УДК 007:681.3.00

А.С. БАЛАБАНОВ

ФОРМИРОВАНИЕ МИНИМАЛЬНЫХ *d*-СЕПАРАТОРОВ В СИСТЕМЕ ЗАВИСИМОСТЕЙ

Ключевые слова: *АОГ-модель, вероятностные зависимости, ациклический орграф, условная независимость, байесовская сеть, критерий d-сепарации, коллайдер, цепь, локально-минимальный сепаратор, идентификация ребер модели, подбор локально-минимального сепаратора*.

ВВЕДЕНИЕ

Анализируем свойства вероятностных моделей зависимостей, структурированных ациклическими ориентированными графами (АОГ-моделей). Различают три основные разновидности АОГ-моделей: байесовские, гауссовы и гибридные сети. В байесовских сетях переменные — номинальные (дискретные), а в гауссовых сетях — действительные (зависимости — линейные). В гибридных сетях используются оба типа переменных. Результаты статьи применимы ко всем разновидностям АОГ-моделей, поскольку базируются только на общих свойствах этих моделей — ацикличности орграфа и критерии *d*-сепарации.

Вероятностные модели зависимостей на основе графов — актуальная тема исследований на стыке многомерного статистического анализа, аппарата графов, теории информации и искусственного интеллекта. Графовые модели зависимостей — строгий язык представления знаний о зависимостях в условиях неполноты и неопределенности. Наиболее популярными являются АОГ-модели. К достоинствам АОГ-моделей относятся компактное представление систем зависимостей, способность отображать причинно-следственные связи, тестируемость любого фрагмента модели и вычислительная эффективность вероятностного вывода от свидетельств (в вероятностных экспертных системах) [1–6]. АОГ-модели применяются в решении таких задач, как медицинская и техническая диагностика, распознавание речи, прогнозирование последствий решений и действий, анализ данных в эконометрике, социометрике, микробиологии и т.д.

Для решения познавательно-исследовательских и прогнозно-аналитических задач необходимо выводить структуру модели на основе статистической выборки данных наблюдений за моделируемой системой. Задача вывода АОГ-модели из данных (или статистических отношений) в общем случае — комбинаторно сложная, что ограничивает практические применения уже при числе переменных модели порядка нескольких десятков. Поэтому на практике аналитик часто вынужден прибегать к жестким предположениям и ограничениям, что может привести к потере истинной модели. Но даже если удалось избежать ошибочных предположений, для вывода адекватной модели необходимо обеспечить надежность процедур.

Проблема поиска и подбора сепараторов в АОГ-модели возникает при выводе модели методами так называемого constraint-based подхода [2, 3, 6–8]. Этот подход состоит в идентификации (эмпирических) сепараторов или установлении фактов их отсутствия. Поиск сепараторов в сложных структурах — переборная задача, особенно трудная случае дискретных моделей. Отметим, что на момент поиска сепаратора структура модели (почти) неизвестна. (Более того, часто неизвестен и темпоральный порядок переменных.) Поэтому методы анализа графа модели (в частности, предложенные в [9]) неприменимы. Не исчерпывают проблемы и другие известные результаты, изложенные, например, в [10]. Желательно находить минимальные сепараторы и как можно раньше выявлять факты отсутствия искомых сепараторов. В настоящей статье показаны новые возможности в поиске (подборе) сепараторов в АОГ-моделях и средства повышения эффективности вывода модели. Ужесточив требования к членам (элементам) сепараторов, можно отсеять большее число кандидатов в сепаратор и тем самым упростить задачу. Взаимосвязи между составами «сопряженных» («соприкасающихся») минимальных сепараторов позволяют направлять и (адаптивно) оптимизировать поиск сложных минимальных сепараторов, исходя из знания уже найденных в «окрестности» простых сепараторов и паттернов зависимостей. Оригинальность предложенных решений вытекает из идеи взаимодействия процессов поиска сепараторов. Покажем, что можно узко направить (сфокусировать) поиск сепаратора для заданной пары переменных, установив, в частности, что одна из переменных заданной пары служит сепаратором для второй и некоторой третьей переменных. Разработанные положения дополняют и расширяют результаты работ [6, 11], где предложен и обоснован ряд требований к членам локально-минимальных сепараторов и признаки принадлежности (непринадлежности) некоторой вершины к искомому локально-минимальному сепаратору.

В разд. 1 излагаются базовые результаты в графовом аппарате. В разд. 2 рассмотрены принципы объединения установленных правил в алгоритм вывода модели. Вопросы приложения результатов к анализу данных и подходящие версии предположения необманчивости рассмотрены в разд. 3.

1. ПРАВИЛА ФОРМИРОВАНИЯ ЛОКАЛЬНО-МИНИМАЛЬНЫХ СЕПАРАТОРОВ В АОГ

АОГ-модель определяется как (G,θ), где G — ациклический ориентированный граф, а θ — совокупность локально заданных параметров (условных распределений зависимых переменных) [1–4]. Каждой вершине в G соответствует переменная. АОГ-модель определяет совместное распределение вероятностей переменных. Ввиду взаимно однозначного соответствия термины «переменная» (модели) и «вершина» (графа) употребляются как взаимозаменяемые согласно контексту. (Терминология и обозначения взяты из работы [6].)

Напомним необходимые графовые понятия. Если в орграфе G есть дуга $x \to y$, то вершина x называется родителем вершины y, а вершина y называется ребенком вершины x. Ребро — это дуга, ориентация которой неизвестна или игнорируется. Вершины x и y называются смежными, если они соединены ребром. Ребро обозначается как $x — y$, а его отсутствие — как $\neg(x — y)$. Путь в орграфе — это последовательность смежных ребер без повторения вершин. Как исключение, первая и последняя вершины пути (из трех или более ребер) могут совпадать, и тогда этот путь называют циклом. Орпуть (т.е. строго ориентированный путь) — это путь, на котором все ребра ориентированы в направлении одного и того же конца пути $(x \to \ldots \to y)$. Ациклический ориентированный граф — это орграф, в котором не существует ориентированных циклов. Далее под графом будем подразумевать АОГ.

Определение 1. Коллайдером (коллизором) в графе называется фрагмент вида $x \to y \leftarrow z$. Если коллайдер $x \to y \leftarrow z$ является частью пути π в орграфе, то y называется коллайдерной вершиной на пути π. Бесколлайдерный (безколлизорный) путь (или цепь) в орграфе — это путь, не содержащий ни одного коллайдера.

Ключевую роль в теории АОГ-моделей играет критерий d-сепарации [1, 3, 4, 6]. В формулировке критерия d-сепарации будем использовать термин «кондиционирование» (введение в условие). В оригинале определения используется конструкция given $\mathbf{Z}$. Этой операции на графе соответствует одноименная операция над данными (или над распределением вероятностей). Неформально это означает фиксацию состояния, причем под фиксацией понимается не вмешательство в моделируемый объект, а взятие подвыборки данных, кондиционированной по переменным $\mathbf{Z}$. Когда по контексту роль множества $\mathbf{Z}$ понятна, можно просто сказать «с помощью (множества вершин) $\mathbf{Z}$».

