УДК 004.8

А.А. Фильченков, А.Л. Тулупьев, А.В. Сироткин

ОСОБЕННОСТИ АНАЛИЗА ВТОРИЧНОЙ СТРУКТУРЫ АЛГЕБРАИЧЕСКОЙ БАЙЕСОВСКОЙ СЕТИ

Фильченков А.А., Тулупьев А.Л., Сироткин А.В. **Особенности анализа вторичной структуры алгебраической байесовской сети.**
Аннотация. Цель данной работы — обобщение результатов структурного анализа минимальных графов смежности, представляющих вторичную структуру алгебраической байесовской алгебраической сети, на графы смежности общего вида, представляющие эту же структуру. Сформулирована система терминов, расширяющая существующую систему для МГС на графы смежности в целом. Исследованы новые свойства графов смежности. Сформулированы и доказаны две леммы, характеризующие оммаж (результат сжатия минимального графа смежности) как минимальную курию (результат сжатия графа смежности). Упрощено доказательство теоремы о множестве минимальных графов смежности.
Ключевые слова: алгебраические байесовские сети, вторичная структура, графы смежности, автоматическое обучение, структурный синтез.

Filchenkov A.A., Tulupyev A.L., Sirotkin A.V. **Algebraic Bayesian networks secondary structure analysis features.**
Abstract. The goal of this work is to generalize structure analysis of minimal join graphs representing algebraic Bayesian network secondary structure on common join graphs representing the same structure. A term system spreading the existing one with common join graphs is designed. New join graph properties are researched. Two lemmas characterizing an homage (product of a minimal join graph compression) as a minimal curia (product of a join graph compression) are proven. Theorem on minimal join graph proof is simplified.
Keywords: algebraic Bayesian networks, secondary structure, join graphs, automated learning, machine learning, structure synthesis.

1. Введение.Алгебраические байесовские сети (АБС) представляют собой логико-вероятностную графическую модель [1–2, 14, 19] сложных систем знаний с неопределенностью. В теории АБС неопределенность знаний характеризуется с помощью оценок вероятностей (вероятность задается на пространстве пропозициональных формул, через которые мы представляем утверждения; это позволяет говорить о вероятности истинности). В классе вероятностно-графических моделей алгебраические байесовские сети,как и байесовские сети доверия, позволяют осуществить синтез согласованных оценок и пропагацию свидетельств, а также поддерживают непротиворечивость [8, 15].Выделяют первичную и вторичную структуру АБС [6, 14, 16].

Выбор вторичной структуры АБС является существенным для эффективности логико-вероятностного вывода всех видов, более того, он обусловливает саму возможность таких выводов, потому что они для определенного класса вторичных структур вообще неосуществимы [10–13, 15]. Одним из вариантов представления вторичной структуры одной и той же АБС являются графы смежности и их семейства [4–7]. С точки зрения потребностей теории АБС, особенно важным является исследование минимальных графов смежности [6]. В частности, вторичная структура любой ациклической алгебраической сети, важность которых была показана в [9], представима минимальным графом смежности. Но свойства этих графов изучены недостаточно. Установлено только, что множество минимальных графов смежности допускает матроидное представление[3], а также через подграфы максимального графа смежности[17]. При этом необходимо более глубокое понимание свойств графов смежности.

Последняя из указанных работ исследует множества минимальных графов смежности, вводя ряд операций на подграфах графов смежности. Основным результатом работы является теорема о множестве минимальных графов смежности, которая утверждает, что множество минимальных графов смежности является декартовым произведением множеств особых подграфов максимального графа смежности. Этот результат является теоретической основой для алгоритмов синтеза указанного множества, а также для дальнейшего исследования свойств минимальных графов смежности.

Цель данной работы — обобщить и уточнить результаты работы [17], связанные с теоремой о множестве минимальных графов смежности, а также уточнить некоторые термины, сделав их более легкими для понимания и удобными в применении. Указанные обобщения и уточнения позволяют получить новые результаты о минимальных графах смежности, упорядочить исследуемую область и сделать ее более доступной для улучшения освоения и применения.

2. Основные определения и обозначения. В данном разделе приводятся определения так, как они были введены в [17].

