СДІТ
МАТЕМАТИЧНІ МЕТОДИ, МОДЕЛІ, ПРОБЛЕМИ І ТЕХНОЛОГІЇ ДОСЛІДЖЕННЯ СКЛАДНИХ СИСТЕМ

УДК 007:681.3.00

СТРУКТУРНЫЕ СТАТИСТИЧЕСКИЕ МОДЕЛИ: ИНСТРУМЕНТ ПОЗНАНИЯ И МОДЕЛИРОВАНИЯ

Ф.И. АНДОН, А.С. БАЛАБАНОВ

Даны введение и краткий обзор перспективного класса моделей — вероятностных моделей зависимостей на основе ациклических ориентированных графов (АОГ), прежде всего — байесовских сетей. Приведены характеристика выразительных и познавательных возможностей АОГ-моделей, их способности отображать причинно-следственные связи. В сопоставлении с другими подходами к выводу знаний и идентификации моделей показаны роль и место байесовских сетей как инструмента анализа и обобщения эмпирических данных, связь с логикой и проблемой индукции.

ВВЕДЕНИЕ

Вероятностные модели систем зависимостей на основе графов возникли на стыке многомерного статистического анализа, теории вероятностей, теории графов, теории информации и искусственного интеллекта. Данный класс моделей играет роль строгого языка представления знаний в условиях неопределенности (в частности, в экспертных системах нового поколения) и эффективного аппарата решения разнообразных аналитических задач. Наиболее привлекательны модели на базе ациклических ориентированных графов (АОГ), т.е. АОГ-модели. Достоинства АОГ-моделей: наглядность, способность отображать причинно-следственные связи и прогнозировать последствия действий (решений), компактное представление систем зависимостей, вычислительная эффективность вероятностного вывода по свидетельствам [1–4]. Эти свойства обеспечивают эффективное применение таких моделей в самых разных сферах, включая медицинскую и техническую диагностику, социометрический, эконометрический и эпидемиологический анализ, моделирование генетических механизмов, распознавание речи и т.д. Можно сказать, что такие модели пронизывают информационные технологии формирующегося ныне комплекса дисциплин e-Science. В работе [5] показано, что формализм АОГ-моделей зависимостей адекватно отображает когнитивные механизмы обучающихся детей при индукции каузальных отношений из опыта.

Простой иллюстративный пример АОГ-модели дан на рис. 1. Стрелки между узлами (вершинами) модели отображают непосредственное влияние одних переменных (событий) на другие. В частности, показано, что мостовая может стать мокрой в результате того, что прошел дождь или лопнула труба водопровода. Примером реального приложения АОГ-модели является нормативная модель «Hailfinder», используемая для предсказаний сильного летнего града на северо-востоке Колорадо. Модель содержит 56 дискретных переменных, связанных с помощью 66 дуг.

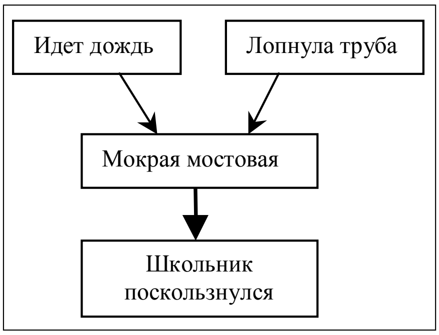

Рис. 1. АОГ-модель

Вывод подобных моделей из данных способствует получению инсайта о связях процессов в предметной области. Например, был проведен анализ причин и следствий бедности в 80 развивающихся странах мира, исходя их данных Мирового Банка [6]. В результате построена модель (рис. 2), которая показывает факторы рождаемости. В частности, видно, что GDP — величина дохода домашнего хозяйства (на душу населения) — непосредственно не влияет на рождаемость, но связана с ней индексом несвободы и Gini-Index.

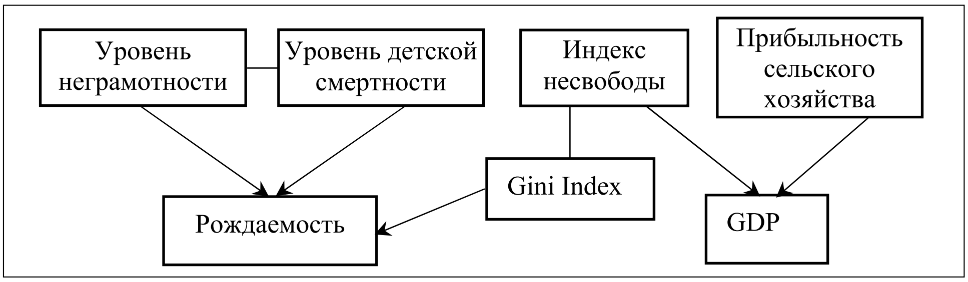

Рис. 2. Структура факторов рождаемости в бедных странах

Приведем также пример из медико-биологической области. Были собраны данные экспрессии генов образцов клеток простаты — раковых и контрольных образцов. Из этих данных выведена модель [7], показывающая, что при патологии механизмы генной регуляции перестраиваются (рис. 3). В частности, при патологии (рис. 3,*б*) онкоген ERG начинает влиять на коллаген COL2A1 и, кроме того, ген TGFβ (фактор роста) прямо влияет на α–methylacyl-CoA racemase (известный как маркер дифференциации опухоли), а в норме — нет. При патологии ген иммунной системы

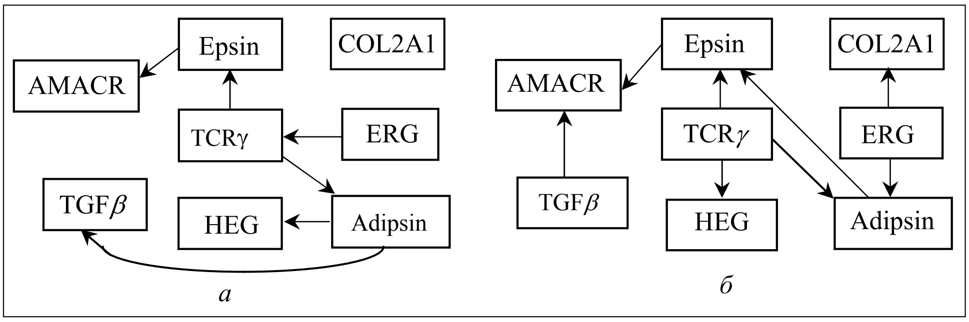

Рис. 3. Взаимодействие генов клетки простаты: *а* — норма; *б* — опухоль; AMACR — α-methylacyl-CoA racemase; HEG — Human enkephalin gene

Adipsin непосредственно влияет на Epsin (ген роста), а в норме — нет. (В норме они ассоциированы только через TCRγ — ген клеточной защиты.)

КРАТКАЯ РЕТРОСПЕКТИВА СТРУКТУРНЫХ МОДЕЛЕЙ

Проблемы познания традиционно рассматривались на философском уровне. Вопросами эмпирической методологии занималась философия науки, привлекая при этом аппарат логики и теории вероятностей. С появлением мощных информационно-аналитических инструментов на базе компьютерных технологий (в том числе интеллектуальных) эти вопросы приобретают непосредственно практическое значение. В помощь и на смену традиционным процедурам и схемам познания, где одна из центральных ролей отводится процедурам обсуждения, коллективного контроля и согласования теорий, приходят автоматизированные технологии «ускоренного познания». Сформировалась новая область научных исследований — открытие знаний в данных, лежащая на пересечении разных дисциплин. В настоящее время область приложений этих исследований в основном сосредоточивается там, где фундаментальная наука (в чистом виде) не дает нужного результата [8, 9] (например, в бизнесе), но сфера таких приложений расширяется. На смену первым простым моделям (правилам классификации и распознавания, пропозициональным формулам и т.п.) идут более сложные модели и системные представления. Одним из таких классов моделей являются графовые вероятностные модели зависимостей.