Определение 2 (d-сепарация). Путь π в АОГ-модели называют d-закрытым (d-блокированным) с помощью (кондиционирования) множества вершин $\mathbf{Z}$, если и только если существует вершина x, $x \in \mathbf{Z}$, $x \in \pi$ (причем на пути π имеется дуга $x \rightarrow$ либо $\leftarrow x$) или на пути π лежит хотя бы один коллайдер $\rightarrow y \leftarrow$ (причем $y \notin \mathbf{Z}$ и не существует никакой вершины $w \in \mathbf{Z}$ такой, что существует орпуть $y \rightarrow \cdots \rightarrow w$). Множество вершин $\mathbf{Z}$ d-сепарирует вершины x и y, если и только если все пути между вершинами x и y являются d-закрытыми с помощью множества вершин $\mathbf{Z}$. Будем обозначать такую d-сепарацию предикатом $\mathrm{Ds}\,(x \perp \mathbf{Z} \perp y)$. Когда $\mathbf{Z} = \varnothing$, то запишем такой предикат как $\mathrm{Ds}\,(x \perp \perp y)$. Если хотя бы один путь между x и y не является d-закрытым с помощью $\mathbf{Z}$, то говорят, что вершины x и y d-соединены (d-зависимы). Это записывается как $\neg \mathrm{Ds}\,(x \perp \mathbf{Z} \perp y)$.

Определение 3 (сепаратор). Если в АОГ предикат $\mathrm{Ds}\,(x \perp \mathbf{Z} \perp y)$ истинен, то множество вершин $\mathbf{Z}$ называют (графовым) сепаратором (или d-сепаратором) для пары (x, y).

Графовые модели можно идентифицировать на основе d-зависимостей (d-сепараций). Первичное свойство сепарации в АОГ можно выразить как

$$(x - y) \Leftrightarrow \forall\, \mathbf{Z}(x, y \notin \mathbf{Z}) : \neg \mathrm{Ds}\,(x \perp \mathbf{Z} \perp y). \tag{1}$$

Таким образом, идентификация ребер связана с идентификацией сепараторов.

В графовой модели (в том числе и более общей, чем АОГ) существование ребра $x - y$ означает, что отсутствует сепаратор для (x, y). Обратная импликация верна для АОГ, но не верна (в общем случае) в расширенных моделях, отражающих влияние скрытых переменных, и в моделях с орциклами.

Определение 4 [6, 12]. Локально-минимальным сепаратором для пары вершин (x, y) в АОГ называется такой сепаратор $\mathbf{Z}$, при удалении из которого любого его члена полученное множество вершин не будет сепаратором для (x, y). Формально это записывается так:

$$\mathrm{Ds}\,(x \perp \mathbf{Z} \perp y);\ \ \forall\, w \in \mathbf{Z} : \neg \mathrm{Ds}\,(x \perp \mathbf{Z} \setminus \{w\} \perp y).$$

Учитывая следствие 2 из [10], можно сформулировать эквивалентное определение: сепаратор $\mathbf{Z}$ для пары вершин (x, y) является локально-минимальным, если для всех $\mathbf{Z}' \subset \mathbf{Z}$ имеем $\neg \mathrm{Ds}\,(x \perp \mathbf{Z}' \perp y)$.

Определение 5. Назовем сепаратор $\mathbf{Z}^*$ для пары вершин (x, y) в АОГ минимальным сепаратором, если для (x, y) не существует сепаратора меньшей кардинальности, т.е. для всех других сепараторов $\mathbf{Z}$ для пары (x, y) верно $|\mathbf{Z}| \geq |\mathbf{Z}^*|$.

Обозначим минимальный и локально-минимальный сепараторы для пары вершин (x, y) соответственно $S_{\min}(x, y)$ и $S_{\mathrm{lom}}(x, y)$. Очевидно, когда $S_{\mathrm{lom}}(x, y)$ состоит из одной вершины, он является также минимальным. Но если $S_{\mathrm{lom}}(x, y)$ состоит из двух или более вершин, такая импликация неверна в общем случае. Ясно, что единственный непустой $S_{\mathrm{lom}}(x, y)$ всегда состоит из подмножества родителей вершины x или y.

Напомним некоторые свойства и факты из работ [6, 11]. Ясно, что каждый минимальный сепаратор является также и локально-минимальным. В одном и том же АОГ может быть несколько минимальных сепараторов для пары вершин (x, y). В то же время единственный сепаратор $S_{\min}(x, y)$ может не иметь ни одного общего члена с некоторым $S_{\mathrm{lom}}(x, y)$. Каждый член каждого локально-минимального сепаратора $\mathbf{Z}$ является неколлайдерной вершиной на том пути (тех путях), который он блокирует, причем этот путь (как минимум один из тех путей) не блокируется никаким другим членом $\mathbf{Z}$. В составе каждого непустого локально-минимального сепаратора для пары вершин (x, y) имеется, как минимум, одна вершина, которая лежит на некоторой цепи между вершинами x и y.

Ввиду множества ссылок на утверждения из [6] будем именовать все представленные ниже утверждения предложениями. В [6] доказано нижеследующее утверждение (под номером 7), которое было названо правилом двойного 1-отсечения (double 1-cutting). Более удобно назвать его правилом аппендикса.

Предложение 1 (правило аппендикса). Если для заданной пары вершин (x, y) существует такая вершина z, что выполняются $\mathrm{Ds}(w \perp z \perp x)$ и $\mathrm{Ds}(w \perp z \perp y)$, то вершина w не входит в состав никакого локально-минимального сепаратора для пары вершин (x, y).

Из правила, представленного в [6] как факт 13, вытекает такое следствие.

Предложение 2 (правило невовлекаемой вершины (uninvolved vertex)). Если в АОГ выполняются $\mathrm{Ds}(w \perp \perp x)$ и $\mathrm{Ds}(w \perp y \perp x)$, то $w \notin S_{\mathrm{lom}}(x, y)$.

Как известно, если для заданной пары вершин (x, y) имеется сепаратор и существуют множества $\mathbf{S}, \mathbf{R}$ такие, что $\mathrm{Ds}(w \perp \mathbf{S} \perp x)$ и $\mathrm{Ds}(w \perp \mathbf{R} \perp y)$, то существует сепаратор для пары (x, y), который не содержит вершины w. Однако исключение вершины w из числа кандидатов в сепаратор может привести к потере минимального сепаратора для пары (x, y). Более того, такая потеря возможна даже в случае, когда в роли отсекающих множеств выступают некоторые одиночные вершины, т.е. выполняется условие

$$\exists\, q, z : \mathrm{Ds}(w \perp q \perp x) \,\&\, \mathrm{Ds}(w \perp z \perp y). \tag{2}$$

На рис. 1 дан пример модели, в которой исключение вершины w из числа кандидатов в сепаратор (на основании (2)) приводит к потере минимального сепаратора. Для данной модели имеем $S_{\min}(x, y) = \{q, w, z\}$. При этом выполняются $\mathrm{Ds}(w \perp q \perp x)$ и $\mathrm{Ds}(w \perp z \perp y)$. Наилучшими среди сепараторов для пары (x, y), которые не включают вершину w, являются $\{q, r_1, r_2, \ldots, r_m\}$ и $\{z, t_1, t_2, \ldots, t_k\}$. (Если присутствуют вершины, показанные на рис. 1 пунктиром, то будут существовать еще два сепаратора, равноценные указанным.) Названные сепараторы намного сложнее минимального. Если оценить этот пример как экзотический, то в большинстве практических ситуаций, вероятно, будет целесообразно исключить вершину w из поиска сепаратора, выявив свидетельство (2).

