# БАЙЕСОВСКИЕ СЕТИ ДОВЕРИЯ: ДЕРЕВО СОЧЛЕНЕНИЙ И ЕГО ВЕРОЯТНОСТНАЯ СЕМАНТИКА♦

А. В. Сироткин

Санкт-Петербургский институт информатики и автоматизации РАН

СПИИРАН, 14-я линия ВО, д. 39, Санкт-Петербург, 199178

<sirotkin@hotbox.ru>

УДК 681.3

*Сироткин А. В.* **Байесовские сети доверия: дерево сочленений и его вероятностная семантика** // Труды СПИИРАН. Вып. 3, т. 1. — СПб.: Наука, 2006.

**Аннотация.** *Рассматриваются байесовские сети доверия (БСД) с бинарными переменными в вершинах. В предположении условной независимости сравниваются семантики глобального распределения вероятностей и локальных распределений, соответствующих узлам дерева сочленений.* — Библ. 12 назв.

UDC 681.3

*Sirotkin A. V.* **Bayesian Belief Networks: Junction Tree and It's Probabilistic Semantics** // SPIIRAS Proceedings. Issue 3, vol. 1. — SPb.: Nauka, 2006.

**Abstract.** *We consider Bayesian belief networks (BBN) with binary variables in nodes. We assume the condition independence and compare semantics of global joint probability and a set of local joint probabilities corresponded to junction tree nodes.* — Bibl. 12 items.

## 1. Введение

Объектом изучения в данной статье являются байесовские сети доверия. Впервые данное понятие вводится в работе [1] и в [2] подвергается достаточно полному и систематическому анализу. В [2] рассматриваются ациклические графы (полидеревья), на которых вводится структура байесовской сети доверия. В этой же работе говорится, что рассмотрение байесовских сетей доверия на графе с циклами — довольно сложная задача. За время, прошедшее с начала исследований байесовских сетей доверия, удалось создать метод, который позволяет оперировать с БСД, построенными на направленных графах, не содержащих направленных циклы, но, возможно, содержащих ненаправленные. Схемы алгоритмов пропагации[1] приводятся, например, в [3]. Дерево сочленений является основным видом структуры данных для алгоритмов в сетях, содержащих ненаправленные циклы.

В статье мы рассмотрим байесовские сети доверия и деревья сочленений с точки зрения вероятностной семантики. В качестве вероятностной семантики мы выбрали совместное распределение вероятностей над множеством всех переменных, входящих в байесовскую сеть доверия. Цель работы — доказать, что в рамках логико-вероятностного подхода мы получаем одно и тоже вероятностное распределение как для традиционного построения распределения в байесовской сети доверия, так и для построения совместного распределения

♦Часть результатов, представленных в настоящей работе, получена в рамках работ по госконтракту № 02.442.11.7289, шифр 2006-РИ-19.0/001/211.

[1]Пропагация — (от англ. propagation — распространение) в контексте байесовских сетей обозначает вычисление маргинальных вероятностей (первичная пропагация) или вычисление условных вероятностей не присутствующих в исходной сети (пропагация свидетельств). При пропагации происходит распространение влияния условных вероятностей отдельных узлов и поступивших свидетельств на всю сеть.

на основе дерева сочленений, согласованного с байесовской сетью доверия, при предположении условной независимости между локальными распределениями над переменными в узлах дерева сочленений при заданных означиваниях переменных, входящих в сепаратор.

## 2. Вероятность пропозициональной формулы по Н. Нильссону

Существуют различные способы введения вероятности произвольной пропозициональной формулы. Мы воспользуемся подходом Н. Нильссона [4, 5]. При работе с байесовскими сетями доверия мы всегда оперируем конечным набором вершин. В рамках данной статьи каждой вершине соответствует атомарная пропозициональная переменная[2], поэтому можно предполагать, что мы изначально ограничены конечным множеством пропозициональных переменных $A = \{x_1, \ldots, x_n\}$.

Суть подхода Н. Нильссона описана, например, в [6] и состоит в следующем. Дан конечный набор пропозициональных формул. Все возможные истинностные означивания этого набора назовем *возможными мирами*. Все непротиворечивые означивания назовем *допустимыми мирами*. Рассмотрим множество допустимых миров как множество элементарных событий. Зададим на нем распределение вероятностей, подчиняющихся двум требованиям: вероятности элементарных событий неотрицательны, и сумма вероятностей всех элементарных событий равна единице. Тогда вероятностью истинности формулы будем считать сумму вероятностей тех допустимых миров, в которых она принимает значение «истина».

В качестве множества элементарных событий выберем множество конъюнкции из $n$ переменных со всевозможными означиваниями (квантов) $Q = Q(A) = \{\tilde{x}_1 \tilde{x}_2 \ldots \tilde{x}_n\}$. Здесь $\tilde{x}_i$ означает одно из двух возможных состояний: $x_i$ или $\bar{x}_i$. По теореме о представлении пропозициональной формулы в совершенной дизъюнктивной нормальной форме (СДНФ) любая пропозициональная формула представима в виде дизъюнкции формул, входящих в выбранное нами множество элементарных событий. Введем функцию $S : F \to 2^Q$, $S(f) = S_f$, $S(f) = S_f$, где $S_f$ — множество квантов, участвующих в СДНФ пропозициональной формулы $A' \subseteq A$. Теперь зададим на множестве элементарных событий $Q$ такое вероятностное распределение: $p^\circ : Q \to [0;1]$, что $\forall q \in Q\ p^\circ(q) \geq 0$, $\sum_{q \in Q} p^\circ(q) = 1$. Введем вероятность $p : 2^Q \to [0;1]$ следующим образом:

$$\forall (S \subseteq Q)\ \ p(S) = \sum_{q \in S} p^\circ(q).$$

На этом шаге мы получили вероятностное пространство $\left\langle Q, 2^Q, p \right\rangle$. Определим вероятностную меру на множестве всех пропозициональных формул следующим образом: $p(f) = p(S(f))$. Таким образом, мы ввели вероятность на произвольной пропозициональной формуле.

