависимостей.

Пусть X – множество объектов, Y – множество атрибутов. Обозначим P_i – долю всех объектов, обладающих атрибутом:

$$X = \{1, \dots, n\}$$

$$X = \{1, \dots, n\} Y = \{1, \dots, m\} P_i = \frac{|A|}{|X|},$$

где $A = \{x | (x, i) \in I\} \sum_{i \in Y} P_i = 1.0$

Отметим, что атрибуты j , где $P_j = 1.0$ можно отсеять как тривиальные (присущие абсолютно всем гипотезам) и, аналогично, $P = 0.0$ как не входящие ни в одну гипотезу. Пронумеруем все оставшиеся атрибуты по убыванию: $P_1 > P_2 > P_q$, где q – число не тривиальных атрибутов. Тогда для любого экстенционала

замыкание которого содержит k атрибутов, получим следующую оценку:

$$|B|=k, A=B^\downarrow, |A|=N_0 \prod_{i \in B} P_i$$

$$\prod_{i \in B} P_i \leq \prod_{i=1}^k P_i = \left(\sqrt[k]{\prod_{i=1}^k P_i} \right)^k$$

$$U_k = \sqrt[k]{\prod_{i=1}^k P_i}, |A|=N_0 U_k^k$$

$$U_k = \sqrt[k]{\prod_{i=1}^k P_i}, |A|=N_0 U_k^k, U_k U_k < \sup < P_i P_i < 1.0$$

где U_k – есть среднее геометрическое от плотностей произвольных k атрибутов данной выборки.

Несложно заметить, что наибольший U_k (от первых k атрибутов) не больше наибольшей плотности атрибута, которая, в свою очередь, меньше единицы.

Плотность же экстенционала описывается как U_k^k , т.е. экспоненциально убывает с ростом k . Приведем графики рассчитанных показателей $U_k(k)$ и $U_k^k(k)$ для выборки Mushroom. На рис. 7 и 8 приведены наибольшие, стохастические средние и минимальные величины параметров.

Рассмотрим также простейшие случаи зацепления атрибутов. Простейший случай – синонимы, такие пары (i,j) , где выполняется $j \in \{i\}^\downarrow$ и $i \in \{j\}^\downarrow$, с точки генерации гипотез их можно принимать за один атрибут. В общем случае где для k атрибутов $B = \{i_1, \dots, i_k\}$ замыкание $D = \{i_1, \dots, i_k\}^\downarrow \uparrow = \{i_1, \dots, i_k, j_1, \dots, j_q\}$, т.е. дает q дополнительных атрибутов. Плотность такого множества можно оценить, как произведение наибольших плотностей наименьшего числа атрибутов, замыкание которых даст D . Это не нарушает общей закономерности, но замедляет падение U_k для случая зависимых атрибутов.

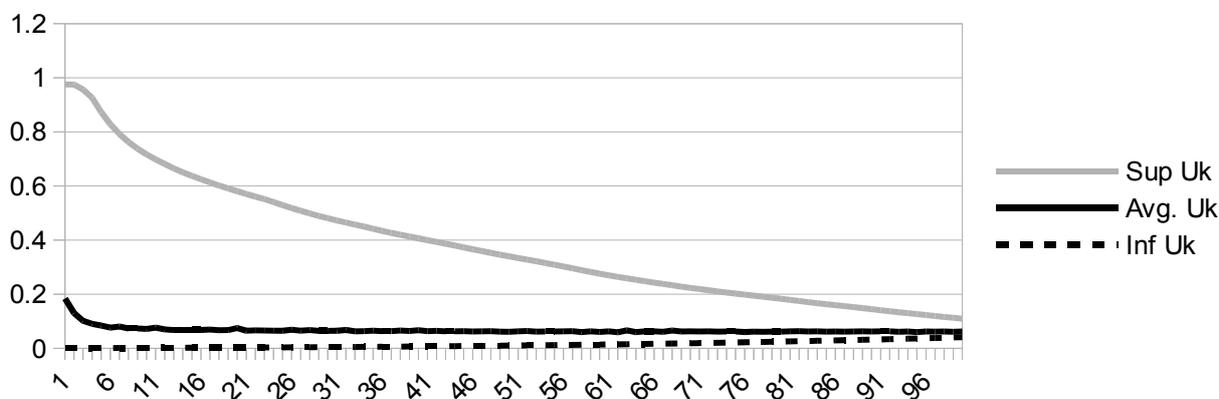


Рис. 7. График $U_k(k)$ для выборки Mushroom.