В [6] факт 6 был дан без доказательства. Докажем это правило, несколько видоизменив формулировку. (Ниже будет дано его обобщение.)

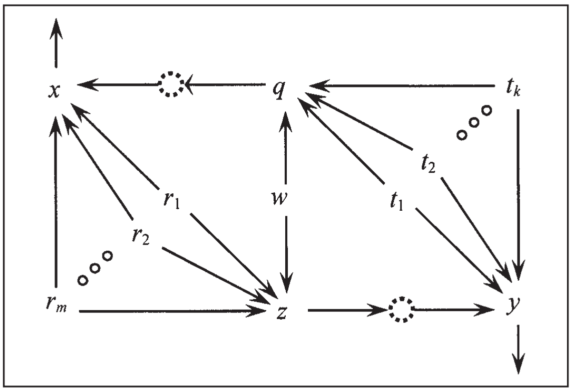

Рис. 1

Предложение 3 (правило единственной общей близкой (single common covariate)). Если в АОГ верно $\neg \mathrm{Ds}(x \perp \perp y)$ и имеется только одна вершина z, d-соединенная с обеими вершинами x и y, то либо существует ребро x — y, либо вершина z входит в состав всех сепараторов для пары вершин (x, y), и тогда если нет ни одной вершины $w\,(w \neq x, y)$, d-соединенной с вершиной z, то $\{z\}$ является единственным локально-минимальным сепаратором для (x, y).

Доказательство. Первая часть следует из правила, записанного в [6] как утверждение 2а. (Иллюстрацией служит рис. 2, *а*.) Для второй части требуется доказать, что невозможна альтернатива $\neg \mathrm{Ds}(x \perp \{z\} \perp y)$ и $\mathrm{Ds}(x \perp \mathbf{S} \perp y)$, где $\mathbf{S} \neq \{z\}$. Предположим, эта альтернатива имеет место. Из условий ясно, что существует цепь $x — z — y$ и не существует других цепей между x и y. Кондиционирование вершины z d-блокирует единственную цепь между x и y. Чтобы в результате этого кондиционирования вершины x и y остались d-соединенными, необходимо, чтобы открылся некоторый путь между x и y. При этом открываемый путь должен содержать хотя бы одну вершину, отличную от x, y, z и соединенную с вершиной z.

Однако условия предложения 3 исключают существование вершин (кроме x и y), d-соединенных с z. (Иллюстрацией служит рис. 2, *б*.) □

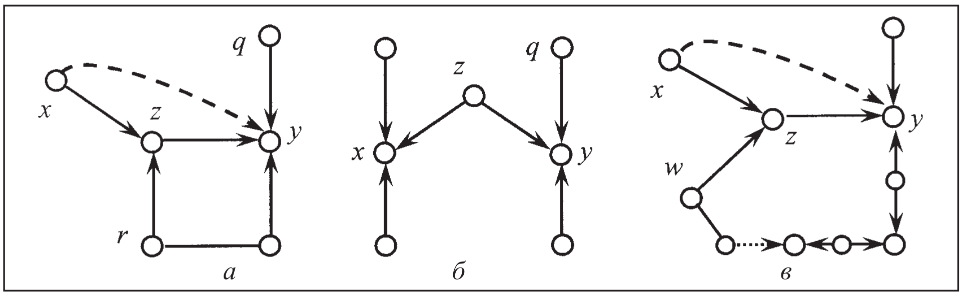

Рис. 2

Докажем предложение, представленное в [6] как факт 8, уточнив формулировку.

Предложение 4 (множество изолированных общих близких (set of isolated common covariates)). Пусть в АОГ верно $\neg\mathrm{Ds}(x\perp\perp y)$ и $\mathbf{R}$ — множество всех вершин, зависимых одновременно от x и y, т.е. $\mathbf{R}=\{r\,|\,\neg\mathrm{Ds}(r\perp\perp x)\,\&\&\,\neg\mathrm{Ds}(r\perp\perp y)\}$. Если для каждой вершины $r\in\mathbf{R}$ и всех $z\neq x, y, r$ верно $\mathrm{Ds}(r\perp\perp z)$, то либо существует ребро x — y, либо множество $\mathbf{R}$ — единственный локально-минимальный сепаратор для пары (x, y).

Доказательство. Из $\neg\mathrm{Ds}(x\perp\perp y)$ следует, что имеется ребро x — y или существует не менее одной цепи между x и y. Очевидно, если $\mathbf{R}$ — пустое, то имеется ребро x — y. Допустим, ребра x — y не существует. Ясно, что все вершины, лежащие на цепях между x и y, входят в непустое множество $\mathbf{R}$. Если бы некоторая из этих цепей состояла из трех или более ребер, т.е. содержала бы не менее двух промежуточных вершин, то эти вершины были бы взаимозависимы, что нарушило бы условие предложения 4. Следовательно, все цепи между x и y должны состоять из двух ребер. Если $|\mathbf{R}|=1$, то доказательство тривиально (предложение 3). Далее, если $|\mathbf{R}|\geq 2$, то все цепи между x и y должны иметь вид $x\leftarrow r\rightarrow y$, чтобы выполнилось условие изолированности всех промежуточных вершин r на этих цепях. Ясно, что (в условиях предложения 4) в АОГ нет ни одной дуги вида $x\rightarrow z$ или $y\rightarrow z$, поскольку такая вершина z нарушила бы условие изолированности. Поэтому в $\mathbf{R}$ не могут входить потомки x или y. Наконец, если предположить, что во множество $\mathbf{R}$ входит некоторая вершина w, лежащая на одноколлайдерном пути вида $x\leftarrow\cdots\rightarrow w\leftarrow\cdots\rightarrow y$, то в таком АОГ будет не менее двух вершин, отличных от x и y и зависимых от w, что нарушит условие предложения 4. (Заметим, что такая вершина w подпадает под условия правила изолятора (см. факт 12 из [6]).) Итак, если ребра x — y не существует, то множество $\mathbf{R}$, удовлетворяющее указанным условиям, состоит точно из вершин r, лежащих на цепях вида $x\leftarrow r\rightarrow y$. Следовательно, $\mathbf{R}$ — единственный локально-минимальный сепаратор для пары (x, y). □

Заметим, что можно легко уточнить резюмирующую часть предложения 4 с помощью следующей простой процедуры. Если для произвольно выбранной вершины $r\in\mathbf{R}$ верно $\mathrm{Ds}(r\perp x\perp y)$ или для произвольно выбранной $t\in\mathbf{R}$ верно $\mathrm{Ds}(x\perp y\perp t)$, то существует ребро x — y. Иначе если оба указанных теста дали отрицательный результат, то существует только один локально-минимальный сепаратор для пары (x, y), а именно $\mathbf{R}$.

Дополним предложение 4 следующим правилом.

Предложение 5 (изолированная общая близкая в контексте (isolated common covariate in context)). Пусть в АОГ выполняются $\neg\mathrm{Ds}(x\perp\perp y)$, $\neg\mathrm{Ds}(r\perp\perp x)$ и $\neg\mathrm{Ds}(r\perp\perp y)$. Если для вершины r и всех $z\neq x, y, r$ верно $\mathrm{Ds}(r\perp\perp z)$, то либо существует ребро x — y, либо r входит в состав каждого сепаратора для пары (x, y).

В отличие от предложения 4 теперь между вершинами x и y допускаются цепи из трех и более ребер и одноколлайдерные пути, состоящие из четырех или более ребер.