Заметим, что к известным различиям между традиционной и компьютерной методологиями научного открытия необходимо добавить различие среды открытия. Прежде чем сделать открытие, человек долгие годы учится, готовится, специализируется и аккумулирует достижения науки в данной области. Компьютерные же технологии полагаются на ускоренные автоматические процедуры «феноменологического» познания, которые за один «проход» охватывают (усваивают) огромный массив данных и трансформируют их в модель, пригодную для объяснения. Поэтому в работе [8] при сопоставлении различных парадигм и методологий открытия в качестве одной из осей диаграммы взята характеристика «объем используемых знаний». А, вообще, правильнее не противопоставлять эти две методологии, а говорить об их взаимном проникновении и дополнении [8–10].

Аппарат формальной логики и аппарат вероятностных (статистических) моделей зависимостей на протяжении своей истории долго развивались почти самостоятельно. Но попытки чисто логическими средствами описать реальные ситуации и задачи искусственного интеллекта натолкнулись на серьезные трудности и стимулировали разработку различных расширений и обобщений логики и языков представления знаний. Трудности возникали и на пути прямой интеграции логики и вероятности. В то же время за последние десять–пятнадцать лет была показана важная роль и плодотворность понятия условной независимости. Успешно развивались теория и практическое применение статистических моделей на основе графов, особенно, моделей зависимостей на основе ациклических орграфов. Такие модели играют важную роль в методологии открытия знаний в базах данных

(БД). В связи с этим становятся актуальными вопросы переосмысления проблемы обобщения и индукции.

Потребность ввести аппарат вероятности в эпистемологию и познавательный инструментарий осознавали давно. Философские концепции создавали, в частности, К. Поппер, Р. Карнап и Х. Путнам. Предпринимались попытки синтеза вероятности и логики. Например, П. Гайек, Т. Гавранек [11] предложили частотные градации квантора всеобщности. Но главное направление развития пошло путем введения вероятностной меры истинности утверждений и теорий, в сочетании с характеристикой точности количественных описаний [12, 13].

Неприменимость классической логики к решению большинства практических задач, а также потребность вероятностных расширений способов рассуждений можно проиллюстрировать очень простой схемой. Пусть мы знаем, что истинно правило «Если *A*, то *B*». Пусть теперь мы получили факт *B*. В классической логике какой-либо вывод из этого факта отсутствует. А вот здравый смысл, используя «слабый силлогизм» [14], позволяет вывести, что наступление факта *B* делает утверждение *A* более правдоподобным. (В более широкой трактовке такой способ рассуждения — логика подтверждения гипотезы или теории.) Но в какой степени повышается правдоподобность утверждения *A*? Для ответа на этот вопрос необходима вероятностная модель.

Индуктивная логика не может работать без вероятностной модели предметной области. Давно сложилась традиция различать частотный и байесовский подходы к вероятности. Именно байесовский аппарат обычно привлекается к построению рассуждений. Ряд авторов, в частности, Е. Джейнес [14], рассматривают байесовский аппарат теории вероятностей как логику рассуждений в научном познании. (Большая часть примеров в указанной монографии взята из математики или ситуаций обыденной жизни, так что они не репрезентуют полномасштабных проблем научного познания.) Байесовский вывод из множества свидетельств (длинная цепочка рассуждений) порождает разрастание громоздких условий. Это (среди прочего) поясняет важность свойства условной независимости.

Аппарат условной независимости, формализованной в теории графоидов и полуграфоидов [1], имеет несколько экспликаций. Это прежде всего: вероятностная условная независимость, нерелевантность в логической системе [15] и нерелевантность в реляционных базах данных (multivalued dependency). Их объединяют общие аксиомы. Заметим, что модели зависимостей на основе ациклических орграфов могут рассматриваться и как инструмент работы с субъективными вероятностями, и как статистические модели на основе частотных вероятностей. Модели, используемые для субъективных рассуждений, обычно называются belief networks [1] и строятся на основе экспертных знаний. Аппарат графовых вероятностных моделей зависимостей объединил в себе несколько направлений, развивавшихся ранее автономно — статистические модели, системы структуральных уравнений (в эконометрике), сети вероятностных рассуждений (в искусственном интеллекте). Зарождение этого аппарата можно усмотреть в путевых диаграммах С.Райта (начало 20-го века).

Не претендуя на всесторонний анализ, рассмотрим условную независимость и модели зависимостей на основе ациклических орграфов с точки зрения логической и познавательной перспектив, а также взаимоотношение индуктивного вывода и идентификации структуры статистических моделей, некоторые возможные варианты «сплава» логики и аппарата графовых вероятностных моделей зависимостей.

ОБОБЩЕНИЕ И ИНДУКЦИЯ

Обобщение — это процесс получения (вывода) более общих описаний (абстрактных, универсальных и компактных спецификаций) определенного объекта (системы, предметной области, сферы или мира) на основе частных описаний.

Понятие индукции употребляется главным образом в логике (хотя еще Ф. Бэкон применял его в контексте эмпирического познания). Но в последнее время смысл этого термина становится более свободным. Поэтому будем различать индукцию в узком и широком смыслах. В широком смысле — это вывод теории или общего утверждения на основе частных случаев. В узком смысле она определяется в рамках формально-логических языков. В работе [16] показано семантическое различие между задачами дедукции, абдукции и индукции, которые встречаются в логическом программировании и искусственном интеллекте. Индукция определена там как вывод модели неизвестной теории, согласующейся со всеми заданными положительными и отрицательными интерпретациями. В пропозициональном исчислении индукцию можно понимать как обоснованный переход от множества выражений с константами к выражению, где константа заменена на переменную. В индуктивном логическом программировании выводится формула, составленная в заданном базисе и объясняющая («покрывающая») заданное множество фактов [17].

Для наших целей определим локальную индукцию (в теориях первого порядка) как обоснованный переход от множества логических выражений с константами на месте предметных переменных в некоторых позициях к сходному по форме логическому выражению, в котором на этих позициях стоят переменные (не связанные кванторами существования). Можно сказать, что при этом вводится новый квантор всеобщности.

Чисто логическая индукция эффективна, по-видимому, только в приложении к искусственным системам или «мирам» (например, при верификации программ) или в очень абстрактных отображениях реальных предметных областей. Известные примеры успешного приложения индуктивного логического программирования (ILP) можно объяснить тем, что был выбран изолированный участок предметной области, подчиняющийся детерминистическим связям (например, клетка), и удачно сформулирована задача. (В последние годы разрабатываются вероятностные версии ILP.) Широко известно, что для реального бесконечного мира строгих и надежных правил индукции не существует, т.е. никакое конечное множество примеров (фактов, частных утверждений) не может обосновать абсолютно верное обобщенное утверждение. Принцип математической (арифметической) индукции — это просто правило сворачивания рекурсивной схемы доказательства, он не имеет прямого отношения к проблеме научной индукции. В реальных бесконечных доменах не работает правило переноса утверждения с одного объекта (индивида) на следующий (и нет даже естественного упорядочения объектов). Поэтому научное сообщество (прежде всего, философы и теоретики computer science) пришли к концепции PAC-learning, т.е., вероятно приблизительно корректного вывода обобщенных утверждений из множества частных [12].