---

[2] В общем случае вершине байесовской сети может соответствовать многозначная переменная, но в рамках логико-вероятностного подхода она может быть представлена как набор пропозициональных переменных. При этом каждое означивание этого набора переменных будет соответствовать одному значению многозначной переменной.



Труды СПИИРАН. 2006. Вып. 3. Т. 1. ISSN 2078-9181 (печ.), ISSN 2078-9599 (онлайн)
SPIIRAS Proceedings. 2006. Issue 3. V. 1. ISSN 2078-9181 (print), ISSN 2078-9599 (online)
www.proceedings.spiiras.nw.ru

## 3. Определение байесовской сети доверия

Основу байесовской сети доверия составляют два объекта. Первый — это ациклический направленный граф. Второй — это набор распределений (тензоров) условных вероятностей, заданных для каждой из вершин графа. Начнем с первого объекта и ряда связанных с ним вопросов, после чего перейдем ко второму. Более подробное описание аппарата байесовских сетей доверия представлено в [3, 7, 8].

Ациклический направленный граф — это направленный граф, в котором нет направленных циклов. Заметим, что в работе [2] основоположник аппарата байесовских сетей доверия Дж. Пиерл отказывается от любых циклов, — и от направленных, и от ненаправленных, — работая только с полидеревьями. Однако циклы встречаются в практических задачах, например в турбокодировании [9, 10]. Поэтому современные исследователи не отказываются от ненаправленных циклов, но при этом оговаривают, что происходит усложнение алгоритмов обработки. Ациклический направленный граф используется для того, чтобы представить непосредственные причинно-следственные связи между переменными, лежащими в смежных вершинах графа. Эти переменные представляют собой различные атомарные утверждения об объектах из предметной области. Мы будем рассматривать случай, когда вершины данного графа — это атомарные пропозициональные переменные, а ребра соответствуют зависимостям между ними. Можно рассматривать общий случай, когда в вершинах графа стоят переменные с произвольным конечным числом значений (возможно, разным для разных переменных), но это выходит за рамки логико-вероятностного подхода, принятого в нашем изложении. На рис. 1 изображена некая причина $x$ и ее следствие $y$ (здесь и далее в статье для обозначения вершин и пропозициональных переменных, соответствующих вершинам, будут использоваться маленькие латинские буквы). Важно оговорить, что при построении байесовской сети доверия не всегда просто определить, какое из двух событий — причина, а какое — следствие[3]. Решение этого вопроса выходит за рамки математики и относится к предметной области. Важно отметить тот факт, что взаимосвязанными могут являться не только причина и следствие, но и события связанные более сложным образом. Чтобы их описать выделим три типа связи (рис. 2).

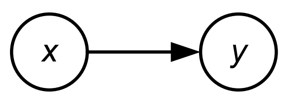

Рис. 1. Простая БСД.

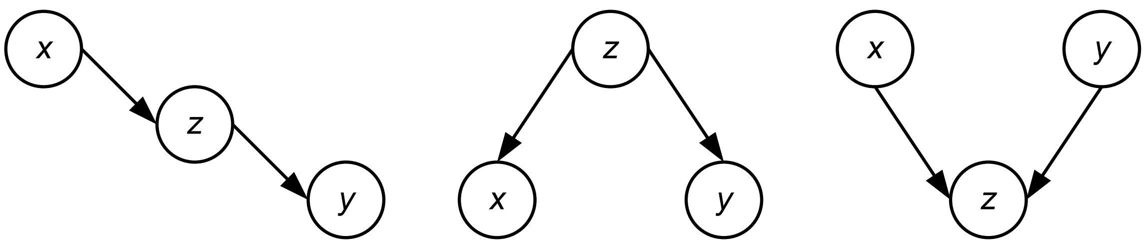

Рис. 2. Различные типы связей, встречающиеся в БСД.
*а* — последовательная связь; *б* — расходящаяся связь; *в* — сходящаяся связь.

[3]Например, в рассуждении «у меня температура, значит, я простудился» причиной является простуда, а следствием температура, а не наоборот, как может показаться по форме рассуждения.



Труды СПИИРАН. 2006. Вып. 3. Т. 1. ISSN 2078-9181 (печ.), ISSN 2078-9599 (онлайн)
SPIIRAS Proceedings. 2006. Issue 3. V. 1. ISSN 2078-9181 (print), ISSN 2078-9599 (online)
www.proceedings.spiiras.nw.ru

*Последовательная связь* — связь, представленная на рис. 2а: связаны сын и родитель одной и той же вершины. *Расходящаяся связь* — на рис. 2б: между двумя сыновьями одного и того же элемента. *Сходящаяся связь* — на рис. 2в: между двумя родителями одного и того же элемента.

Заметим, что эти три типа связи исчерпывают все возможности соединения трех вершин на произвольном пути графа. Теперь перейдем к определению понятия $d$-разделимости согласно [3].