Определение 6. Назовем стержнем сепаратора (pivot of separator) для пары зависимых вершин (x, y) вершину, которая является членом хотя бы одного локально-минимального сепаратора для (x, y) и лежит на некоторой цепи между вершинами x и y.

Ясно, что вершина z на цепи $x — z — y$ обязательно является стержнем любого локально-минимального сепаратора для пары (x, y). Заметим, что некоторые из вершин, лежащих на некоторых цепях между x и y, могут не входить ни в какой $S_{\mathrm{lom}}(x, y)$. В то же время обязательным членом каждого сепаратора для (x, y) может быть вершина, не лежащая ни на какой цепи между x и y. (Пример дан в [11].)

Понятно, что в составе каждого непустого локально-минимального сепаратора для пары вершин (x, y) существует, как минимум, одна вершина, которая лежит на некоторой цепи между вершинами x и y. Следовательно [6], если имеется непустой сепаратор для пары вершин (x, y), то существует, как минимум, одна общая близкая вершина z $(z \neq x, \quad z \neq y)$, которая не отстраняется от пары (x, y), т.е. верно $\neg\mathrm{Ds}(z \perp \perp x)$, $\neg\mathrm{Ds}(z \perp \perp y)$, $\neg\mathrm{Ds}(z \perp x \perp y)$ и $\neg\mathrm{Ds}(z \perp y \perp x)$.

Определение 7. Вершина z называется потенциальным стержнем сепаратора (potential pivot of separator) для пары вершин (x, y), если верно $\neg\mathrm{Ds}(x \perp \perp y)$ и $\neg\mathrm{Ds}(z \perp \perp x) \& \neg\mathrm{Ds}(z \perp \perp y) \& \neg\mathrm{Ds}(z \perp x \perp y) \& \neg\mathrm{Ds}(z \perp y \perp x)$.

Как показано в [11] и подтверждается предложением 3, если существует только один потенциальный стержень сепаратора для (x, y), то он обязательно является членом любого сепаратора для (x, y). (Подразумевается, что ребро $x — y$ отсутствует.) Если же имеется несколько потенциальных стержней сепаратора для (x, y), то не обязательно каждый из них является членом какого-нибудь локально-минимального сепаратора для (x, y).

Определение 8. Назовем претендентом в стержни сепаратора для пары вершин (x, y) вершину, которая лежит на некоторой цепи между вершинами x и y. Ясно, что любой претендент в стержни сепаратора для пары вершин (x, y) является также и потенциальным стержнем сепаратора для (x, y).

Претендент в стержни сепаратора для пары вершин (x, y) может не быть членом никакого $S_{\mathrm{lom}}(x, y)$. Более того, существуют модели, где претендент в стержни сепаратора для (x, y), будучи одновременно родителем вершины y, не входит в состав никакого $S_{\mathrm{lom}}(x, y)$.

Можно расширить условия предложения 3 и установить следующее свойство.

Предложение 6 (простой потенциальный стержень сепаратора (simple potential pivot of separator)). Пусть в АОГ существует зависимость $\neg\mathrm{Ds}(x \perp \perp y)$ и имеется только одна вершина z, d-соединенная с обеими вершинами x и y. Тогда если для всех вершин w $(w \neq x, y)$, d-соединенных с вершиной z, верно $\neg\mathrm{Ds}(w \perp \perp y)$ и $\mathrm{Ds}(w \perp z \perp y)$, то либо существует ребро $x — y$, либо $\{z\}$ является единственным локально-минимальным сепаратором для (x, y).

Доказательство. Пусть ребра $x — y$ не существует. Тогда из $\neg\mathrm{Ds}(x \perp \perp y)$ следует, что существует цепь между x и y, причем в указанных условиях возможна единственная такая цепь и она имеет вид $x — z — y$. Кондиционирование вершины z d-блокирует единственную цепь между x и y. В результате либо получим $\mathrm{Ds}(x \perp z \perp y)$, что завершает доказательство, либо при кондиционировании вершины z открывается некоторый новый путь T между x и y. В последнем случае все коллайдеры на пути T будут открыты при кондиционировании z, и поэтому из каждой коллайдерной вершины q на пути T будет орпуть вида $q \to \cdots \to z$ (критерий d-сепарации). Возьмем одну из этих вершин q, ближайшую к вершине y. Поскольку эта q является d-соединенной с вершиной z, то согласно условиям предложения 6 должно быть $\mathrm{Ds}(q \perp z \perp y)$. Значит, между вершинами q и y нет открытых путей при кондиционировании z. Следовательно, путь T не может быть открытым при кондиционировании z. Получили противоречие. (Иллюстрация дана на рис. 2, *в*.) □

В ходе доказательства не уточнялись ориентации ребер, поскольку рассматривалась только проблема смежности. Ясно, что условия предложения 6 подразумевают независимость $\mathrm{Ds}(x \perp\perp w)$. Отсюда следуют ориентации $x \to z$ и $z \to y$ (а также $x \to y$, если это ребро существует). Заметим, что в случае невыполнения условия $\neg\mathrm{Ds}(w \perp\perp y)$ на вершины w и z распространялось бы правило чужого гена (факт 11 из [6]). Из условий предложения 6 также следует, что каждый «обходной» путь между w и z (в обход ребра $z \to y$) должен содержать не менее двух коллайдеров. Кроме того, смежными вершине x могут быть только вершина z и, возможно, y. (Но в этом смысле правило можно расширить.)

Далее расширим условия предложения 6 и получим полезный вывод. В отличие от предложения 6 не будем требовать единственности общей близкой переменной (covariate) и исключим квантор «для всех». Тогда получим следующее правило.

Предложение 7 (правило полуаппендикса). Пусть в АОГ выполняются $\neg\mathrm{Ds}(x \perp\perp y)$, $\mathrm{Ds}(w \perp\perp x)$, $\neg\mathrm{Ds}(w \perp\perp y)$ и пусть существует вершина z такая, что верно $\mathrm{Ds}(w \perp z \perp y)$. Тогда:

а) если верно $\mathrm{Ds}(w \perp z \perp x)$, то вершина w не является членом никакого локально-минимального сепаратора для (x, y);

б) если $\neg\mathrm{Ds}(w \perp z \perp x)$, то ребра x — y не существует и вершина w не является членом минимального сепаратора для (x, y).