В англоязычной литературе понятие *обобщение* часто связано с понятием learning. В русскоязычной и украиноязычной традиции такой ассоциации нет, вероятно потому, что слово *обучение* охватывает более широкую семантику. (Понятие *самообучение* несколько ближе, но тоже не является точным эквивалентом. *Научение* — лучше.)

Неформально, в широком контексте, обобщение — это переход (обоснованный или аргументированный) от частных случаев (примеров) к общим, более универсальным утверждениям (знаниям) о некотором объекте, системе, предметной области, сфере или мире в целом. Например, в социальных и медико-биологических исследованиях это имеет форму вывода знаний о популяции в целом на основе фактов, касающихся отдельных членов популяции. Когда такой процесс вывода (знаний) носит скачкообразный характер (полученный результат существенно отличается от исходной информации по глубине, компактности, степени универсальности) и касается широкой сферы или даже всего универсума, то говорят о научном открытии. В большинстве случаев научные открытия были сделаны благодаря интуитивному акту постижения истины (озарению). Более того, опыт человечества говорит, что совершить открытие (по крайней мере, в естественных науках) путем формальной процедуры, которая гарантировала бы от ошибок, невозможно. И даже если ограничиться утверждениями в строгом логическом языке (например, в логике первого порядка), то все равно невозможно выполнить индуктивный вывод из конечного числа примеров с полной гарантией от ошибки. Поэтому теоретики приняли концепцию вероятностного аппроксимационного обучения на эмпирическом опыте [12, 13].

ВЫЯВЛЕНИЕ СТРУКТУРЫ ВЕРОЯТНОСТНОЙ МОДЕЛИ КАК ИНДУКЦИЯ

Наука в первую очередь изучает связи. Относительно новым аппаратом, позволяющим явно описывать связи между процессами, являются графовые вероятностные модели зависимостей [1–4,18,19]. Рассмотрим наиболее популярный класс — модели на базе ациклических ориентированных графов (АОГ-модели). Поскольку именно дискретные модели позволяют «перебросить мостик» к логике, сосредоточим внимание на АОГ-моделях с дискретными переменными, т.е. на так называемых байесовских сетях. Соответственно, говоря о БД, будем иметь в виду базы с дискретными атрибутами.

Байесовская сеть (БС) определяется:

1) структурой ациклического орграфа G (где каждой вершине графа соответствует переменная) и

2) совокупностью параметров ϑ, т.е. локально заданных условных распределений вероятностей значений переменных $p(R \,|\, \pi(R))$, где $\pi(R)$ — графовые «родители» для R (вершины, из которых выходят дуги, входящие в R).

Модель описывается совместным распределением

$$p(R_1, R_2, ..., R_n) = \prod_i p(R_i \mid \pi(R_i)) . \tag{1}$$

Формула (1) подразумевает выполнение каузального марковского допущения, которое в графовых терминах выражается формулами условной независимости вида

$$p(R \,|\, \text{ все атрибуты, за исключением потомков } R) = p(R \,|\, \text{ родители } R).$$

(Отношение потомки–предки определяется транзитивным замыканием отношения дети–родители.) Графовые модели зависимостей иногда называют сетями вероятностной условной независимости.

В чем же заключается обобщение в процессе идентификации структуры вероятностной АОГ-модели на основе данных? Можно ли считать индуктивным выводом процесс восстановления модели зависимостей в виде ациклического орграфа (исходя из данных)?

Кратко обрисуем задачу идентификации (восстановления) байесовской сети по данным наблюдений (причем рассмотрим «сепарационный» подход к решению задачи) [3,19]. Задано: а) статистические данные наблюдений за системой (объектом), для которой мы ищем модель; б) предположение, что адекватная модель относится к классу АОГ-моделей; в) несколько концептуальных методологических постулатов, необходимых для индуктивного вывода (в частности, предположение необманчивости распределения вероятностей модели и стандартное предположение любого статистического моделирования – *i.i.d.*-характер выборки данных). Результатом идентификации должна стать спецификация структуры графа G и совокупности параметров ϑ.

На первый взгляд может быть далеко не очевидно, в чем же заключается в данном случае обобщение или индукция, т.е. какое общее знание получено. Может показаться, что просто выполнен синтез компактного описания исходных данных. Модель выведена, но в чем состоит обобщение?

Напомним, что каждая запись исходных данных соответствует случаю, прецеденту, т.е. поведенческому акту объекта, к которому относятся данные. В разных предметных областях такие случаи описываются как кортежи состояний переменных, разных по природе, характеру, длительности и уровню. Например, записям (случаям) могут соответствовать:

- индивиды человеческой популяции (в медико-эпидемиологических задачах);
- транзакции или запросы, проходящие через процессор (в компьютерных системах);
- сообщения (в телекоммуникациях);
- сигналы или импульсы (в электронных или электротехнических устройствах);

- изделия на выходе производственной линии (например, отливки металла, штампованные детали, экземпляры интегральных схем и т.д.).

Для последнего примера (производство БИС) запись данных будет состоять из результатов контроля БИС, параметров технологического процесса и других измерений и показателей, предположительно релевантных для анализа. В данном примере записям соответствуют экземпляры однотипных изделий. В других приложениях все данные могут относиться к одному и тому же устройству (обычно — сложному и долговечному), и тогда случаи описывают разные циклы, интервалы, акты функционирования этого устройства. А элементами записи (атрибутами случая) выступают значения измеряемых величин в соответствующие моменты (интервалы) времени.

Рассмотрим синтаксическую форму выражения полученных знаний. Пусть мы хотим записать результат обобщения в форме логического (импликативного) правила, предсказывающего значение одного атрибута, исходя из значений других. Содержательно такое выражение — конъюнкция одноместных предикатов равенства, причем во все предикаты входит одна и та же предметная переменная *x*, а перед всем выражением стоит (подразумевается) квантор всеобщности. В полной логико-подобной форме фрагмент байесовской сети следовало бы записать в виде комплекта квазиправил вида

$$(A(x) = a_1 \,\&\, B(x) = b_2 \,\&\ldots\&\, F(x) = f_i) \Rightarrow \text{Распределение атрибута } R(x)\,. \quad (2)$$

Чтобы придать правилам, описывающим байесовскую сеть, еще большее сходство с привычными логическими выражениями, можно ввести в левую часть правила дополнительную переменную H, имеющую вероятностную природу и обеспечивающую однозначное значение в правой части. Тогда получаются правила, в которых нет распределений вероятностей, например:

$$(A = a_1)\,\&\,(B = b_2)\,\&\,(H = h_j) \Rightarrow R = r_k\,.$$

(Переменная H имеет больше значений, чем R.) Такая форма представления не только внешне привлекательна, но и отражает механизм возникновения вероятностной природы отношения — существование ненаблюдаемых аргументов (факторов).

Любая запись данных в полной форме выглядит как

$$\left\langle A(x) = a_1 \,\&\, B(x) = b_j \,\&\ldots\&\, F(x) = f_s \ldots \right\rangle. \quad (3)$$

(Чтобы сделать запись еще более привычной для логиков, надо заменить равенства на двуместные предикаты с аргументами типа *объект* и *значение атрибута*.) Но записывать данные в логико-подобной форме (3) нецелесообразно, и на практике записывают просто кортеж значений.