*Две вершины $x$ и $y$ направленного графа называются* $d$-разделенными*, если для любого пути из $x$ в $y$ (без учета направления ребер) существует такая вершина $z$, не совпадающая с $x$ и $y$, что либо связь в вершине $z$ расходящаяся или последовательная, и $z$ получило означивание, либо связь сходящаяся, и ни $z$, ни какой-либо из его потомков не получили означивания.*

Поясним смысл этого определения на следующем примере. Обратимся к рис. 3, где изображена простая структура байесовской сети доверия. Предположим, мы знаем, что горло болит. Это, естественно, повышает вероятность простуды, при этом также возрастет и вероятность того, что есть температура. Но если мы уже знаем, что простужены, то наблюдение того, болит горло или нет, никаким образом не повлияет на вероятность того, что есть температура. Таким образом, мы столкнулись с ситуацией, когда от наличия или отсутствия свидетельства будет зависеть, влияет ли наше знание о состоянии одной вершины на оценку вероятности истинности другой. Для разделения не влияющих друг на друга вершин и было введено фундаментальное понятие $d$-разделимости [3]. Нас будет интересовать лишь логико-вероятностное описание ограничений, которые она накладывает. Если два множества переменных (вершин БСД) связаны путем, проходящим через вершину $z$, и связь в этой вершине последовательная или расходящаяся, то требование $d$-разделимости будет эквивалентно тому, будет эквивалентно тому, что означивания этих множеств условно независимы при заданном $z$ как случайные величины. Если же эти два множества связаны путем со сходящейся связью и не связаны другими путями, то при отсутствии свидетельств условие $d$-разделимости будет эквивалентно тому, что означивания этих множеств независимы как случайные величины.

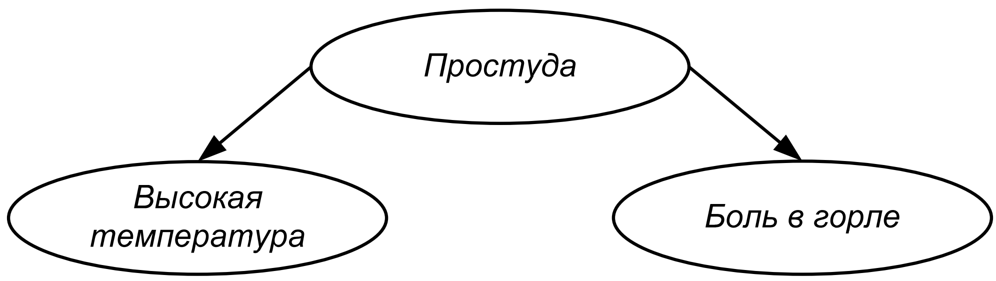


Рис. 3. Пример простой структуры байесовской сети доверия.

Перейдем к описанию второй важной составляющей байесовской сети доверия — набору тензоров условных вероятностей, приписанных каждой вершине. Условные вероятности определяют тесноту взаимосвязи причины и следствия. Для каждой вершины мы задаем условные вероятности истинности пропозиции, соответствующей этой вершине, при всех возможных означиваниях родителей данной вершины. Здесь важно подчеркнуть, что недостаточно задать условную вероятность для *пар* родитель–сын. Эти сведения не определяют



Труды СПИИРАН. 2006. Вып. 3. Т. 1. ISSN 2078-9181 (печ.), ISSN 2078-9599 (онлайн)
SPIIRAS Proceedings. 2006. Issue 3. V. 1. ISSN 2078-9181 (print), ISSN 2078-9599 (online)
www.proceedings.spiiras.nw.ru

взаимосвязь двух родителей и поэтому не определяют однозначно распределение вероятностей. Например, в графе на рис. 2в для вершины $z$ мы должны будем задать следующий набор условных вероятностей:

$$p(z \mid xy),\ p(z \mid \bar{x}y),\ p(z \mid x\bar{y}),\ p(z \mid \bar{x}\bar{y}),\ p(\bar{z} \mid xy),\ p(\bar{z} \mid \bar{x}y),\ p(\bar{z} \mid x\bar{y}),\ p(\bar{z} \mid \bar{x}\bar{y}).$$

Для удобства, когда нам надо перечислить все эти условные вероятности, мы будем использовать сокращенную запись вида $p(\tilde{z} \mid \tilde{x}\tilde{y})$.

Заметим, что, задавая набор условных вероятностей, можно выписывать только половину из указанных значений. Это связано с тем, что оставшиеся значения в силу свойств вероятности могут быть вычислены по формуле

$$p(z \mid \tilde{W}) = 1 - p(\bar{z} \mid \tilde{W}).$$

Для вершин, не имеющих родителей, мы получим условную вероятность при пустом условии, то есть зададим маргинальную вероятность.

Основная задача логико-вероятностного вывода (первичной пропагации) в байесовских сетях доверия — это вычисление маргинальных вероятностей какой-либо пропозиции при заданных в каждой вершине условных вероятностях. Для случая БСД на основе полидеревьев алгоритмы пропагации дают однозначный результат при вычислении маргинальных вероятностей. Это связано с тем, что между любыми двумя вершинами существует только один путь, и влияние свидетельства может распространяться только по нему. При наличии циклов для каких-то пар вершин будут существовать хотя бы два различных пути, и, соответственно, при пропагации свидетельства мы должны будем корректно совместить влияния свидетельства, распространившиеся разными путями. В работе [11] было показано, что, если разрешить направленные циклы, то можно построить противоречивую сеть, то есть сеть, для которой нельзя будет построить набор маргинальных вероятностей над конъюнкциями всех атомарных переменных сети, по которым бы восстанавливались исходные условные. Но в нашем случае направленные циклы запрещены, а разрешены только ненаправленные, и, как показано, например, в [7], мы можем построить соответствующее маргинальное распределение.