Доказательство. Факты $\mathrm{Ds}(w \perp\perp x)$ и $\neg\mathrm{Ds}(w \perp\perp y)$ имплицируют запрет дуги $x \leftarrow y$ (и орпути $x \leftarrow \cdots \leftarrow y$). Пункт а) повторяет правило аппендикса (специальный случай дан на рис. 3, *а*). Рассмотрим пункт б). Факты $\mathrm{Ds}(w \perp z \perp y)$ и $\neg\mathrm{Ds}(w \perp z \perp x)$ имплицируют запрет дуги $x \to y$. (Следовательно, ребра x — y не существует.) Пусть $w \in S_{\mathrm{lom}}(x, y) = \mathbf{S}_w$. Идея доказательства состоит в следующем. Из $\mathrm{Ds}(w \perp\perp x)$ и $\neg\mathrm{Ds}(w \perp z \perp x)$ следует, что существуют цепи $w - \cdots \to z$ и $x - \cdots \to z$ (возможно, с общими дугами). Из $\neg\mathrm{Ds}(w \perp\perp y)$ и $\mathrm{Ds}(w \perp z \perp y)$ следует, что вершина z — неколлайдерная на всех цепях между w и y. Сопоставляя последнее с фактом наличия $w - \cdots \to z$, заключаем, что существует орпуть $z \to \ldots \to y$. Следовательно, получаем цепь (цепи) вида $x - \cdots \to z \to \cdots \to y$, которую блокирует некоторый член множества $\mathbf{S}_w$. Поэтому $\mathbf{S}_w$ не является минимальным сепаратором для (x, y). Существует меньший сепаратор $\mathbf{S}_z$, в котором вершина z замещает и вершину w, и (как минимум) еще одну вершину из состава $\mathbf{S}_w$. Две модели, иллюстрирующие пункт б) предложения 7, даны на рис. 3, *б*, *в*. □

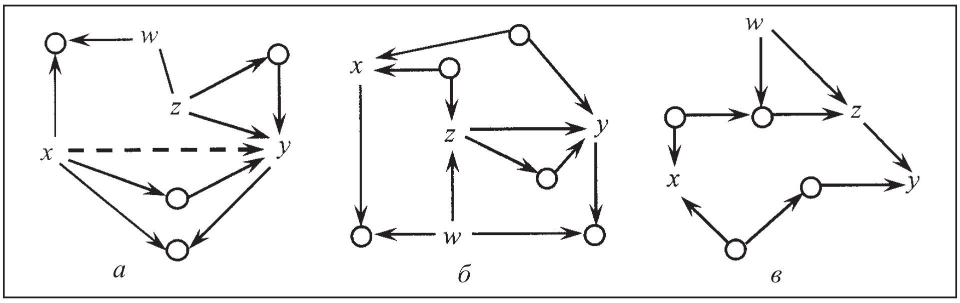

Рис. 3

Обобщим правило запрета дуги (факт 5 из [6]).

Предложение 8 (нетранзитивность первого ранга). Если в АОГ имеем $\mathrm{Ds}(w \perp z \perp x)$ и $\neg\mathrm{Ds}(w \perp z \perp y)$, где $z \neq x, y, w$, то не существует дуги $x \leftarrow y$.

Доказательство. В случае существования дуги $x \leftarrow y$ ее стыковка (splicing) с путем между вершинами w и y, открытым при кондиционировании z, дала бы открытый (при кондиционировании z) путь между вершинами w и x. Иными словами, из $\neg\mathrm{Ds}(w \perp z \perp y)$ и $x \leftarrow y$ следует $\neg\mathrm{Ds}(w \perp z \perp x)$. Но это противоречит условию. □

Заметим, что в условиях предложения 8 не существует ни одного орпути $x \leftarrow \cdots \leftarrow y$, который не проходит через вершину z. Кроме того, предложение подразумевает $\neg \mathrm{Ds}(x \perp\perp y)$ и обычно предполагается $\neg \mathrm{Ds}(x \perp\perp z)$, $\neg \mathrm{Ds}(y \perp\perp z)$. Предложение 8 можно обобщить, подставив множество вершин вместо одиночной вершины z.

Предложение 8 логически имплицирует следующее правило, являющееся обобщением утверждения 1 (правило непоглощения) из [6].

Предложение 9 (квазиинструментальная пара). Если в АОГ для некоторых разных вершин x, y, z, t, v, w (где может быть $v \equiv w$) выполняются $\mathrm{Ds}(v \perp z \perp x)$, $\neg \mathrm{Ds}(v \perp z \perp y)$, $\neg \mathrm{Ds}(w \perp t \perp x)$ и $\mathrm{Ds}(w \perp t \perp y)$, то не существует ребра $x — y$.

В дополнение к сформулированным условиям подразумевается $\neg \mathrm{Ds}(x \perp\perp y)$, $\neg \mathrm{Ds}(x \perp\perp z)$, $\neg \mathrm{Ds}(y \perp\perp z)$, $\neg \mathrm{Ds}(x \perp\perp t)$ и $\neg \mathrm{Ds}(y \perp\perp t)$. В условиях предложения не существует никаких орпутей между вершинами x и y, которые не проходят через вершину z или t. Более того, если существует цепь между x и y в обход вершин z и t, тогда все цепи между x и y (в том числе цепи, которые проходят через z и t) имеют вид $x \leftarrow \cdots \rightarrow y$. Кроме того, будет выполнятся $\neg \mathrm{Ds}(x \perp y \perp z)$ и $\neg \mathrm{Ds}(z \perp x \perp y)$.

Уточним, что в указанных условиях существуют три альтернативы:

а) вершина w не является членом никакого $S_{\mathrm{lom}}(x, y)$;

б) вершина w входит в состав некоторого минимального (локально-минимального) сепаратора для пары вершин (x, y);

в) вершина w сама (одна) является минимальным сепаратором для (x, y).

В последнем случае вершины z и t также являются минимальными сепараторами для (x, y). Пример альтернативы а) можно построить согласно правилу чужого гена (факт 11 из [6]). Пример альтернативы б), когда $v \equiv w$, дан на рис. 1.

Предложение 10 (изоляция единственного возможного стержня сепаратора). Пусть в АОГ для пары вершин (x, y) имеем $\neg \mathrm{Ds}(x \perp\perp y)$ и $\mathbf{R} = \{r \mid \neg \mathrm{Ds}(r \perp\perp x) \& \neg \mathrm{Ds}(r \perp\perp y)\}$. Допустим, для всех вершин $r \in \mathbf{R}$, за исключением одной r^*, верно $\mathrm{Ds}(r \perp x \perp y)$ или $\mathrm{Ds}(r \perp y \perp x)$, или $\mathrm{Ds}(r \perp r^* \perp \{x, y\})$. Пусть $\mathbf{Z} = \{z \mid \neg \mathrm{Ds}(z \perp\perp r^*),\ z \notin \mathbf{R},\ z \neq x,\ z \neq y\}$. Если для всех вершин $z \in \mathbf{Z}$ выполняется $\mathrm{Ds}(z \perp r^* \perp x)$ или $\mathrm{Ds}(z \perp r^* \perp y)$, то либо существует ребро $x — y$, либо $\{r^*\}$ является единственным локально-минимальным сепаратором для (x, y).

Схема доказательства. Если выполняется $\mathrm{Ds}(w \perp x \perp y)$ или $\mathrm{Ds}(w \perp y \perp x)$, то следует применить утверждение 8 из [6] (отстранение вершины (placing aside)). Если выполняется $\mathrm{Ds}(w \perp r^* \perp x)$ или $\mathrm{Ds}(w \perp r^* \perp y)$, то следует повторить рассуждения из доказательства предложения 6. Заметим, что условия подразумевают, что выполняется $\neg \mathrm{Ds}(z \perp\perp x)$ или $\neg \mathrm{Ds}(z \perp\perp y)$, иначе получим ситуацию правила чужого гена [6]. Предложение допускает частный случай $\mathbf{Z}=\varnothing$.

Следующее предложение обобщает несколько правил, в том числе правило изолятора [6].

Предложение 11 (изолятор отстраняемой вершины (isolator under placing aside)). Если в АОГ имеем $\neg \mathrm{Ds}(x \perp\perp y)$, $\mathrm{Ds}(x \perp\perp z)$, $\neg \mathrm{Ds}(z \perp\perp w)$, $\mathrm{Ds}(r \perp x \perp y)$ и $\neg \mathrm{Ds}(r \perp x \perp w)$, где все вершины x, y, r, z, w различны, то вершина w не является членом никакого локально-минимального сепаратора для (x, y).