Граф — пара $\langle V, E\rangle$, где V — множество вершин графа, а E — множество ребер, каждое из которых является неупорядоченной парой (v_i, v_j), $i \neq j$, v_i, $v_j \in V$. Для удобства будем через V и E обозначать функции от графа, возвращающие множество его вершин и множество его ребер соответственно:

$$V(G') = V';\ E(G') = E', \qquad \text{где } G' = \langle V', E'\rangle.$$

Внесем дополнительную определенность: $\subseteq$ — (нестрогое) включение:

$$(A \subseteq B) \Leftrightarrow (\forall a \in A \quad a \in B);$$

$\subset$ — строгое включение:

$$(A \subset B) \Leftrightarrow (\forall a \in A, a \in B \ \ \exists b \in B: b \notin A).$$

Алфавитом будем называть множество атомарных пропозициональных формул $A = \{x_1, \dots, x_n\}$, над которым будут заданы максимальные фрагменты знаний. *Слово* V — подмножество алфавита: $V \subseteq A$.

Множество главных конъюнктов максимальных фрагментов знаний (МФЗ), вошедших в АБС, — это такое множество слов $V^* = \{V_i \subseteq A\}_{i=1}^{i=m}$, что:

1) оно не содержит несобственное подмножество алфавита:

$$V_i \neq A, V_i \neq \emptyset;$$

2) никакое слово полностью не содержит никакое другое слово:

$$\forall\ i \neq j\ (V_i \not\subset V_j) \text{ и } (V_j \not\subset V_i).$$

Граф максимальных фрагментов знаний — ненаправленный граф G, вершины которого соответствуют элементам множества главных конъюнктов МФЗ, вошедших в алгебраическую байесовскую сеть, а ребра удовлетворяют условию

$$\left(V_i, V_j\right) \in E(G) \Rightarrow V_i \cap V_j \neq \emptyset.$$

Введем понятия веса для вершины, для ребра и для подграфа. *Вес* $W(V_i)$ *вершины* $V_i \in V(G)$ — множество атомов алфавита, вошедших в V_i:

$$W(V_i) = \{x_i | x_i \in V\}.$$

Вес $W\left(\{V_i, V_j\}\right)$ *ребра* $\{V_i, V_j\} \in E(G)$ графа G определяется как пересечение весов тех вершин, которые соединены этим ребром:

$$W\left(\{V_i, V_j\}\right) = W(V_i) \cap W\left(V_j\right).$$

Вес $W(H)$ *подграфа* $H \subseteq G$ — наибольшее по включению слово, которое входит в веса всех его вершин:

$$W(H) = \bigcap_{V \in H} W(V).$$

Магистральный путь $B: V_b \dashrightarrow V_e$ от вершины V_b до вершины V_e, пересечение весов которых непусто, — это такой путь от вершины V_b до вершины V_e, что вес любой принадлежащей ему вершины содержит пересечение весов начальной и конечной вершин. $B: V_b \dashrightarrow V_e = P: V_b \rightsquigarrow V_e$, такой, что

$$\forall V_i \in B\ W(V_b) \cap W(V_e) \subset W(V_i).$$

Граф *магистрально связен*, если между каждой парой несовпадающих вершин, веса которых содержат общие элементы, существует магистральный путь. Будем обозначать множество магистрально связных графов через **BCG.**

Благодаря введенным понятиям *граф смежности* определяется как магистрально связный граф МФЗ. При этом магистрально связный граф не обязательно связен (например, граф из двух вершин ab и cd, в котором нет ребер, тем не менее, является магистрально связным).

Минимальный граф смежности — граф смежности, число ребер в котором минимально (рис. 1). Будем обозначать множество минимальных графов смежности через $\mathbf{MJG}$.

Максимальный граф смежности $G_{\max}$ — наибольший по числу ребер граф смежности (рис. 1). Так как в графе МФЗ возможны не все ребра, а только те, которые соединяют вершины, пересечение весов которых непусто, то максимальный граф смежности вовсе необязательно совпадает с полным подграфом. В [14] было доказано, что для заданного множества вершин существует и при этом единственный максимальный граф смежности.

3. Клики, сужения и владения. В данном разделе приводятся определения так, как они были даны в [18].