## 4. Дерево сочленений

Начнем с определения дерева смежности. *Дерево смежности* — это ненаправленный связный ациклический граф (дерево) с заданными весами в вершинах, удовлетворяющий следующим требованиям:

1) веса вершин и ребер — множества переменных;
2) вес ребра, соединяющего две вершины, — пересечение весов вершин;
3) любые две вершины соединены путем;
4) любой вес (ребра или вершины), лежащий на пути от одной вершины к другой, содержит пересечение весов этих вершин;
5) ни для какой вершины ее вес не является подмножеством веса другой вершины.

Мы определили абстрактное дерево смежности. Теперь покажем, как дерево смежности связано с байесовской сетью доверия. Вершины дерева смежности мы будем называть узлами, чтобы отличать их от вершин байесовской сети доверия. Основное требование, которое мы будем предъявлять к дереву смежности байесовской сети доверия — это то, что для любой вершины БСД есть узел дерева смежности, в котором эта вершина лежит вместе со всеми своими родителями. Кроме этого, когда говорят о вычислительных аспектах, предпочитают говорить об оптимальном дереве смежности. Здесь оптимальность обычно устанавливается как минимизация числа элементов в самом большом весе.

Дерево сочленений отличается от дерева смежности тем, что узлам, кроме множеств, сопоставляются тензоры условных вероятностей, заданные в исходной БСД. Тензор условной вероятности из БСД приписывается узлу дерева сочленений, если вес узла содержит все переменные, фигурирующие в этом тензоре. Таким образом, любая условная вероятность, заданная в БСД, будет приписана хотя бы одному узлу дерева сочленений.

Пример байесовской сети доверия и соответствующего ей дерева смежности изображен на рис. 4.

Выше мы описали два различных способа — байесовская сеть доверия и дерево сочленений — представления данных с неопределенностью на основе тензоров условных вероятностей. Теперь можно строго сформулировать основную задачу статьи. Рассмотрим вероятностную семантику двух описанных объектов (байесовская сеть доверия и дерево сочленений). Под вероятностной семантикой мы подразумеваем совместное распределение вероятностей над конъюнкциями всех переменных, встречающихся в БСД. Для того чтобы показать совпадение семантик мы вначале опишем структуру вероятностных распределений, возникающих в узлах дерева сочленений. Потом покажем, что локальные распределения в дереве сочленений согласованы с вероятностями, заданными в байесовской сети доверия. После описания локальных распределений мы построим общее распределение для всех переменных и покажем, что распределения над конъюнкциями всех атомарных пропозиций совпадут для БСД и для дерева сочленений.

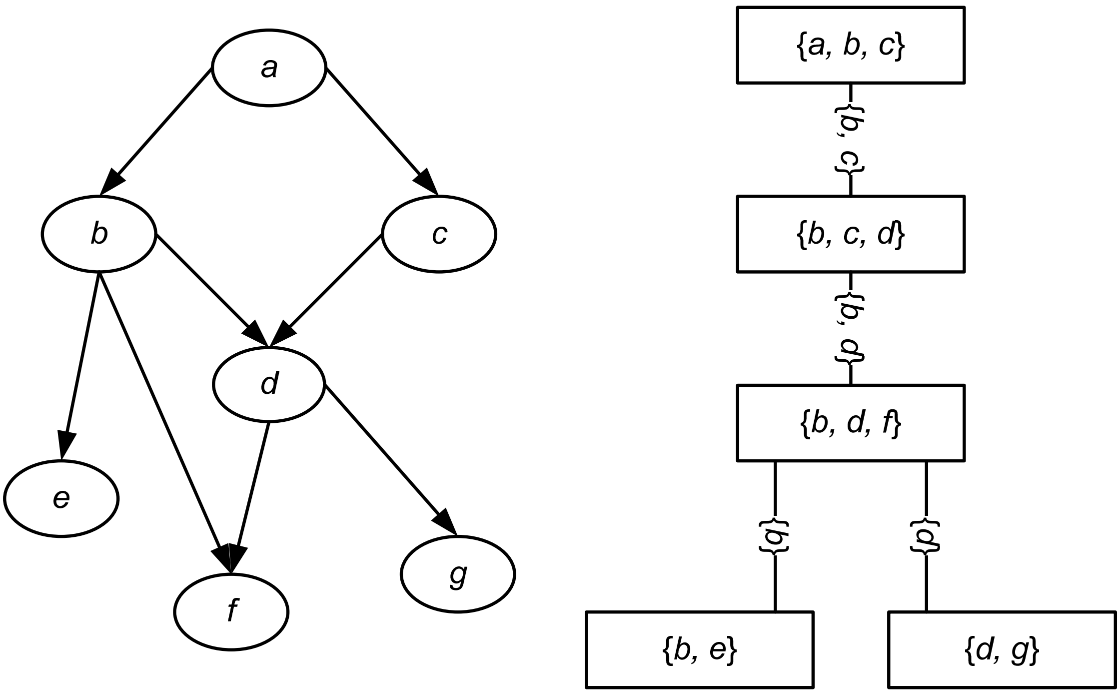


Рис. 4. Пример байесовской сети доверия и соответствующего ей дерева смежностей.