Таким образом, результатом идентификации (восстановления) является модель, описывающая (в вероятностном смысле) поведение *всех* экземпляров популяции. Естественно, экстраполяция описания на новые (еще не наблюдавшиеся) случаи подразумевает статистическую значимость утверждений, на основе которых сконструирована модель, а также определенную степень стационарности и стабильности системы (источника данных). Итак, можно сказать, что при идентификации модели *акт обобщения*, или индукции, произошел.

Ясно, что конструкции, подобные (2), ближе к пропозициональному исчислению, чем к выражениям логики первого порядка [20]. Символ предметной переменной обычно никогда не вписывают ни в текст данных, ни в результат вывода. (Все случаи считаются безымянными.) Даже если в БД присутствуют уникальные идентификаторы случаев (индивидов популяции), модель их игнорирует. Переменными в таких моделях обычно называют атрибуты. Однако назвать такую запись чисто пропозициональной тоже нельзя, ибо тогда можно потерять связующую ось, т.е. отсылку всех компонентов правила к одному и тому же случаю (экземпляру популяции).

Заключенное в утверждении условной независимости обобщающее свойство можно явно выразить формально-логически кванторами всеобщности по экземплярам популяции, которые отражают возможность делать прогноз целевого атрибута для неизвестного (нового) экземпляра популяции. (Ясно, что для экземпляров из имеющейся БД такой прогноз является огрублением, ибо каждый известный экземпляр имеет однозначное значение всех атрибутов, а не распределение вероятностей. Но квазиправило и тут работает корректно: оно дает ту часть информации о целевом атрибуте экземпляра, которую можно объяснить.)

Помимо этого квантора всеобщности, перед формулой утверждения можно поставить также кванторы всеобщности по значениям атрибутов, как это показано в варианте вероятностной логики первого порядка Ф. Бэхуса [21]. Иными словами, для заданного целевого атрибута R все левые части комплекта квазиправил вида (2) имеют одинаковый формат. (Разумеется, выведенное описание такого вида содержит элемент открытия только тогда, когда в левые части квазиправил входят не все не-потомки атрибута R. Т.е. множество родителей не совпадает с множеством не-потомков.)

И формула (1) тоже имеет неизменный формат для всех состояний узлов сети. Знания, заключенные в байесовской сети, можно представить в аппарате хорновских предложений. Понятие *предметная переменная* в аппарате БС требует пояснения, ибо не всегда соответствует обычному в логике представлению (как о неком объекте). Например, может быть выведена модель, содержащая, в частности, такую зависимость: если идет дождь, то выручка магазина падает. В этом утверждении неназванной предметной переменной выступают пространственно-временные координаты (город и день).

Уже по самому формату (3) записей данных понятно, что здесь нельзя переносить утверждения с одного объекта на следующий (равно как и на любой другой), т.е. не работает принцип математической индукции. Аналогично принципу математической индукции можно было бы оформить принцип статистической индукции (рис. 4).

Мы рассмотрели внешнюю форму процесса, названного обобщением. Проанализируем модель по существу, как результат обобщения. Этим результатом в нашем случае является граф G, аннотированный соответствующими параметрами ϑ. Такая модель — форма представления совместного распределения вероятностей. Можно было бы описать это совместное распределение вероятностей просто в виде таблицы. Может показаться, что мы выполнили просто компиляционную обработку данных (и степень обобщения очень слабая). Напомним, что байесовская сеть выражает совокупность отношений условной независимости. Если полученный граф полносвязный, то можно констатировать, что отношений условной независимости не выявлено. Т.е. открытия не произошло, и более того, показано, что соответствующих закономерностей в предметной области нет. (Или они намного сложнее, другого характера и требуют другого типа выражения.) Если же отношения условной независимости обнаружены, значит, получены определенные знания. Суть обнаруженных знаний (семантику выявленных отношений условной независимости) рассмотрим ниже. А сейчас констатируем, что индукции в классическом смысле не произошло, нет стандартной схемы индукции (нет введения в логическое выражение квантора всеобщности по обычной переменной). Более того, вывести отношение (не)зависимости по схеме математической индукции невозможно, так как это отношение невозможно выразить на уровне одного экземпляра (случая).

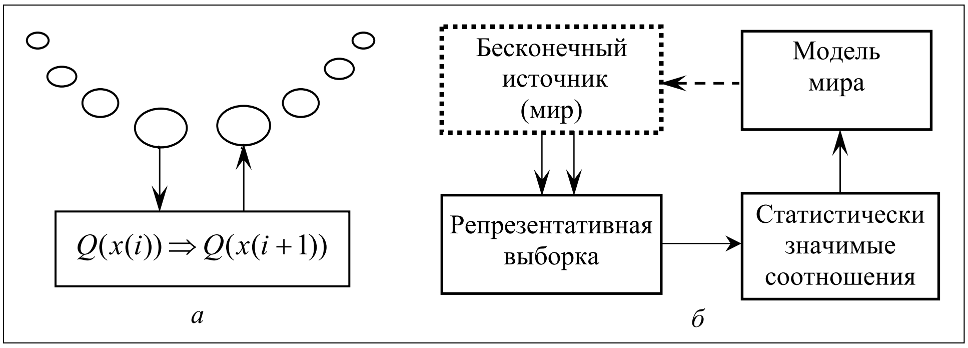

Рис. 4. Схемы индукции: а — математическая индукция; *б* — статистическая

Классическая схема индукции сработает при детерминистических отношениях. Тогда одна запись может выступать примером общего отношения. Но в этом случае не ясно, что надо выводить. Далее для краткости вместо $A=a_i$ будем писать просто A_i. Например, пусть имеем запись $\langle A_1, B_2, C_3 \rangle$. Без задания определенных предварительных ограничений (априорных знаний) на базе этой записи как примера можно было бы построить много гипотетических правил, в частности:

$$A_1 \Rightarrow B_2,\ B_2 \Rightarrow A_1,\ A_1 \Leftrightarrow B_2,\ C_3 \Rightarrow B_2,$$

$$A_1 \,\&\, B_2 \Rightarrow C_3,\ A_1 \,\&\, C_3 \Rightarrow B_2,\ B_2 \,\&\, C_3 \Rightarrow A_1 \text{ и т.д.}$$

Уместно говорить о различии примеров и гипотез, но для нас важнее сейчас другие (методологические) аспекты. Кстати, для примера $\langle A_1, B_2, C_3 \rangle$ самой естественной гипотезой будет «всегда $A_1 \,\&\, B_2 \,\&\, C_3$». Но таким образом гипотезу не формируют ввиду очевидной наивности. При канонической схеме индукции форма утверждения предопределяется нашими знаниями. В индуктивном логическом программировании схема несколько другая: гипотеза явно не задана, но заданы «строительные блоки» для конструирования модели (индуктивного утверждения). При этом данные распадаются на положительные и отрицательные примеры, что предопределено чисто логической формой утверждения и детерминистичностью отношений.

При идентификации (восстановлении) байесовской сети из данных ситуация другая. На входе не задается никакой конкретной информации об отношениях между элементами, за исключением «рамочных» ограничений (ацикличность схемы, отсутствие дублирующих дуг и т.д.). Для более ясного понимания сути вопроса целесообразно рассмотреть его в контексте родственных задач.