Сужение $G \downarrow U$ ненаправленного графа G на слово U — это ненаправленный граф, в который входят только те вершины и ребра исходного графа G, веса которых содержат или равны U:

$$G \downarrow U = \langle \{V_i | V_i \in V(G), U \subseteq W(V_i)\}, \{E_i | E_i \in E(G), U \subseteq W(E_i)\} \rangle.$$

Любое слово, являющееся весом какого-либо ребра графа G, будем называть *значимым словом графа* G, а сужение графа G на такое слово — *значимым сужением*. Для любого значимого сужения $G \downarrow U$ выполняется $W(G \downarrow U) = W(U)$.

Кликой U будем называть значимое сужение максимального графа смежности на вес U (рис. 2). Вес любой клики является значимым словом, а сама она является полным подграфом графа $G_{\max}$. Множество всех таких клик будем обозначать как Clique.

В [17] было доказано, что если сужение $G_{\max}$ на произвольное слово непусто, то оно является максимальным подграфом. Отсюда следует, что клика является полным подграфом, как ее обычно и понимают в теории графов.

Граф клик — направленный граф, вершинами которого являются клики из множества Clique (рис. 3). Данный граф клик отличен от того, что иногда под ним могут понимать — от графа максимальных клик. Ребро из вершины P в вершину Q проведено, если клика P содержит клику Q, и не существует клики R, такой, что клика P содержит клику R и клика R содержит клику Q: G_{Clique} — пара $\langle \text{Clique}, E_{\text{Clique}} \rangle$, где

$$\forall P, Q \in \text{Clique}, (P, Q) \in E \Leftrightarrow Q \subset P \text{ и } \nexists R \in \text{Clique}: Q \subset R \subset P.$$

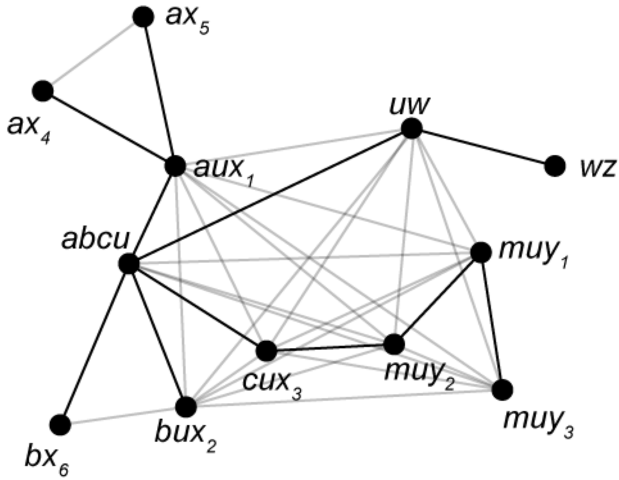

Рис. 1. Максимальный граф смежности
и минимальный граф смежности.
Графы смежности, построенные на вершинах $abcu$,ax_4, ax_5, aux_1, bux_2, cux_3, muy_1, muy_2, muy_3,uw, wz. Минимальный граф смежности состоит из черных ребер, максимальных граф смежности — из черных и серых.

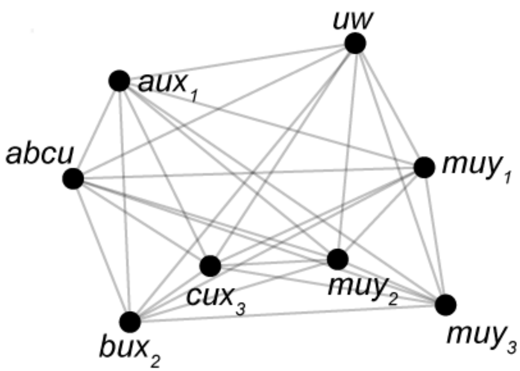

Рис. 2. Кликаa.
Клика, образованная сужением максимального графа с рис. 1 на вес u.

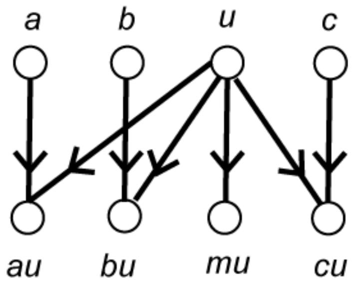

Рис. 3. Граф клик.
Граф клик, построенный для графа на рис. 1.