Доказательство. Сопоставляя условия $\mathrm{Ds}(x \perp\perp z)$ и $\neg \mathrm{Ds}(z \perp\perp w)$, заключаем, что не существует ни одного орпути $x \leftarrow \cdots \leftarrow w$. Предположим от противного, что $w \in S_{\mathrm{lom}}(x, y)$. Тогда согласно базовой теореме (утверждение 5 из [6]) или теореме 2 из [10] и ввиду отсутствия орпути $x \leftarrow \cdots \leftarrow w$ должен существовать орпуть $w \rightarrow \cdots \rightarrow y$, не проходящий через вершину x. Условие $\neg \mathrm{Ds}(r \perp x \perp w)$ означает, что между вершинами r и w существует путь, открытый при кондиционировании x. Этот путь без образования коллайдера стыкуется с орпутем $w \rightarrow \cdots \rightarrow y$, который не проходит через вершину x. Ввиду такой стыковки констатируем, что существует путь между вершинами r и y, открытый при кондиционировании x. Это противоречит условию $\mathrm{Ds}(r \perp x \perp y)$. □

Замечание к предложению 11. Вершина r также не входит в состав $S_{\text{lom}}(x, y)$. В условия предложения 11 включены минимально необходимые факты. Оно предназначено для ситуаций, где другие известные правила не применимы, и прежде всего для того, чтобы исключить из списка кандидатов в сепаратор вершины, которые являются потенциальными стержнями сепаратора для (x, y). Значит, подразумеваются факты $\neg\text{Ds}(w\perp\perp x)$, $\neg\text{Ds}(w\perp\perp y)$, $\neg\text{Ds}(w\perp y\perp x)$, $\neg\text{Ds}(w\perp x\perp y)$. Далее подразумеваем $\neg\text{Ds}(z\perp\perp y)$, иначе действует правило чужого гена. Из $\text{Ds}(x\perp\perp z)$ и $\neg\text{Ds}(z\perp\perp y)$ заключаем, что дуга $x\leftarrow y$ невозможна. Подразумевается $\neg\text{Ds}(w\perp\perp y)$ и $\neg\text{Ds}(w\perp x\perp y)$, иначе w отстраняется от пары (x, y). Сочетание фактов $\neg\text{Ds}(w\perp x\perp y)$, $\neg\text{Ds}(r\perp x\perp w)$ и $\text{Ds}(r\perp x\perp y)$ объясняет название предложения 11 (вершина w изолирует r от y). Поскольку указанные факты не включены в условия предложения 11, то предложение охватывает и некоторые из правил, рассмотренные ранее, и новые специфичные случаи. Три модели, где применимо предложение 11, даны на рис. 4.

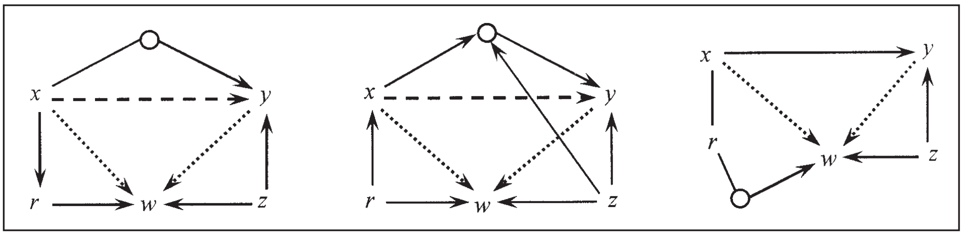

Рис. 4

2. ПРИНЦИПЫ ВЫВОДА АОГ-МОДЕЛЕЙ ИЗ ПАТТЕРНОВ ЗАВИСИМОСТЕЙ

Рассмотренные выше предложения и правила (а также результаты из [6, 11]) непосредственно предназначены для решения двух видов задач. Во-первых, они могут использоваться для вывода скелета АОГ-модели (т.е. всех ребер) из совокупности заданных d-зависимостей или d-независимостей. Другими словами, предложенные средства служат для компиляции (синтеза) структуры модели, исходя из знаний об отношениях зависимости (независимости) в модели. (В некоторых ситуациях предложенные правила можно применить непосредственно при выводе структуры АОГ-модели из данных, когда известны фрагменты структуры модели.) Во-вторых, предложенные инструменты применимы для анализа модели, в частности, для планирования схемы рассуждений на заданной модели с целью выполнения вероятностного вывода от свидетельств к целевым переменным (в экспертных системах). В таких рассуждениях информация распространяется по структуре модели через вершины, входящие во все локально-минимальные сепараторы $S_{\text{lom}}(x, y)$, где y — целевая переменная, а x — переменная, входящая в состав свидетельства.

Для объединения предложенных правил (при разработке алгоритма вывода (синтеза) структуры модели) необходимы базовые принципы. Алгоритм PC (наиболее известный из алгоритмов constraint-based подхода) для сокращения перебора использует простой принцип поиска сепаратора: сепаратор для пары (x, y) подбирается среди множества вершин, которые считаются (возможно) смежными к x или y на текущем этапе вывода модели. Это значительно сужает поиск и гарантирует нахождение сепаратора (если он существует). Но такая простая тактика поиска приводит к серьезным недостаткам. Алгоритм PC далеко не всегда находит минимальные сепараторы. Кроме того, PC не оснащен никакими средствами распознавания ребра (или выявления его отсутствия), и поэтому продолжает искать сепаратор до полного исчерпания всех возможностей, даже когда сепаратора не существует. Вследствие этого алгоритм выполняет излишнюю работу и рискует допустить ошибку при тестировании условной независимости с большой кардинальностью пробного сепаратора.

Принцип вывода модели, принятый в алгоритме PC, можно выразить следующим образом. Если испытаны и отвергнуты все подмножества из множества вершин, возможно, смежных к x (соответственно к y), то существует ребро $x — y$. Поскольку мы вводим дополнительные требования к кандидатам в члены сепаратора и к составу сепаратора в целом, то появляется возможность повысить эффективность подбора сепараторов, соответственно уточнив указанный принцип. Предложенные требования и правила часто позволяют отсеять значительно больше кандидатов в сепаратор (иногда даже всех кандидатов). Тогда поиск сепаратора прекращается на раннем этапе. Названный принцип вывода модели перерастает в новую форму, записанную ниже.

Резолюция смежности. Если все кандидаты в сепараторы для пары ассоциированных вершин (x, y) отвергнуты или испытаны все подмножества из множества неотверженных кандидатов в сепаратор, то существует ребро $x — y$.

Как было показано в [6, 11], в составе каждого непустого сепаратора должен быть хотя бы один потенциальный стержень сепаратора. Но некоторые (иногда даже все) потенциальные стержни сепаратора для (x, y) могут быть отвержены, т.е. могут быть найдены свидетельства, согласно которым эти потенциальные стержни не являются членами никакого локально-минимального сепаратора $S_{\text{lom}}(x, y)$. Такое отвержение дают, в частности, правила чужого гена, изолятора [6], аппендикса, полуаппендикса и изолятора отстраняемой вершины. Поэтому резолюция смежности может быть «перекрыта» (outperformed) следующим новым принципом.