ОТ ДОКАЗАТЕЛЬСТВА К ОТКРЫТИЮ. СХЕМЫ ФОРМИРОВАНИЯ ЗНАНИЙ «СНИЗУ–ВВЕРХ»

Сопоставление разных постановок задач и «схем вывода» знаний в формальной системе можно выразить схемой (рис. 5). Результат доказательства по уровню общности не выше, чем аксиомы на входе, и это есть вывод «сверху-вниз». Такой результат в принципе потенциально уже содержится в аксиомах на входе. Доказательство лишь разворачивает сжатую информацию. В эмпирических исследованиях просматривается внешне похожая схема, но содержание блоков там другое, а место доказательства в ней занимает экспериментальная проверка (верификация). Для нас здесь важен вопрос: откуда берется гипотеза? В научном фольклоре под индуктивным познанием часто понимают как раз формирование правдоподобной гипотезы на основе данных, а не ее проверку или доказательство.

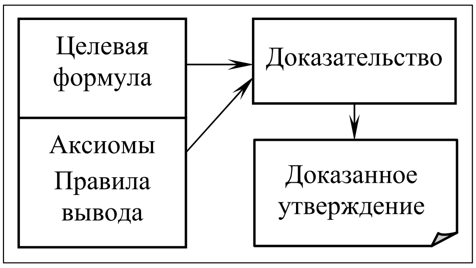

Рис. 5. Дедукция

Формирование правил (формул или концепций) в ILP показано на рис. 6. Формула правила (или концепции) на входе не задана, но указан соответствующий атрибут в примерах или разметка примеров (целевая переменная). Разбиение на положительные и отрицательные примеры соответствует случаю, когда формируется импликативное правило с логической формулой в правой части или когда строится бинарный классификатор.

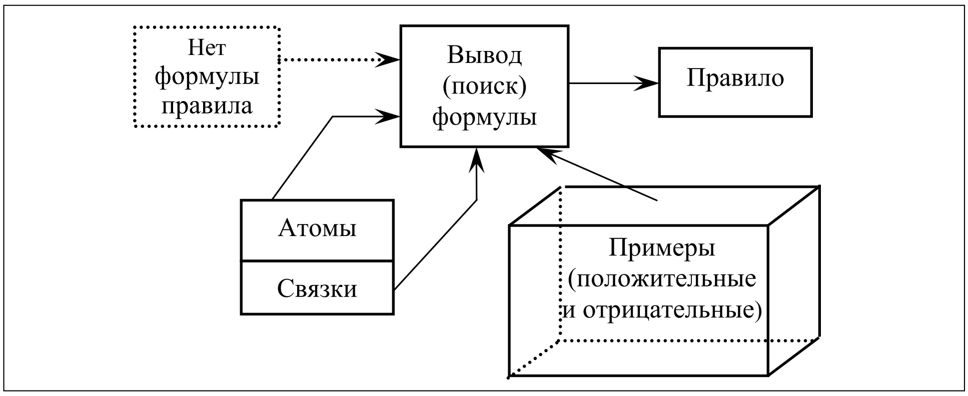

Рис. 6. Вывод в ILP

Обнаружение множества функциональных зависимостей в БД с дискретными атрибутами выполняется по схеме рис. 7. Несмотря на детерминистический характер, эта задача является NP-сложной (иногда она даже сложнее аналогичной задачи в вероятностных условиях).

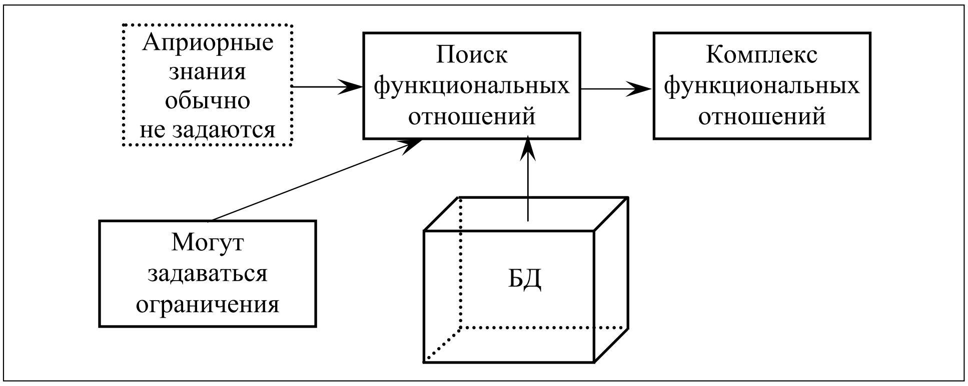

Рис. 7. Вывод функциональных отношений в БД

При идентификации АОГ-модели гипотезы формируются автоматически в ходе поиска, причем мы не указываем зависимую (целевую) переменную, т.е. не предопределяем, что есть фактор, а что — эффект (отклик). (Кстати, дискриминировать факторы и эффекты часто не удается именно в ситуации, когда связи детерминистические. Аналогичная проблема может возникнуть при формировании базового набора функциональных отношений. Ведь суперпозиция функций тоже функция.)

Для практики более существенна следующая особенность задачи. Обычно оказывается, что в наблюдаемой части мира (в БД) нет ни одного строго функционального отношения. Например, в той части среды, которая доступна роботу или программному агенту, не наблюдается ни одной детерминистической связи. Более того, в БД (дискретных) могут присутствовать все мыслимые комбинации значений (хотя и с разной повторяемостью). Именно с такой ситуацией работает методология идентификации графовых вероятностных моделей зависимостей. Практически в каждой предметной области можно легко выявить связи, но не детерминистические, а более тонкие, такие, какими они выглядят после вычленения подсистемы из системы. Часто эти связи можно выразить отношениями вероятностной условной независимости. Выявление таких отношений требует намного бóльших объемов данных, но зато эти отношения значительно робастнее, ибо имеют статистическую значимость. Даже один новый пример может опровергнуть выведенное в ILP импликативное правило. Для модели БС это исключено.

Восстановление АОГ-моделей из статистических данных показано на рис. 8. Можно заметить внешнюю аналогию между задачей идентификации БС и выводом набора базовых теорем в формальной системе. Наиболее известные алгоритмы вывода структуры АОГ-модели из данных — алгоритмы PC, FCI, TPDA [2,3]. Именно PC использован для построения упомянутой во введении модели причин и следствий бедности в развивающихся странах [6].

БС можно назвать квазитеорией предметной области, описывающей закономерности поведения объектов (индивидов) в форме влияния одних свойств объекта (процесса) на другие, причем всем связям и закономерностям приписывается непрерывная мера истинности. АОГ-модель является не научной, а феноменологической теорией. Ее можно рассматривать как множество классификационных и регрессионных моделей, естественным образом интегрированых в систему. Это гибкий инструмент вывода, связывающий множество свидетельств и следствий (в роли которых может выступать каждый атрибут).

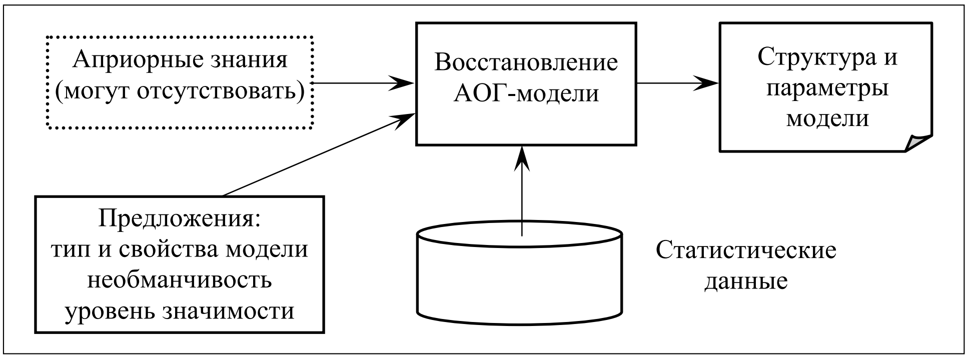

Рис. 8. Реконструкция графовой вероятностной модели зависимостей

Практически утверждения условной независимости можно рассматривать как правила для упрощения выполнения запроса к БД (сокращения условия поиска). При этом ответом на запрос будет распределение значений для указанной переменной (однозначного детерминистического ответа нет). БС обеспечивает вычислительные преимущества при отработке запроса к базе. Смысл этой задачи — пассивное интерполяционное предсказание новой записи данных, т.е. заполнение пропусков в запросе по сопутствующим компонентам запроса (с использованием формулы Байеса). Это обобщенный запрос к БД. Выполнение такого запроса можно было бы делать и по таблице совместного распределения вероятностей, но это повлекло бы избыточное расходование вычислительных ресурсов.