Труды СПИИРАН. 2006. Вып. 3. Т. 1. ISSN 2078-9181 (печ.), ISSN 2078-9599 (онлайн)
SPIIRAS Proceedings. 2006. Issue 3. V. 1. ISSN 2078-9181 (print), ISSN 2078-9599 (online)
www.proceedings.spiiras.nw.ru

## 5. Локальные распределения в вершинах дерева сочленений

Перейдем непосредственно к анализу вероятностных распределений, возникающих при работе с БСД и деревом сочленений. Основное предположение, которое используется при работе с байесовской сетью доверия (chain rule [3]), заключается в том, что общее распределение определяется по формуле:

$$p(\tilde{x}_1\tilde{x}_2\ldots\tilde{x}_n)=\prod_{i=1}^{n}p(\tilde{x}_i\mid\tilde{\wedge}\mathrm{pa}(x_i)),$$

где $\mathrm{pa}(x_i)$ означает конъюнкцию множества родителей вершины $x_i$. Символ $\tilde{\wedge}$ мы используем для описания представителя семейства конъюнкций элементов множества со всевозможными означиваниями. Например, $\tilde{\wedge}\{x,y\}=\tilde{x}\tilde{y}\in\{xy,x\overline{y},\overline{x}y,\overline{x}\overline{y}\}$. Иными словами, общее распределение равно произведению всех тензоров, фигурирующих в исходной байесовской сети доверия.

Основная цель пропагации (первичной или свидетельства) — это вычисление распределения на подмножестве переменных, входящих в байесовскую сеть доверия. Если у нас полностью задано распределение, то достаточно просто просуммировать это распределение по всем возможным означиваниям переменных, не входящих во множество, для которого мы хотим получить распределение. Однако размер задачи растет экспоненциально от количества переменных, и подобное суммирование в явном виде может оказаться невозможным на практике.

Рассмотрим распределение вероятностей над переменными произвольного узла дерева сочленений. Его можно вычислить на основе общего распределения, проведя маргинализацию (суммирование по всем означиваниям оставшихся переменных):

$$p(\tilde{y}_1\tilde{y}_2\ldots\tilde{y}_m)=\sum_{\tilde{x}_1\tilde{x}_2\ldots\tilde{x}_n}p(\tilde{x}_1\tilde{x}_2\ldots\tilde{x}_n\tilde{y}_1\tilde{y}_2\ldots\tilde{y}_m).$$

Заметим, что в силу коммутативности сложения определенные подобным образом вероятностные распределения для двух пересекающихся множеств переменных будут согласованны на общей части.

Описанный выше метод — это один из способов определять локальные распределения. Теперь предположим, что у нас не определено общее распределение, а заданы локальные распределения. При этом на основе локальных распределений можно рассчитать условные вероятности. Если рассчитанные вероятности совпадут с вероятностями, заданными в байесовской сети доверия, то мы будем говорить, что заданные локальные распределения *согласованы* с байесовской сетью доверия.

Кроме того, будем требовать условной независимости в соответствии с $d$-разделимостью: если два множества вершин байесовской сети доверия $\{x_1,x_2,\ldots,x_n\}$ и $\{y_1,y_2,\ldots,y_m\}$ $d$-разделены вершиной $z$, и связь в этой вершине расходящаяся или последовательная, то для распределений будет выполняться требование условной независимости:

$$p(\tilde{x}_1\tilde{x}_2\ldots\tilde{x}_n\tilde{z})p(\tilde{y}_1\tilde{y}_2\ldots\tilde{y}_m\tilde{z})=p(\tilde{x}_1\tilde{x}_2\ldots\tilde{x}_n\tilde{y}_1\tilde{y}_2\ldots\tilde{y}_m\tilde{z})p(\tilde{z}).$$

Для сходящейся связи при отсутствии свидетельств требование $d$-разделимости для двух непересекающихся множеств будет эквивалентно просто требованию независимости распределений.

Проанализируем множество означиваний $\tilde{y}_1\tilde{y}_2\ldots\tilde{y}_m$, соответствующее одному узлу дерева сочленений. Для множества переменных $\{y_1, y_2, \ldots, y_m\}$ в байесовской сети доверия есть связывающие их ребра. Рассмотрим следующее произведение вероятностей:

$$\prod_{i=1}^{m} p(\tilde{y}_i \mid \tilde{\wedge}\mathrm{pa}_Y(y_i)).$$

Здесь $\mathrm{pa}_Y(y_i)$ означает множество родителей, лежащих в анализируемом множестве (в одном узле дерева сочленений). Так как множество $\mathrm{pa}_Y(y_i)$ может не совпадать с множеством $\mathrm{pa}(y_i)$, то часть произведения окажется неопределенной. Если бы у нас было задано локальное распределение, то тогда на его основе можно было бы вычислить недостающие условные вероятности. Теперь предположим, что нам заданы эти условные вероятности. Они нам нужны только для проведения формальных вычислений: как будет показано позже, все условные вероятности, не входящие в байесовскую сеть доверия, сократятся. Считая, что вероятности вида $p(\tilde{y}_i \mid \tilde{\wedge}\mathrm{pa}_Y(y_i))$ заданы для всех $i$, определим локальное распределение как их произведение (напоминаем, что речь идет об одном конкретном означивании участвующих в распределении переменных). Покажем, что такое определение локального распределения соответствует условной независимости $d$-разделенных вершин. Для этого перепишем это произведение следующим образом:

$$\prod_{i=1}^{m} \frac{p(\tilde{\wedge}\mathrm{fam}_Y(y_i))}{p(\tilde{\wedge}\mathrm{pa}_Y(y_i))},$$

здесь $\mathrm{fam}_Y(y_i)$ означает $\mathrm{pa}_Y(y_i) \cup \{y_i\}$. Мы по-прежнему рассматриваем множество родителей, попадающих в один узел дерева сочленений. Начнем «сворачивать» это произведение «сверху», то есть начнем с элементов, у которых нет родителей в анализируемом множестве (в рассматриваемом узле дерева сочленений). Под сворачиванием мы понимаем процесс постепенного построения общего распределения. На первом шаге у нас есть ряд маргинальных вероятностей. На втором шаге мы можем рассчитать маргинальные распределения элементов, для чьих родителей у нас были заданы маргинальные вероятности. На каждом следующем шаге у нас будет ряд маргинальных распределений на непересекающихся множествах переменных. При этом каждое такое множество будет связным. Рассмотрим вершину, для которой все родители уже маргинализованы. Тогда возможны два случая.