Экспресс-резолюция смежности. Если для пары вершин (x, y) не осталось ни одного неотверженного (незабракованного) потенциального стержня сепаратора, то существует ребро $x — y$.

Если имеется возможность идентифицировать не только потенциальные стержни сепаратора, но и претенденты в стержни сепаратора, то названный принцип можно усилить следующим образом.

Специальная экспресс-резолюция смежности. Если для пары вершин (x, y) все претенденты в стержни сепаратора отвергнуты, то существует ребро $x — y$.

Ряд рассмотренных правил (непоглощения, квазиинструментальной пары, полуаппендикса) подсказывает необходимость концептуально нового принципа вывода, который дает возможность избежать поиска соответствующего сепаратора и отсеивания кандидатов в сепаратор. Предложим новую формулировку принципа.

Резолюция несмежности. Если набор статистических фактов свидетельствует о невозможности присутствия ни дуги (орпути) $x \to y$, ни дуги (орпути) $y \to x$, то ребро $x — y$ отсутствует.

Открытым остается вопрос: следует ли сохранять тактику алгоритма PC (т.е. ограничивать множество кандидатов в сепаратор для пары (x, y) только вершинами, смежными к x или к y) или отбросить это ограничение и использовать только обоснованные выше требования к членам локально-минимальных сепараторов. Как было показано на примере (см. рис. 1), такое ограничение может привести к потере минимальных сепараторов. Ввиду этого целесообразно не ограничивать поиск сепаратора смежными вершинами, а использовать для сокращения списка кандидатов только обоснованные выше правила. С одной стороны, разработанные правила (в отличие от алгоритма PC) иногда позволяют исключить вершину, смежную к одной из сепарируемых вершин. С другой стороны, эти правила обычно не могут исключить из кандидатов все вершины, не смежные к обеим сепарируемым вершинам. В результате список кандидатов в сепаратор получится более объемным. Можно дать следующую рекомендацию. Если приоритетом является сокращение списка кандидатов в сепаратор и сокращение числа тестов, то следует сохранить тактику алгоритма PC (соответственно дополнив ее). Если же приоритетом является минимальность сепаратора, то целесообразно отказаться от тактического ограничения алгоритма PC. При решении этого вопроса необходимо учитывать размер выборки данных, требования к надежности и знания о модели. Также следует иметь в виду возможность разработки новых правил формирования локально-минимальных сепараторов.

Ясно, что когда существует минимальный сепаратор кардинальности, равной единице, алгоритм PC обязательно найдет этот сепаратор (один из минимальных). В таких ситуациях рассмотренные правила не могут повысить эффективности алгоритма PC. Однако это не означает бесполезности применения тех правил, в резюмирующей части которых записана (как одна из альтернатив) условная независимость первого ранга. Важная роль таких правил — остановка поиска сепаратора. (Эти правила могут быть использованы также и в других целях.)

В разд. 1 рассмотрен пример, где тесты первого ранга исключают члена минимального сепаратора кардинальности, равной трем. Такому синдрому потери минимального сепаратора подвержен, в частности, и алгоритм PC. Заметим, что в этом примере происходит «встречное» открывание коллайдерных путей. Если ограничиться случаями, где нет открывания коллайдерных путей, то в алгоритме PC тесты первого ранга, по-видимому, не приведут к потере минимального сепаратора. Простейший пример потери минимального сепаратора алгоритмом PC (при условии, что нет открывания коллайдерных путей) дан в [11]. В рассмотренном там примере «ошибочное» исключение члена минимального сепаратора (который имеет кардинальность, равную трем) происходит в результате выполнения тестов второго ранга. (Одной из предпосылок такой «ошибки» является то, что все цепи между вершинами x и y имеют вид $x \leftarrow \cdots \rightarrow y$.) Значит, простейший минимальный сепаратор, который может быть не найден алгоритмом PC (без замены равноценным сепаратором), имеет кардинальность, равную трем. (Алгоритм PC также может потерять минимальный сепаратор из двух переменных в результате выполнения тестов второго ранга, но только при определенной последовательности тестов.)

3. ПОДБОР ЭМПИРИЧЕСКИХ СЕПАРАТОРОВ И ВЫВОД РЕБЕР МОДЕЛИ ИЗ ДАННЫХ

Изложенные результаты пригодны для методов constraint-based подхода к выводу структуры модели из статистических данных [2, 3, 7, 8]. Основной этап вывода структуры модели из данных — построение ее скелета — состоит в идентификации всех ребер. Наличие или отсутствие ребер верифицируется с помощью тестирования статистических фактов условной независимости переменных в выборке данных.

Условную независимость переменной x от переменной y при кондиционировании множества переменных $\mathbf{Z}$ $(x, y \notin \mathbf{Z})$ обозначим $\Pr(x \perp \mathbf{Z} \perp y)$. При этом множество переменных $\mathbf{Z}$ будем называть эмпирическим (статистическим) сепаратором для пары (x, y). Для дискретных переменных условная независимость означает $p(xy|\mathbf{Z}) = p(x|\mathbf{Z}) \cdot p(y|\mathbf{Z})$. В линейных моделях условная независимость проявляется как нулевое (незначимое) значение коэффициента частной корреляции. Безусловную независимость $\Pr(x \perp \varnothing \perp y)$ будем записывать в виде $\Pr(x \perp\perp y)$, а безусловную зависимость — в виде $\neg\Pr(x \perp\perp y)$. (Заметим, что в конечной выборке слабая зависимость неотличима от «шума» при независимости.)

Известно [13], что критерий d-сепарации определяет марковские свойства АОГ-модели (т.е. отношения условной независимости, выполняющиеся при любой параметризации модели):

$$\mathrm{Ds}(x \perp \mathbf{Z} \perp y) \Rightarrow \Pr(x \perp \mathbf{Z} \perp y). \quad (3)$$

Для обоснования методов вывода структуры из данных требуются соответствующие предположения. Обычно таким является предположение каузальной необманчивости (faithfulness) распределения вероятностей переменных модели относительно АОГ [2, 3, 7, 10]. (Слово «каузальная» далее будем опускать.) Полная форма предположения необманчивости распределения вероятностей выражается как

$$\Pr(x \perp \mathbf{Z} \perp y) \Rightarrow \mathrm{Ds}(x \perp \mathbf{Z} \perp y), \quad (4)$$

т.е. как обратная импликация по отношению к (3). Объединяя (3) и (4), получаем

$$\mathrm{Ds}\,(x \perp \mathbf{Z} \perp y) \Leftrightarrow \mathrm{Pr}\,(x \perp \mathbf{Z} \perp y)\,. \tag{5}$$

Выполнение (4) обеспечивает структурно-поведенческий изоморфизм модели.

Предположение (4) является излишне сильным. Многие методы вывода скелета АОГ-модели используют ослабленную форму предположения необманчивости:

$$\mathrm{Pr}\,(x \perp \mathbf{Z} \perp y) \Rightarrow \neg (x - y)\,. \tag{6}$$

Импликацию (6) можно назвать реберной необманчивостью. Учитывая первичное свойство (1) сепарации в АОГ, получим единый принцип идентификации ребер модели

$$\forall\, \mathbf{Z}\,(x, y \notin \mathbf{Z}) : \neg \mathrm{Pr}\,(x \perp \mathbf{Z} \perp y) \Leftrightarrow (x - y)\,. \tag{7}$$

Предположение реберной необманчивости (6) является достаточным для обоснования результативного алгоритма вывода скелета модели из данных.