Выше формулой (2) описан фрагмент модели. Композиция всех фрагментов в формуле (1) оправдывается отношением условной независимости. Ребра модели являются каналами передачи информации. Вывод в соответствии со структурой модели сокращает вычисления благодаря равенству соответствующих условных вероятностей. Заметим, что квазиправила вида (2) нельзя считать полными правилами предсказания (вывода) по свидетельствам. Они обеспечивают вычисление безусловной совместной вероятности всех переменных в процессе перемещения по дугам от корневых вершин к потомкам. Но эффекты (следствия), когда эти правила выступают как свидетельства, также несут информацию о своих причинах. Поэтому, когда решается задача вывода по свидетельствам (т.е. вывода условной вероятности), информация передается от свидетельств к переменной запроса в любом направлении, в том числе и против стрелок орграфа G. (В диагностических задачах вывод идет как раз от следствий к причинам.)

Таким образом, для вывода значения переменной запроса R следует учитывать не только «родителей» переменной R, но и ее «детей». Заметим, что информация может передаваться также от одного корня к другому, если имеется свидетельство об общем их потомке. Это феномен «провоцированной зависимости» [19]. В общем случае вопрос информативности свидетельств определяется отношением условной независимости, согласно критерию d-сепарации [1,4,19]. Заметим, что процесс вывода от свидетельств продвигается не по отдельным ребрам графа, а по его сечениям, фронтом. Для этого граф преобразуется в дерево юнкций.

Прослеживается аналогия между ролью условной независимости для вероятностного вывода и ролью отношения условной нерелевантности в строго логическом выводе [15].

Условная независимость (нерелевантность) в логической системе определяется так [15].

В домене Ω множество утверждений $\mathbf{Y}$ условно независимо от множества утверждений $\mathbf{X}$ при справедливости известных утверждений $\mathbf{Z}$, если и только если выполняется

$$\Omega, \mathbf{Z}, \mathbf{X} = \left|\mathbf{Y} \Leftrightarrow \Omega, \mathbf{Z}\right| = \mathbf{Y}.$$

Чтобы увидеть аналогию, можно выразить вероятностную условную независимость $p(Y \mid Z, X) = p(Y \mid Z)$ через Шенноновскую взаимную информацию, а именно, как $\operatorname{Inf}(Y, (Z X)) = \operatorname{Inf}(Y, Z)$. Таким образом, в обоих случаях выражена избыточность информации X (логической либо Шенноновской).

Таблица совместного распределения вероятностей пригодна для решения задач только первого рода (вывод неизвестных значений в «пассивной» постановке). Задача формулируется так: каким правдоподобно должно быть значение атрибута C у экземпляра популяции x, если известно, что значения атрибутов A и B у этого экземпляра x известны. Смысл задачи — правдоподобное заполнение пропусков в БД. Если же идентифицируется структура вероятностной модели зависимостей с точностью до ориентации дуг (направленности стрелок) графа, то появляется возможность решать качественно иную задачу.

Принципиально другая задача — предсказание активного типа, которое становится возможным благодаря восстановленным строго ориентированным путям в графе модели. Вывод в «активной» постановке (задача второго рода) соответствует рассуждению: что было бы, если бы в тех же статистических условиях среды агент придал бы (принудительно установил извне) указанному атрибуту экземпляра популяции k-е значение? При этом наши суждения выходят за пределы БД и относятся к миру (среде, из которой получены данные). Более того, мы прогнозируем события в частично (локально) измененном мире: каким он будет после выполнения планируемого воздействия на объект. (Напомним, что сама модель обычно выводится из данных, собранных как пассивные наблюдения за объектом.)

Прогноз последствий действий предполагает знание каузального характера связей. Действительно, пусть в результате вывода мы получили модель, где вершины A и B непосредственно ассоциированы, т.е. соединены ребром, но истинная ориентация этого ребра неизвестна. Этот недостаток знаний не мешает нам решать задачу пассивного предсказания значений переменной *A*, исходя из значения переменной B. Другое дело, когда мы хотим спрогнозировать значение B после активного воздействия на объект, которое описывается установкой переменной A в значение a_0. Если каузальная связь в объекте адекватно описывается дугой $A \to B$ модели, то прогноз воздействия будет точно такой, как при пассивном прогнозе, а именно $p(B \mid A = a_0)$. Если же каузальная связь выглядит как $A \leftarrow B$, то активное воздействие на A не изменит значения переменной *B* (т. е. оно по-прежнему будет описываться тем же самым маргинальным распределением). Технически можно представить, что когда мы извне устанавливаем значение переменной A, то все дуги модели, входящие в узел A, удаляются [4]. Возвращаясь к примеру (см. рис. 1), можно представить ситуацию, что перед школой вылили цистерну воды. Ясно, что после такого внешнего вмешательства следует проигнорировать связь между мокрой мостовой и дождем, и не нужно искать причину того, что школьник поскользнулся среди предшествующих узлов модели.

После удаления дуг, входящих в манипулированный узел, вывод последствий вмешательства для других переменных (если нет свидетельств) выполняется в соответствии с квазиправилами вида (2).

С целью большей наглядности многие авторы предлагают явно включать в граф модели переменные, отображающие внешние вмешательства или решения человека. Полученную модель называют диаграммой влияния.

СУЩНОСТЬ ЗНАНИЙ, ОТКРЫТЫХ В ФОРМЕ БС

Локальный фрагмент БС (семейство вершин графа) содержит два типа информации: 1) качественные знания, выражаемые дугами вида $A \to B$, и 2) количественная информация (параметры), привязанная к семейству вершин. Поскольку на входе алгоритма идентификации БС не задавалось ничего конкретного о связях и отношениях, то уместно говорить об открытии. Получены важные знания о предмете, которых в явном виде нет ни в одном отдельно взятом случае. И знания эти могут касаться даже каузальных отношений.

Ни в каком отдельно взятом примере (случае) не содержится явная информация о связях и влияниях (разумеется, исключая вульгарное понимание связи как видимого на фотографии механического коннектора, рычага). Информация о влиянии проявляется только через поведение. По одной мгновенной фотографии нельзя восстановить знания о влиянии одних объектов на другие. А вот большое их число делает это возможным. О влиянии одного объекта на другой можно было бы судить по малому числу примеров, полученных в ходе активного эксперимента. Но и в этой ситуации одна запись содержит настолько мало информации, что ее нельзя выразить словесно или отобразить формулой. Наша постановка задачи обобщения сложнее, ибо мы отталкиваемся от данных пассивных наблюдений. Возвращаясь к сопоставлению с логикой, видим, что БС дает верные конфигурации правил вида $A = a_1 \; \& \; B = b_2 \Rightarrow$ целевой атрибут C. При этом A и B должны быть непосредственными причинами для C, т.е. родителями в графе. Однако правую часть вместо логического выражения составляет распределение вероятностей.