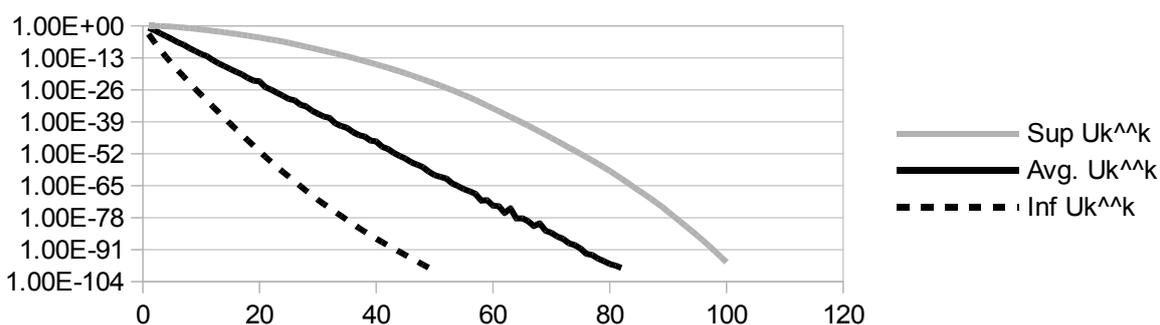


Рис. 8. График $U_k(k)$ для выборки Mushroom.

ЗАКЛЮЧЕНИЕ

В настоящей статье наглядно показано влияние выбора структур данных на скорости работы реализаций алгоритмов. В отличие от прочих статей по этой тематике, также изучены комбинации структур данных. В частности, одна из комбинаций, сортированные массивы для экстенционала и битовые векторы для интенционала,

проявили себя значительно лучше прочих. Причина подобного успеха неоднородных комбинаций структур разобрана и сводится к разным операциям, которые мы применяем для экстенционала и интенционала. Более того, автор полагает, что этот успех можно развить, создав специализированные структуры данных для экстенционала и интенционала, оптимизированные под конкретные операции.

СПИСОК ЛИТЕРАТУРЫ

1. Финн В.К. Индуктивные методы Д.С. Милля в системах искусственного интеллекта. Часть I // Искусственный интеллект и принятие решений. – 2010. – №3. – С. 3–21; Finn V.K. J.S. Mill's Inductive Methods in Artificial Intelligence Systems. Part I // Scientific and Technical Information processing. – 2011. – Vol. 38, №6. – P. 241-261.
2. Финн В.К. Индуктивные методы Д.С. Милля в системах искусственного интеллекта. Часть II // Искусственный интеллект и принятие решений. – 2010. – №4. – С. 14–40; Finn V.K. J.S. Mill's Inductive Methods in Artificial Intelligence Systems. Part II // Scientific and Technical Information processing. – 2012. – Vol. 39, №5. – P. 241-261.
3. Финн В.К. Об определении эмпирических закономерностей посредством ДСМ-метода автоматического порождения гипотез // Искусственный интеллект и принятие решений. – 2010. – № 4. – С. 41-48; Finn V.K. On the Definition of Empirical Regularities by the JSM Method for the Automatic Generation of Hypotheses // Scientific and Technical Information Processing. – 2012. – Vol. 39, №5. – P. 261-267.
4. Kuznetsov S.O., Obiedkov S.A. Comparing Performance of Algorithms for Generating Concept Lattices // Journal of Experimental and Theoretical Artificial Intelligence. – 2002. – Vol. 14. № 2–3. – P. 189–216.
5. Аншаков О.М. ДСМ-метод: теоретико-множественное объяснение // Научно-техническая информация. Сер. 2. – 2012. – №9. – С. 1-19.
6. Chein M. Algorithme de recherche des sous-matrices premie`res d'une matrice // Bull. Math. Soc.Sci. Math. R.S. Roumanie. – 1969. – Vol.13. – P. 21-25.
7. Norris E.M. An Algorithm for Computing the Maximal Rectangles in a Binary Relation // Revue Roumaine de Mathématiques Pures et Appliquées. – 1978. – № 23(2). – P. 243-250.
8. Bordat J.P. Calcul pratique du treillis de Galois d'une correspondance // Math. Sci. Hum. – 1986. – Vol.96. – P. 31-47.
9. Кузнецов С.О. Быстрый алгоритм построения всех пересечений объектов из конечной полурешетки // научно-техническая информация. Сер. 2. – 1993. – № 1. – С. 17-20; Kuznetsov S.O. Fast algorithm for construction of all intersections of objects from a finite semilattice // Automatic Documentation and Mathematical Linguistics. – 1993. – Vol. 27, № 1. – P. 23-28.
10. Ganter B., Wille R. Formal Concept Analysis: Mathematical Foundations. – Berlin: Springer-Verlag, 1999.
11. Kuznetsov S.O., Obiedkov S.A. Comparing Performance of Algorithms for Generating Concept Lattices // Journal of Experimental and Theoretical Artificial Intelligence. – 2002. – Vol. 14, № 2–3. – P. 189–216.
12. Krajca P., Vychodil V. Comparison of Data Structures for Computing Formal Concepts // Modeling Decisions for Artificial Intelligence. MDAI 2009. Lecture Notes in Computer Science. – 2009. – Vol. 5861.
13. Krajca P., Outrata J., Vychodil V. Advances in Algorithms Based on CbO // Proc. CLA. – 2010. – P. 325-337.
14. Outrata J., Vychodil V. Fast Algorithm for Computing Fixpoints of Galois Connections Induced by Object-Attribute Relational Data // Information Sciences. – 2012. – P. 114–127.
15. Andrews S. A 'Best-of-Breed' approach for designing a fast algorithm for computing fixpoints of Galois Connections // Information Sciences. – 2015. – P. 633-649.
16. Ольшанский Д.Л. Подбор алгоритма для параллельной реализации метода сходства в интеллектуальных ДСМ-системах // Научно-техническая информация. Сер.2 – 2015. – №7. – С.10-19; Ol'shanskii D.L. Selection of an Algorithm for the Parallel Implementation of the Similarity Method in Intelligent DSM Systems // Automatic Documentation and Mathematical Linguistics. – 2015. – Vol. 49, №4. – P. 109-116.
17. Lichman M. UCI Machine Learning Repository. – Irvine, CA: University of California, School of Information and Computer Science, 2013. – URL: <http://archive.ics.uci.edu/ml>
18. Andrews S., Orphanides C. F. A formal context creator // ICCS / eds. M. Croitoru, S. Ferre, D. Lukose. – 2010. – Vol. 6208 of LNCS. Springer – 2010.