Если принять полную форму предположения необманчивости, то согласно (5) можно транскрибировать все правила, обоснованные в предыдущем разделе и в [6], и получить эмпирические двойники (counterparts) этих правил для использования в качестве инструментов вывода скелета модели из данных. Но даже реберная необманчивость (6), т.е. существенно ослабленная форма предположения необманчивости, может изредка нарушаться (особенно в небольшой выборке данных), что создает риск ошибки. Чтобы обосновать эмпирические версии предложенных правил формирования сепараторов, недостаточно предположения (6). В то же время для такого обоснования не обязательна полная форма предположения необманчивости (4).

Например, для прямой идентификации всех цепей в АОГ-модели с помощью процедуры Геном-1 [12] достаточно предположения, формализуемого в виде

$$\mathrm{Pr}\,(x \perp\, \perp y) \Rightarrow \mathrm{Ds}\,(x \perp\, \perp y)\,, \tag{8}$$

которое можно назвать безусловной (маргинальной) цепной необманчивостью. Некоторые правила формирования сепараторов требуют близкого к (8) предположения. Действительно, во многих доказательствах использована импликация вида «бесколлайдерная стыковка» (splicing) двух цепей. В терминах d-сепарации эта импликация корректна. Но аналогичный способ аргументации для эмпирических зависимостей (в выборке реалистического объема) часто приводит к ошибке (ввиду слабости транзитной зависимости). Заметим, что для представленных доказательств можно слегка ослабить предположение (8) до более локальной версии.

Очевидно, безусловная цепная необманчивость недостаточна для тех правил, где используется условная независимость первого ранга. Для таких правил достаточно добавить предположение необманчивости первого ранга:

$$\forall\, x, y, z:\ \mathrm{Pr}\,(x \perp z \perp y) \Rightarrow \mathrm{Ds}\,(x \perp z \perp y)\,. \tag{9}$$

Однако (9) является избыточно сильным предположением. Отдельные рассмотренные правила опираются на разные (слегка ослабленные, локальные) версии предположения (9), которые еще предстоит специфицировать.

Подчеркнем, что в применении к конечной выборке данных предположение необманчивости означает, что отрицание вероятностной независимости должно всегда подтверждаться достаточной величиной эмпирической зависимости. Иными словами, из отрицания независимости должно следовать, что величина зависимости превышает порог, который мы выбрали для отсеивания выборочного «шума». Особенно нежелательно ошибочное отвержение единственного стержня сепаратора, поскольку такая ошибка уже не компенсируется в дальнейшем.

Методы сепарационного подхода подвержены риску ошибки при идентификации ребер модели по причине ненадежности тестирования условной независимости (особенно когда выборка данных невелика). Другой недостаток — неудовлетворительная вычислительная эффективность поиска сепараторов. Вычислительная сложность определяется количеством выполняемых тестов и объемом вычисления необходимых статистик. В дискретных моделях объем вычисления статистик, необходимых для теста условной независимости $\Pr(x \perp \mathbf{Z} \perp y)$, определяется рангом теста, т.е. кардинальностью условия $\mathbf{Z}$. По мере роста кардинальности $\mathbf{Z}$ в тесте выборка данных дробится на части и надежность теста снижается. Для выполнения теста условной независимости k–го ранга требуется статистика размерности $(k+2)$. Поэтому целесообразно по возможности обходиться тестами низкого ранга и находить минимальные сепараторы. Кроме того, возможна такая постановка задачи, когда далеко не всякую статистику можно вычислить из имеющихся данных. Тогда аналитик вынужден идентифицировать ребро или его отсутствие косвенно. Желательно как можно раньше выявить отсутствие сепаратора для заданной пары переменных и прекратить бесперспективный поиск.

ЗАКЛЮЧЕНИЕ

Установленное множество паттернов зависимости (независимости) помогает косвенно идентифицировать факты сепарации переменных АОГ-модели и, в частности, позволяют идентифицировать присутствие или выявить отсутствие ребра или принадлежность (непринадлежность) переменной к локально-минимальному сепаратору. Показано, каким образом можно идентифицировать существование (несуществование) или состав сложных сепараторов, исходя из знания набора простых сепараторов в окрестности. В частности, правила позволяют идентифицировать отсутствие ребра косвенно, т.е. без нахождения соответствующего факта независимости. Исходными фактами для разработанных правил вывода выступают пустые сепараторы и сепараторы из одной переменной, что соответствует безусловной независимости и условной независимости первого ранга. Предложены новые принципы индуктивного вывода скелета АОГ-модели.

СПИСОК ЛИТЕРАТУРЫ

1. Pearl J. Graphs, causality, and structural equation models // Sociological Methods & Research. — 1998. — **27**, N 2. — P. 226–284.
2. Scheines R., Spirtes P., Glymour C., Meek C., and Richardson T. The TETRAD project: Constraint based aids to causal model specification // Multivar. Behavior. Res. — 1998. — **33**, N 1. — P. 65–118.
3. Neapolitan R.E. Learning Bayesian networks. — Upper Saddle River: Prentice Hall, NJ, 2004. — 694 p.
4. Pearl J. Causality: models, reasoning, and inference. — Cambridge: Cambridge Univ. Press, 2000. — 526 p.
5. Lauritzen S.L. Causal inference from graphical models // Complex Stochastic Systems / O.E. Barndorff-Nielsen, D.R. Cox, and C. Klüppelberg (Eds.). — London: Chapman and Hall, 2000. — P. 63–107.
6. Балабанов А.С. Минимальные сепараторы в структурах зависимостей. Свойства и идентификация // Кибернетика и системный анализ. — 2008. — № 6. — С. 17–32.
7. Zhang J. Causal inference and reasoning in causally insufficient systems: PhD Thesis / Dep. of Philosophy. — Pittsburgh, PA: Carnegie Mellon University, 2006. — 246 p.
8. Ramsey J., Spirtes P., and Zhang J. Adjacency-faithfulness and conservative causal inference // Proc. of the 22nd Conf. on Uncertainty in Artificial Intelligence. — Oregon: AUAI Press, 2006. — P. 401–408.
9. Acid S. and L.M. de Campos. An algorithm for finding minimum d-separating sets in belief networks / E. Horvitz, F. Jensen (Eds.) // Proc. of the 12th Conf. on Uncertainty in Artificial Intelligence. — San Mateo: Morgan Kaufmann, CA, 1996. — P. 1–22.
10. Tian J., Paz A., and Pearl J. Finding minimal d-separators. — Computer Science Dep., Technical Rep. — R-254. — UCLA, CA, 1998. — 15 p.
11. Балабанов О.С. Правила підбору сепараторів у баєсівських мережах // Проблеми програмування. — 2007. — № 4. — С. 22–33.
12. Балабанов А.С. Восстановление структур систем вероятностных зависимостей из данных. Аппарат генотипов переменных // Проблемы управления и информатики. — 2003. — № 2. — С. 91–99. (http://www.begellhouse.com/journals/)
13. Verma T., and Pearl J. Causal networks: semantics and expressiveness / R. Shachter, T.S. Levitt, L.N. Kanal (Eds.) // Uncertainty in Artificial Intelligence. — Amsterdam: Elsevier Science Publishers, North-Holland, 1990. — **4**. — P. 69–76.

Поступила 29.04.2009