*Случай 1.* Есть одно уже вычисленное маргинальное распределение вероятностей, которое задано на множестве, содержащем всех родителей (в узле дерева сочленений) рассматриваемой вершины. Тогда (пусть $y_1$ — общий сын $y_2, \ldots, y_k$)

$$\frac{p(\tilde{y}_1\tilde{y}_2\ldots\tilde{y}_k)}{p(\tilde{y}_2\ldots\tilde{y}_k)} p(\tilde{y}_2\ldots\tilde{y}_k\tilde{y}_{k+1}\ldots\tilde{y}_l) = p(\tilde{y}_1\ldots\tilde{y}_k\tilde{y}_{k+1}\ldots\tilde{y}_l).$$

Здесь мы воспользовались тем, что $\dfrac{p(\tilde{y}_1\tilde{y}_2\ldots\tilde{y}_k)}{p(\tilde{y}_2\ldots\tilde{y}_k)} = p(\tilde{y}_1 \mid \tilde{y}_2\ldots\tilde{y}_k)$, и тем, что $\tilde{y}_1$ условно независимо от множества $\tilde{y}_{k+1}\ldots\tilde{y}_l$ при известном $\tilde{y}_2\ldots\tilde{y}_k$. Это справедливо, так как от каждой вершины из $\{y_{k+1}, \ldots, y_l\}$ существует путь к вершине

$y_1$, проходящий через ее родителей. В таком случае связь на этом пути в родительской вершине либо последовательная, либо расходящаяся, а, значит, по $d$-разделимости выполняются требования условной независимости. Отдельно надо рассмотреть случай, когда $p(\tilde{y}_2 \ldots \tilde{y}_k) = 0$, но тогда, в силу согласованности распределений, $p(\tilde{y}_2 \ldots \tilde{y}_k \tilde{y}_{k+1} \ldots \tilde{y}_l)$ также равно нулю. И если положить $p(\tilde{y}_1 \ldots \tilde{y}_k \tilde{y}_{k+1} \ldots \tilde{y}_l) = 0$, то написанное выше произведение остается корректным и в этом особом случае.

*Случай 2.* Множество родителей содержится в нескольких непересекающихся множествах, для которых вычислены маргинальные распределения. Тогда, в силу $d$-разделимости, эти распределения должны быть независимы. Это связано с тем, что существует путь от одного множества к другому, проходящий через общего сына и, следовательно, в этом сыне будет сходящаяся связь. А раз эти распределения независимы, то можно определить общее распределение как их произведение. Определенное таким образом распределение будет задано над множеством, содержащим всех родителей рассматриваемой вершины. Теперь можно воспользоваться соображениями из случая 1 и построить совместное распределение.

На каждом шаге мы будем получать совместное распределение для большего множества переменных. Так как дерево конечное, то на одном из шагов мы закончим построение совместного распределения вероятностей.

## 6. Глобальное и локальные распределения

Построение глобального распределения можно было бы провести по только что описанной схеме. Если описанным выше образом постепенно вычислять распределение на всей сети, то мы получим распределение, которое является произведением всех условных вероятностей. В данном случае мы опишем другой способ получить глобальное распределение на основе локальных.

В основе этого способа лежит понятие условной независимости. Выпишем его в общем виде:

$$p(\tilde{x}_1 \ldots \tilde{x}_n \tilde{y}_1 \ldots \tilde{y}_m \tilde{z}_1 \ldots \tilde{z}_k) p(\tilde{y}_1 \ldots \tilde{y}_m) = p(\tilde{x}_1 \ldots \tilde{x}_n \tilde{y}_1 \ldots \tilde{y}_m) p(\tilde{y}_1 \ldots \tilde{y}_m \tilde{z}_1 \ldots \tilde{z}_k).$$

Здесь описана условная независимость $\tilde{x}_1 \ldots \tilde{x}_n$ относительно $\tilde{z}_1 \ldots \tilde{z}_k$ при известном означивании $\tilde{y}_1 \ldots \tilde{y}_m$. Если $p(\tilde{y}_1 \ldots \tilde{y}_m)$ не равно нулю, то можно переписать это соотношение в виде

$$p(\tilde{x}_1 \ldots \tilde{x}_n \tilde{y}_1 \ldots \tilde{y}_m \tilde{z}_1 \ldots \tilde{z}_k) = \frac{p(\tilde{x}_1 \ldots \tilde{x}_n \tilde{y}_1 \ldots \tilde{y}_m) p(\tilde{y}_1 \ldots \tilde{y}_m \tilde{z}_1 \ldots \tilde{z}_k)}{p(\tilde{y}_1 \ldots \tilde{y}_m)}.$$