Отношение условной независимости содержит знание, даже когда нам не удается довести процесс идентификации модели до ориентации дуг. С точки зрения обработки запросов к данным это есть знание об информационной нерелевантности (избыточности) одних атрибутов относительно других при задании третьих. С точки зрения отображения объекта (источника данных) — это знание об отсутствии непосредственной связи соответствующих процессов и явлений.

Какие принципиальные соображения обосновывают возможность рассуждений по схеме: каким было бы значение атрибута C, если бы в заданных условиях агент придал атрибутам A и B указанные значения? Ответ кроется в обоснованности идентификации структуры БС и общенаучном принципе предпочтения самой простой модели среди равно адекватных альтернатив [4, 18, 22]. Этот принцип минимальности модели — проявление общего принципа «Бритвы Оккама», широко известного в очень отвлеченной формулировке, и поэтому требующего аккуратной интерпретации [22]. Другим фундаментальным положением для восстановления структуры из данных служит предположение необманчивости распределения вероятностей переменных АОГ-модели [3, 22], которое можно выразить как следующую импликацию: из вероятностной условной независимости между переменными следует отсутствие путей в АОГ-модели, поддерживающих ассоциацию этих переменных, что в обозначениях [19] записывается (через критерий d-сепарации) как

$$\Pr(A \perp \mathbf{S} \perp B) \Rightarrow Ds(A \perp \mathbf{S} \perp B)\,.$$

Это обратная импликация по сравнению с каузальным марковским допущением для АОГ-модели [2–4]. При пустом условии $\mathbf{S}$ получаем упрощенную (безусловную) форму предположения необманчивости: отсутствие ассоциированности переменных означает отсутствие цепи передачи зависимости между этими переменными в орграфе модели [19]. Предположение необманчивости и каузальное марковское допущение обеспечивают взаимно-однозначное соответствие структуры и свойств модели (структурно-поведенческий изоморфизм).

Процесс идентификации АОГ-модели из данных заключается в построении самой простой модели среди адекватных. Такая модель компактно представляет и объясняет все данные. Обоснованность утверждений условной независимости заключается в статистической значимости соответствующих тестов [3, 19]. Конструирование модели по результатам тестов выглядит более аргументированным выводом, чем максимизация правдоподобия (когда необходимо корректировать правдоподобие с помощью штрафа за сложность). Если процесс генерации рассматриваемых данных проще всего описать именно такой АОГ-моделью, то естественно предположить, что в реальном мире, откуда взяты данные, между соответствующими явлениями (сущностями) действует система механизмов, изоморфная этой модели. Такой гносеологический принцип принят во всех естественных науках. Ясно, что и в других случаях экстраполяция прогнозов, полученных из модели, на будущее (на новые случаи) предполагает стабильность системы в тех аспектах, которые отражены моделью.

Из данных идентифицируются ориентации не всех дуг. А все индуктивно ориентированные дуги можно разделить на два (или более) сорта согласно степени обоснованности их ориентации. В работе [18] эти сорта дуг условно названы каузальными и субкаузальными. Это разделение можно пояснить с помощью скрытых переменных. Соответственно сорту дуг меняются основания для выполнения задач вывода по свидетельствам. Идентифицировать каузальные связи в модели, выведенной из наблюдений, в случае скрытых переменных можно только при специальном паттерне соседних связей в модели [4, 18]. Этот паттерн отображается графически как Y -образная конфигурация. При этом говорят, что мокрая мостовая — причина того, что школьник поскользнулся (см. рис. 1).

ЛОГИЧЕСКИЕ РАСШИРЕНИЯ АППАРАТА БС

Одно из существенных ограничений возможностей АОГ-моделей следует из предположения, что данные — это *i.i.d.*-выборка. Тем самым предопределяется: вся информация о связях заключена внутри одной и той же записи. Расширение АОГ-моделей путем прямолинейного введения времени порождает резкое возрастание размерности, что делает вывод структуры из данных практически нереализуемым. Поэтому для динамических байесовских сетей [23] обычно задаются структурные ограничения, например, через ограничение лага времени (или же структура задается априори).

БС часто характеризуются в литературе как аналог пропозиционального исчисления (хотя это не совсем точно). Этим подчеркивают ограниченность их возможностей. Кстати, БС можно представить как точную теорию в неопределенном контексте [24], т.е. как локальный связный фрагмент полной функциональной модели, отображающий зависимости представленных элементов.

Последнее время активно исследуются варианты расширения БС до уровня, соответствующего логике первого порядка. Логика является одним из самых универсальных языков описания предметных областей и пригодна к расширению и пополнению. Поэтому любую БС, в принципе, можно расширить и пополнить до описания, в котором будет использоваться не только логика первого порядка, но и более богатые языки (например, отражающие время и скорость процессов).

Часто смежные вершины БС соответствуют характеристикам одного и того же реального объекта (предметной переменной), в частности, пациента в медицинской БД. В других моделях смежные вершины (и компоненты одной записи данных) могут соответствовать разным индивидам, например, родственникам в моделях наследственности и генеалогических процессов. Во многих моделях связующей осью каждого рассуждения выступает время и место (как в приведенном выше примере с дождем). А в примере с историей болезни пациента время и место изменяются (согласно биографии пациента), и это не является существенным препятствием для вывода обобщений. В любом случае все записи данных должны быть получены согласно единой схеме измерения, когда между переменными (измеряемыми величинами) действуют одни и те же механизмы и закономерности. При этом данные, подаваемые на вход алгоритма вывода модели зависимостей, являются результатом редуцирования полного исходного описания. Излишняя и ненужная для индуктивного вывода информация убирается, и исходные данные редуцируются до пропозиционального представления (одной плоской таблицы). На самом деле аналитик помнит и подразумевает всю информацию, которая зафиксирована в схеме измерения, специфицирующей точки и отсчеты времени фиксации данных. (Эта схема может быть изменена значительно позже проведения фактических измерений путем предобработки и переформатирования данных.)

Исходную модель отношений, описанную с помощью предикатов, во многих случаях легко представить (вложить) в форме обычной байесовской сети с помощью схемы измерения. Например, когда предикат имеет два аргумента, причем один из них — объект, а другой — характеристика объекта, то объектную переменную выносят в схему измерения, а характеристику явно представляют как вершину БС. Однако в предметной области могут существовать более сложные отношения, которые не удается таким образом редуцировать. Тогда прибегают к логическим расширениям БС. Потребность перехода к моделям с более полным логическим описанием возникает, прежде всего, в связи с намерениями (необходимостью):

1) автоматизировать интерпретацию модели;

2) автоматизировать компиляцию модели из базы знаний (помочь в этом пользователю или эксперту) [21];

3) повторять моделирование, подбирая наиболее эффективную схему измерения.

Кроме того, попытка «загнать» данные в стандартное представление (плоскую таблицу) для алгоритмов вывода БС может привести к разрастанию массивов данных. Еще одна проблема возникает при наличии детерминистических отношений между переменными. Детерминистические отношения порождают трудности для большинства методов вывода структуры АОГ-модели.