Материал поступил в редакцию 17.07.17.

Сведения об авторах

ОЛЬШАНСКИЙ Дмитрий Леонидович – аспирант в Вычислительном центре имени А. А. Дородницына РАН, Москва
e-mail: dmitry.olsh@gmail.com

ИНФОРМАЦИОННОЕ ПИСЬМО И ПРИГЛАШЕНИЕ
МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ К 65-ЛЕТИЮ ВИНТИ РАН
«ИНФОРМАЦИЯ В СОВРЕМЕННОМ МИРЕ»
Москва, 25-26 октября 2017 г.

подробная информация на сайте: <http://www.viniti.ru>

Главный организатор:

Всероссийский институт научной и технической информации
Российской академии наук (ВИНИТИ РАН)

Соорганизаторы:

Российская академия наук
Федеральное агентство научных организаций
Российский фонд фундаментальных исследований
Министерство образования и науки РФ

Проблемно-тематическое направление конференции: современный издательский процесс, интеллектуальная собственность, научные библиотеки, информационное обеспечение научной и инновационной деятельности, информационные технологии для научной и библиотечной отрасли, информационная безопасность, международное сотрудничество и информационный обмен, инфометрия, классификации, стандартизация, образование для отрасли, экономика информации

Основные вопросы, предлагаемые к обсуждению:

- Популяризация научных знаний: Новые модели распространения научной информации
- Редакционно-издательская деятельность в цифровой среде: продукты и сервисы
- Издательские стандарты и технологии
- Перспективы развития книжного дела. Проекты и программы
- Взаимодействие цифровых и печатных ресурсов в научно-технической библиотеке
- Информационно-библиотечное обслуживание: сервисный подход
- Управление данными и навигация в современной научной библиотеке
- Научные библиотечные консорциумы – основные подписчики на научную литературу
- Перспективы развития национальных систем научно-технической информации
- Государственные проекты и программы поддержки информационного обеспечения научно-образовательной деятельности
- Тенденции развития региональных аналитических центров
- Информационное обеспечение экспертной деятельности. Использование информационно-аналитических систем для управления наукой и образованием
- Формальные и неформальные каналы развития современных научных коммуникаций

- Современные агрегаторы научной литературы открытого доступа как источник научной информации
- Машинная обработка данных и аналитические исследования: Приоритеты и сотрудничество
- Использование специальных сервисов компании CrossRef для идентификации научных публикаций
- Роль поисковых систем в современном издательском процессе
- Защита данных от несанкционированного использования. Маркеры безопасности. Политика безопасности открытых систем
- Вопросы достоверности и доверенности при обработке информационного потока
- Межгосударственный обмен научно-технической информацией на евразийском пространстве
- Информационное взаимодействие в рамках СНГ
- Международное партнерство при хранении и обработке больших массивов данных
- Современное состояние систем классификации знаний как инструмента индексирования и поиска данных по перспективным направлениям науки и критическим технологиям
- Современные библиометрические методы определения научных лидеров: Новые математические модели
- Анализ читательской аудитории научной литературы путем вебметрического анализа
- Подготовка специалистов в сфере научно-информационной деятельности
- Мастер-класс по работе с классификационными системами (УДК, ГРНТИ)
- Информация как источник цифрового капитала и фактор социальных изменений
- Информационная деятельность как фактор национальной экономики
- Новейшие бизнес-модели для публикаций открытого и закрытого доступа

На конференции планируются доклады представителей ведущих информационных центров и научно-технических библиотек России, СНГ и дальнего зарубежья.