Сильное сужение $G \not\downarrow U$ — значимое сужение $G \downarrow U$, из которого удалили все ребра веса U:

$$G \not\downarrow U = \langle \{V_i | V_i \in V(G), U \subseteq W(V_i)\}, \{E_i | E_i \in E(G), U \subset W(E_i)\}\rangle.$$

Сильное сужение графа $G_{\max} \not\downarrow U$ представляет собой компоненты связности, на которые разбивается сужение $G_{\max} \downarrow U$ удалением ребер веса U (рис. 4). *Владение* P_U^i — множество вершин i-й компоненты связности сильного сужения $G_{\max} \not\downarrow U$ (рис. 4).

В работе [18] была доказана *теорема о классификации владений*, которая утверждала, что любое владение является либо доменной вершиной, либо вассалом, либо братством (эти три термина мы определим ниже).

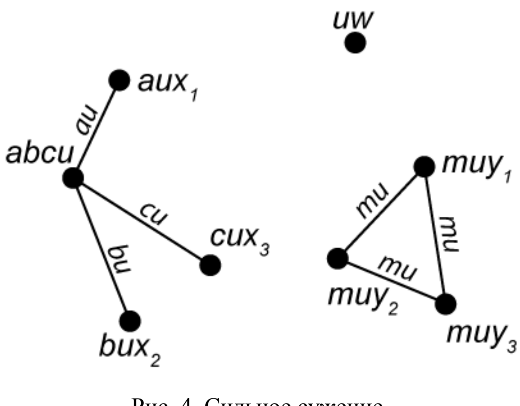

Рис. 4. Сильное сужение.
Сильное сужение графа с рис. 1 на вес u. Оно состоит из 3х владений.

Доменная вершина D_U клики U — вершина, принадлежащая клике U и не принадлежащая ни одному из ее сыновей (рис. 5).

Вассал V_U клики U — множество вершин, входящих в какого-либо сына клики U(рис. 5).

Два вассала V_U^i и V_U^j называются *братьями*, если их пересечение непусто(рис. 5):

$$V_U^i \leftrightarrow V_U^j \Leftrightarrow V_U^i \cap V_U^j \neq \emptyset.$$

Два вассала V_U^i и V_U^j называются *полусиблингами*, если существует такой упорядоченный набор вассалов $\{V_U^{w_1}, \dots V_U^{w_n}\}$, что V_U^i — брат $V_U^{w_1}$, $V_U^{w_1}$ — брат $V_U^{w_2}$, …, $V_U^{w_{n-1}}$ — брат $V_U^{w_n}$, а $V_U^{w_n}$ — брат V_U^j:

$$V_U^i \leftrightarrow V_U^j \Leftrightarrow \exists\{V_U^{w_1}, \dots V_U^{w_n}\}: V_U^b \leftrightarrow V_U^{w_1}, V_U^{w_n} \leftrightarrow V_U^e$$
$$\text{и } \forall i < n \;\; V_U^{w_i} \leftrightarrow V_U^{w_{i+1}}.$$

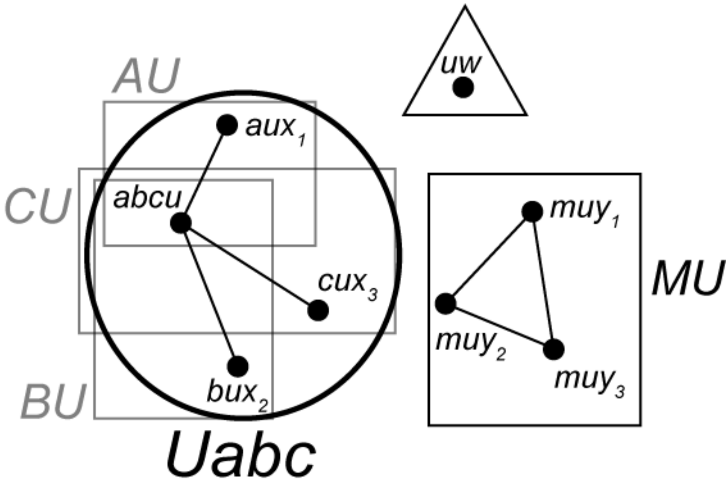

Рис. 5. Владения.