Этой формулой мы и воспользуемся для построения глобального распределения на основе локальных. Если $p(\tilde{y}_1 \ldots \tilde{y}_m) = 0$, то по свойствам вероятностей получаем, что

$$p(\tilde{x}_1 \ldots \tilde{x}_n \tilde{y}_1 \ldots \tilde{y}_m \tilde{z}_1 \ldots \tilde{z}_k) = p(\tilde{x}_1 \ldots \tilde{x}_n \tilde{y}_1 \ldots \tilde{y}_m) = p(\tilde{y}_1 \ldots \tilde{y}_m \tilde{z}_1 \ldots \tilde{z}_k) = p(\tilde{y}_1 \ldots \tilde{y}_m) = 0.$$

В такой ситуации мы положим формально, что

$$\frac{0 \cdot 0}{0} = 0.$$

При таком предположении мы с, одной стороны, формально будем работать с произведением, а, с другой стороны, избежим проблемы деления на ноль. Да-



Труды СПИИРАН. 2006. Вып. 3. Т. 1. ISSN 2078-9181 (печ.), ISSN 2078-9599 (онлайн)
SPIIRAS Proceedings. 2006. Issue 3. V. 1. ISSN 2078-9181 (print), ISSN 2078-9599 (online)
www.proceedings.spiiras.nw.ru

лее мы не будем отдельно оговаривать случай деления на ноль, а будем его обрабатывать указанным образом.

Мы будем исходить из предположения, что локальные распределения соседних узлов дерева сочленений согласованы на сепараторах и условно независимы относительно них. Сепаратором мы называем вес ребра, соединяющего рассматриваемые узлы.

Построим совместное распределение для множества переменных, являющегося объединением весов двух соседних узлов дерева сочленений. Подставим в формулу условной независимости распределения, выраженные через условные вероятности. Получим:

$$p(\tilde{x}_1 \ldots \tilde{x}_n \tilde{y}_1 \ldots \tilde{y}_m \tilde{z}_1 \ldots \tilde{z}_k) =$$
$$= \frac{p(\tilde{x}_1 \mid \tilde{\wedge}\mathrm{pa}_1(x_1)) \cdot \ldots \cdot p(\tilde{x}_n \mid \tilde{\wedge}\mathrm{pa}_1(x_n)) \cdot p(\tilde{y}_1 \mid \tilde{\wedge}\mathrm{pa}_1(y_1)) \cdot \ldots \cdot p(\tilde{y}_m \mid \tilde{\wedge}\mathrm{pa}_1(y_m))}{p(\tilde{y}_1 \mid \tilde{\wedge}\mathrm{pa}_2(y_1)) \cdot \ldots \cdot p(\tilde{y}_m \mid \tilde{\wedge}\mathrm{pa}_2(y_m))} \times$$
$$\times p(\tilde{y}_1 \mid \tilde{\wedge}\mathrm{pa}_3(y_1)) \cdot \ldots \cdot p(\tilde{y}_m \mid \tilde{\wedge}\mathrm{pa}_3(y_m)) \cdot p(\tilde{z}_1 \mid \tilde{\wedge}\mathrm{pa}_3(z_1)) \cdot \ldots \cdot p(\tilde{z}_k \mid \tilde{\wedge}\mathrm{pa}_3(z_k)).$$

В этой формуле

$$\mathrm{pa}_1(x) = \mathrm{pa}(x) \cap \{x_1, \ldots, x_n, y_1, \ldots y_m\},$$
$$\mathrm{pa}_2(x) = \mathrm{pa}(x) \cap \{y_1, \ldots y_m\},$$
$$\mathrm{pa}_3(x) = \mathrm{pa}(x) \cap \{y_1, \ldots, y_m, z_1, \ldots, z_k\}.$$

Рассмотрим произвольный элемент $x$, лежащий в сепараторе. Для него возможны четыре разных расположения родителей. Разберем эти четыре случая:

1) элемент $x$ со всеми своими родителями попал в сепаратор. Тогда множества $\mathrm{pa}_1(x)$, $\mathrm{pa}_2(x)$, $\mathrm{pa}_3(x)$ и $\mathrm{pa}(x)$ совпадают, и можно сократить дробь на $p(\tilde{x} \mid \tilde{\wedge}\mathrm{pa}(x_n))$; при этом в числителе останется множитель $p(\tilde{x} \mid \tilde{\wedge}\mathrm{pa}(x_n))$;
2) элемент $x$ лежит в сепараторе, а все его родители попали во множество $\{x_1, \ldots x_n, y_1, \ldots y_m\}$. Тогда $\mathrm{pa}_2(x) = \mathrm{pa}_3(x)$, а $\mathrm{pa}_1(x) = \mathrm{pa}(x)$, и можно сократить дробь на $p(\tilde{x} \mid \tilde{\wedge}\mathrm{pa}_2(x_n))$; при этом в числителе останется множитель $p(\tilde{x} \mid \tilde{\wedge}\mathrm{pa}(x_n))$;
3) элемент $x$ лежит в сепараторе, а все его родители попали в множество $\{y_1, \ldots, y_m, z_1, \ldots, z_k\}$. Тогда $\mathrm{pa}_1(x) = \mathrm{pa}_2(x)$, а $\mathrm{pa}_3(x) = \mathrm{pa}(x)$, и можно сократить дробь на $p(\tilde{x} \mid \tilde{\wedge}\mathrm{pa}_2(x_n))$; при этом в числителе останется множитель $p(\tilde{x} \mid \tilde{\wedge}\mathrm{pa}(x_n))$;
4) элемент $x$ лежит в сепараторе, и не все его родители не попали во множество $\{x_1, \ldots x_n, y_1, \ldots y_m, z_1, \ldots, z_k\}$. Тогда или $\mathrm{pa}_1(x) = \mathrm{pa}_2(x)$, или $\mathrm{pa}_3(x) = \mathrm{pa}_2(x)$, и можно сократить дробь на $p(\tilde{x} \mid \tilde{\wedge}\mathrm{pa}_2(x_n))$. В числителе останется множитель, соответствующий узлу, содержащему большее число родителей элемента $x$.