В рамках юбилейной конференции состоится научно-практический семинар по классификационным системам «Перспективные направления научных исследований и критические технологии в классификационных системах». Предполагается проведение специализированных обучающих мероприятий по УДК индексированию. Запланировано заседание методического совета пользователей ГРНТИ и УДК. Участники конференции получают свидетельства о повышении квалификации.

Материалы конференции будут опубликованы в сборнике Трудов и на CD-ROM, основные – в сборнике **«Научно-техническая информация»**.

Доклады

Принимаются оригинальные работы, имеющие научное и прикладное значение, соответствующие тематическим направлениям конференции и НЕ ОПУБЛИКОВАННЫЕ ГДЕ-ЛИБО РАНЕЕ.

Предлагаемый доклад должен отвечать следующим требованиям:

1. Необходимо указать название доклада, фамилию, имя, отчество (полностью) авторов/соавторов, название организации, город, страну, выделить автора, который будет представлять доклад.
2. Необходимо наличие аннотации, раскрывающей содержание доклада. Размер аннотации - не более 850 знаков (включая пробелы).
3. Доклады принимаются только в электронной форме; тексты – в формате MS Word; схемы, диаграммы, фотографии, сканированные виды экранов и т. п. - в формате JPG. Объем доклада вместе с аннотацией, рисунками, приложениями и т.п. не более 10 страниц формата А4.
4. Доклад необходимо выслать по электронной почте до 11 сентября 2017 г. в адрес оргкомитета: conf@viniti.ru

Доклады, не соответствующие вышеуказанным требованиям,
НЕ РАССМАТРИВАЮТСЯ.

Программный комитет оставляет за собой право определять статус доклада (пленарный доклад, доклад, стендовый доклад), включать принятые доклады в те или иные секции.

Время для выступления: пленарные доклады – 15–20 мин., доклады на отдельных мероприятиях – до 10 мин. Доклады включаются в Труды на основании решения экспертов оргкомитета.

Контакты: 125190, Москва, ул. Усиевича, 20, ВИНТИ РАН

Телефоны: 8 (499) 152 61 13, 8 (499) 155 42 52, 8 (499) 151 02 61. Факс 8 (499) 943 00 60

Интернет-сайт: <http://www.viniti.ru> Эл. почта: conf@viniti.ru

Центр (Отдел) научно-информационного обслуживания (ЦНИО) ВИНИТИ РАН

Информационные услуги, предоставляемые ЦНИО ВИНИТИ РАН:

- проведение тематического поиска и консультации поисковых экспертов;
- подготовка списков научной литературы;
- подбор, копирование полнотекстовых материалов из первоисточников на бумажном носителе и в электронном виде;
- библиометрическая оценка публикационной активности исследователей и научных организаций с использованием российских и зарубежных баз данных;
- информационное обеспечение информационно-аналитической деятельности по подготовке и предоставлению аналитических обзоров и других научных материалов.

ВИНИТИ РАН располагает следующими информационными ресурсами:

- фондом НТЛ, включающим более 2,5 млн. отечественных и иностранных журналов, книг, депонированных рукописей, авторефератов диссертаций и другой научной литературы, ретроспектива – с 1991 года;
- базами данных и Интернет-ресурсами: БД ВИНИТИ (разработка ВИНИТИ), БД SCOPUS, БД Questel (патенты) и другими реферативными ресурсами;
- полнотекстовыми электронными ресурсами (статьи, патенты, материалы конференций).

Ознакомиться с информацией о доступных полнотекстовых и реферативных ресурсах можно на сайте ВИНИТИ www.viniti.ru

К услугам пользователей – **Электронный Каталог ВИНИТИ** <http://catalog.viniti.ru>
и **служба электронной доставки документов.**

Осуществляется платное информационное обслуживание по разовым заказам и на договорной основе с предоставлением всех необходимых финансовых документов.

Проводится индивидуальное обслуживание пользователей в читальном зале ЦНИО ВИНИТИ.

Обращаться в ЦНИО ВИНИТИ:

- адрес: 125190, Россия, г. Москва, ул. Усиевича, 20;
- телефоны: 8(499) 155 -42 -43, 8(499) 155 -42 -17;
- эл. почта cnio@viniti.ru, fdk@viniti.ru;
- факс 8(499) 930 -60 -00 (для ЦНИО).