Наиболее популярным расширением становятся реляционные БС или вероятностные реляционные модели [25]. При этом информация о каждом типе сущностей хранится в своем соответствующем отношении модели, так что одно отношение описывает все сущности данного типа. Допускаются зависимости между атрибутами одной сущности, а также (другая форма) между атрибутами разных сущностей (через «цепочки»). Исследователи отмечают ограниченность выразительных способностей вероятностных реляционных моделей [27]. Известны дальнейшие расширения этих моделей. Недавно были предложены логические БС [27]. В них определены логические и вероятностные предикаты, что обеспечивает разное представление детерминистических и вероятностных отношений. Для реляционных данных предложены также вероятностные марковские сети [28]. Однако следует заметить, что они соответствуют лог-линейным статистическим моделям, т.е. другому (не графовому) принципу представления совместного распределения вероятностей.

ЗАКЛЮЧЕНИЕ

- Класс вероятностных моделей зависимостей на основе ациклических ориентированных графов (и его расширения) берет на себя роль языка и средства коммуникации аналитиков и специалистов различных дисциплин — математиков, статистиков, экономистов, социологов, психологов, биологов, медиков, инженеров и т.д.
- Как статистические модели, которые идентифицируются на основе эмпирических данных, графовые вероятностные модели выполняют роль инструмента познания, т.е. способны дать исследователю инсайт для понимания механизмов взаимодействия в рассматриваемой предметной области или в любой системе, чье поведение подчиняется неизвестному комплексу закономерностей различной природы.
- Вероятностные модели систем зависимостей на основе ациклических ориентированных графов в компьютерном виде служат эффективным представлением знаний в условиях неопределенности и обеспечивают средство решения прогнозно-аналитических задач и механизм автоматических рассуждений для экспертных систем, агентов и роботов в разных сферах приложений.
- В целом описанный класс моделей и методов поддерживает законченный цикл компьютерных технологий $\{$измерения, наблюдения, данные$\}$ $\to$ модель $\to$ $\{$анализ, прогноз, управление$\}$.

ЛИТЕРАТУРА

1. *Pearl J.* Probabilistic reasoning in intelligent systems: networks of plausible inference. — San Mateo: Morgan Kaufmann, 1988. — 480 p.
2. *Computation,* Causation and Discovery / Eds: C. Glymour and G.F. Cooper. — Menlo Park, CA, Cambridge, MA: AAAI Press/The MIT Press, 1999. — 570 p.
3. *The TETRAD* Project: Constraint Based Aids to Causal Model Specification / R. Scheines, P. Spirtes, C. Glymour et al. // Multivariate Behavioral Research. — 1998. — **33**, № 1. — P. 65–118.
4. *Pearl J.* Causality: Models, Reasoning and Inference. — Cambridge Univ. Press, 2000. — 526 p.
5. *A theory* of causal learning in children: Causal maps and Bayes nets / A. Gopnik, C. Glymour, D.M. Sobel // Psychological Review. — 2004 — **111**, № 1. — P. 1–30.
6. *Bessler D.A.* On World Poverty: Its Causes and Effects. — Food and Agricultural Organization (FAO) of the United Nations Research Bulletin. — Rome. — 2003. — 50 p.
7. *Sebastiani P., Abad M., Ramoni M.F.* Bayesian Networks for Genomic Analysis / In: Genomic Signal Processing and Statistics. — EURASIP Book Series on Signal Processing and Communications. — 2005. — P. 281–320.
8. *Андон Ф.И., Балабанов А.С.* Выявление знаний и изыскания в базах данных: подходы, модели, методы и системы (обзор) // Проблемы программирования. Материалы 2-й междун. конф. «УкрПРОГ’2000». — 2000. — № 1–2. — С. 513–526.
9. *Балабанов А.С.* Выделение знаний из баз данных — передовые компьютерные технологии интеллектуального анализа данных // Математ. машины и системы. — 2001. — № 1/2. — С. 40–54.

10. *Балабанов О.С.* Комп'ютерний інтелект: можливості і реальність // Вісн. Національної Академії наук України. — 1997. — № 9–10. — С. 16–21.
11. *Гаек П., Гавранек Т.* Автоматическое образование гипотез. — М.: Наука, 1984. — 280 с.
12. *Valiant L. G.* A theory of the learnable // Communications of the ACM. — 1984. — **27**, № 11. — P. 1134–1142.
13. *Вапник В.Н., Червоненкис А.Я.* Теория распознавания образов (статистические проблемы обучения). — М.: Наука, 1974. — 416 с.
14. *Jaynes E.T.* Probability theory: the logic of science. — Cambridge Univ.Press, 2002. — 520 p.
15. *Darwiche A.* A logical notion of conditional independence: Properties and applications // Artificial Intelligence. — 1997. — **97**, № 1–2. — P. 45–82.
16. *Denecker M., Martens B., De Raedt L.* On the difference between abduction and induction: a model theoretic perspective // ECAI-96, Workshop on Abductive and Inductive Reasoning. — 1996 / Denecker M., Flach P., Kakas A., eds // Workshop report. — Budapest, Hungary. — P. 1–7.
17. *Muggleton S.* Scientific knowledge discovery using inductive logic programming // Communications of ACM. — 1999. — **42**, № 11. — P. 42–46.
18. *Балабанов О.С.* Відкриття структур залежностей в даних: від непрямих асоціацій до каузальності // Проблемы программирования. Материалы 3-й междунар. конф. «УкрПРОГ'2002». — 2002. — № 1–2. — С. 309–316.
19. *Балабанов А.С.* К выводу структур моделей вероятностных зависимостей из статистических данных // Кибернетика и системный анализ. — 2005. — № 6. — С. 19–31.
20. *Андон Ф.И., Яшунин А.Е., Резниченко В.А.* Логические модели интеллектуальных информационных систем. — Киев: Наук. думка. — 1999. — 396 с.
21. *Bacchus F.* Using First-Order Probability Logics for the Construction of Bayesian Networks / Uncertainty in Artificial Intelligence (UAI-93) // Conf. Proceedings. — Washington: Morgan Kaufmann, 1993. — P. 219–226.
22. *Domingos P.* The Role of Occam's Razor in Knowledge Discovery // Data Mining and Knowledge Discovery. — 1999. — **3**. — P. 409–425.
23. *Learning* the Structure of Dynamic Probabilistic Networks / N. Friedman, K. Murphy, S. Russell // Proceedings of the 14th Conf. on Uncertainty in Artificial Intelligence (UAI'98). — USA: Morgan Kaufmann, 1998. — P. 139–147.
24. *Halpern J., Pearl J.* Causes and explanations: A structural-model approach. Part I. Causes // Proceedings of the 17th Conf. on Uncertainty in Artificial Intelligence. — San Francisco, CA: Morgan Kaufmann, 2001. — P. 194–202.
25. *Learning* Probabilistic Relational Models / N. Friedman, L. Getoor, D. Koller, A. Pfeffer // Proceedings of the 15th Internat. Joint Conf. on Artificial Intelligence (IJCAI). — Stockholm, Sweden, 1999. — San Francisco, CA: Morgan Kaufmann, 1999. — P. 1300–1307.
26. *De Raedt L., Kerstin K.* Probabilistic logic learning // SIGKDD Explorations. — 2003. — **5**, № 1. — P. 31–48.
27. *Logical* Bayesian Networks and Their Relation to Other Probabilistic Logical Models / D. Fierens, H. Blockeel, M. Bruynooghe, J. Ramon // Lecture Notes in Computer Science (Springer, Berlin). — 2005. — **3625**. — P. 121–135.
28. *Richardson M., Domingos P.* Markov Logic Networks // Machine Learning. — 2006. — **62**, № 1–2. (Special Issue: Multi-Relational Data Mining and Statistical Relational Learning.) — P. 107–136.

Поступила 04.04.2006