Этот список полностью исчерпывает возможные варианты расположения родителей элемента из сепаратора. Так как любой вес (ребра или вершины), лежащий на пути от одной вершины к другой, содержит пересечение весов этих вершин (из определения дерева смежности и соответственно дерева сочленений), то условные вероятности, не сократившиеся в случае 4, сократятся на одном из следующих шагов. Кроме того, случаи 1, 2 и 3 гарантируют то, что в числителе останется $p(\tilde{x} \mid \tilde{\wedge}\mathrm{pa}(x_n))$ (на одном из последующих шагов объедине-



Труды СПИИРАН. 2006. Вып. 3. Т. 1. ISSN 2078-9181 (печ.), ISSN 2078-9599 (онлайн)
SPIIRAS Proceedings. 2006. Issue 3. V. 1. ISSN 2078-9181 (print), ISSN 2078-9599 (online)
www.proceedings.spiiras.nw.ru

ния). Таким образом, мы получаем, что, с одной стороны, каждая из вероятностей вида $p(\widetilde{x} \mid \widehat{\ }\mathrm{pa}(x_n))$ попадет в числитель (так как есть узел, где лежат все родители $x$), а, с другой стороны, все вероятности вида $p(\widetilde{x} \mid \widetilde{W})$, не заданные в исходной сети, сократятся. Следовательно, восстановленное подобным образом (по данным из дерева сочленений) распределение совпадет с задаваемым традиционным образом, т. е. произведением тензоров условных вероятностей в байесовской сети доверия.

## 7. Выводы

Мы рассмотрели байесовские сети доверия и соответствующие им деревья сочленений с точки зрения вероятностной семантики. В качестве вероятностной семантики мы выбрали совместное распределение вероятностей над множеством всех переменных сети. Распределение в байесовской сети доверия мы рассчитываем на основе традиционного для БСД правила цепи (chain rule [3]), которое гласит, что совместное распределение равно произведению всех тензоров условных вероятностей, заданных в БСД. Для дерева сочленений глобальное распределение строится на основе предположения условной независимости локальных распределений, соответствующих узлам, относительно сепараторов.

Мы показали, что если локальные маргинальные распределения, заданные на дереве сочленений, определяют те же условные вероятности, которые заданы в байесовской сети доверия, и для соседних узлов дерева сочленений распределения согласованы по сепаратору, то совместное распределение, построенное на основе предположения условной независимости, совпадет с вероятностным распределением, построенным на основе байесовской сети доверия.

Эти результаты открывают перед нами возможности для дальнейшего сравнительного анализа байесовских сетей доверия и алгебраических байесовских сетей через представление последних в виде деревьев смежности [12].

### Литература


1. *Pearl J.* How to Do with Probabilities what People Say You Can't // Artificial Intelligence Applications / Ed. Weisbin C.R., IEEE, North Holland.1985. P. 6–12.
2. *Pearl J.* Probabilistic Reasoning in Intelligent Systems. San Francisco: Morgan Kaufmann Publishers, 1988. 552 p.
3. *Jensen F. V.* Bayesian Networks and Decision Graphs. New York: Springer-Verlag, 2001. 268 p.
4. *Nilsson N. J.* Probabilistic Logic // Artificial Intelligence. 1986. Vol. 47. Amsterdam: Elsevier Science Publishers B.V., 1986. P. 71–87.
5. *Nilsson N. J.* Probabilistic Logic Revisited // Artificial Intelligence. 1993. Vol. 59. Amsterdam: Elsevier Science Publishers B.V., 1993. P. 31–36.
6. *Тулупьев А. Л.* Алгебраические байесовские сети. Логико-вероятностный подход к моделированию баз знаний с неопределенностью. СПб.: СПИИРАН, 2000. 292 с.
7. *Cowell R. G., Dawid A. P., Lauritzen S. L., Spiegelhalter D. J.* Probabilistic Networks and Expert Systems. NY.:Springer-Verlag, 1999.
8. *Korb K. B., Nicholson A. E.* Bayesian Artificial Intelligence. New York: Chapman and Hall/CRC, 2004. 364 p.
9. *Kschischang, F., Frey B.* Iterative decoding of compound codes by probability propagation in graphical models. IEEE Journal on Selected Areas in Communications. 1998. Vol. 16-2. P. 219–230.

10. *MacKay D. J., McEliece R. J., Cheng J. F*. Turbo decoding as an instance of pearl's belief propagation algorithm // IEEE Journal of Selected Areas of Communication, February 1998. P. 140–152.
11. *Tulupyev A. L., Nikolenko S. I.* Directed Cycles in Bayesian Belief Networks: Probabilistic Semantics and Consistency Checking Complexity // MICAI 2005: Advances in Artificial Intelligence**.** 4th Mexican International Conference on Artificial Intelligence, Monterrey, Mexico, November 14–18, 2005, Proceedings Series: Lecture Notes in Computer Science; Subseries: Lecture Notes in Artificial Intelligence, Vol. 3789. / Gelbukh, Alexander; Terashima, Hugo (Eds.) 2005, XXVI. P. 214–223.
12. *Тулупьев А. Л.* Дерево смежности с идеалами конъюнктов как ациклическая алгебраическая байесовская сеть // Труды СПИИРАН. 2006, Вып. 3, т. 1. СПб.: Наука. [В настоящем томе.]