Треугольником обозначена доменная вершина uw; прямоугольником обозначен вассал MU, состоящий из вершин muy_1, muy_2 и muy_3; овалом обозначено братство $Uabc$, состоящее из выделенных полупрозрачными прямоугольниками вассалов AU, BU и CU; вассалы AU, BU и CU приходятся друг другу братьями.

Полусиблинговый путь между двумя родственными вассалами V_U^i и V_U^j — такой набор вассалов $\{V_U^{w_1}, \dots V_U^{w_n}\}$ из определения 6’, что $V_U^b \leftrightarrow V_U^{w_1}$; $V_U^{w_n} \leftrightarrow V_U^e$ и $\forall i < n V_U^{w_i} \leftrightarrow V_U^{w_{i+1}}$.

Братство B_U клики U — непустой набор вассалов $\{V_U^1, \dots V_U^l\}$ клики U, такой, что с каждым вассалом в братство входят все его полусиблинги и только они (рис. 5):

$$B_U = \{V_U^i | (V_U^i \in B_U) \& (V_U^i \leftrightarrow V_U^j) \Rightarrow V_U^j \in B_U; V_U^i, V_U^j \in B_U \Rightarrow$$
$$\Rightarrow V_U^i \leftrightarrow V_U^j\}.$$

4. Множество минимальных графов смежности. В утверждении 3.1 из [17] говорилось, что для любого минимального графа смежности M его сужение на произвольное значимое слово U связно; следствие 3.1 из него утверждало, что это сужение магистрально связно; лемма 3.2 утверждала, что если любое значимое сужение магистрального графа магистрально связно, то оно является графом смежности. Мы можем обобщить все эти факты в одном утверждении.

Утверждение 1. G — граф смежности тогда и только тогда, когда его произвольное значимое сужение $G \downarrow U$ магистрально связно.

Доказательство. 1) $\Rightarrow$. Любые две вершины v_1, v_2 из сужения $G \downarrow U$ в самом графе G магистрально связны. Пускай $U' = W(v_1) \cap W(v_2)$. Так как $U \subseteq W(v_1)$ и $U \subseteq W(v_2)$, то $U \subseteq U'$. Рассмотрим связывающий эти вершины магистральный путь $B: v_1 \rightsquigarrow v_2$.

$$\forall e_i \in B \;\; U' \subseteq W(e_i) \Rightarrow U \subseteq W(e_i) \Rightarrow e_i \in G \downarrow U.$$

Так как все ребра магистрального пути лежат в сужении, значит, v_1 и v_2 магистрально связны. Поскольку мы рассмотрели произвольные вершины, то все сужение магистрально связно.

2)$\Leftarrow$. Возьмем любые две вершины v_1 и v_2, у которых пересечение весов не пусто и равно U. Они обе попадают в сужение графа на U. В этом сужении между v_1 и v_2 существует магистральный путь, а, значит, этот же магистральный путь связывает их в самом графе, а значит, граф является графом смежности.

Пример 1. Давайте рассмотрим граф смежности, изображенный на рисунке 6.а. Его сужение на вес uизображено на рисунке 6.б с выделенными владениями (братством, вассалом и доменной вершиной). Для дальнейшего анализа удобно изменять сужение, представляя владения в виде вершин и соединяя их ребрами в том случае, если сами владения соединены.

В статье [17] мы ввели операцию сжатия (определение 3.7), которая вводилась на графе для его подграфа. Смысл был в том, чтобы каждое владение P_U^i в одну вершину f_i, так, что ребра, выходящие из P_U^i и ведущие к другим владениям, станут ребрами, выходящими из f_i и ведущими к другим вершинам. Однако подобное определение было достаточно неудобно: граф, состоящий из сжатых компонент связности, представлял собой композицию многих сжатий (то есть композицию сжатий для каждого такого компонента связности). При подготовке спецификаций на разработку программного обеспечения такой подход оказался неудобным.

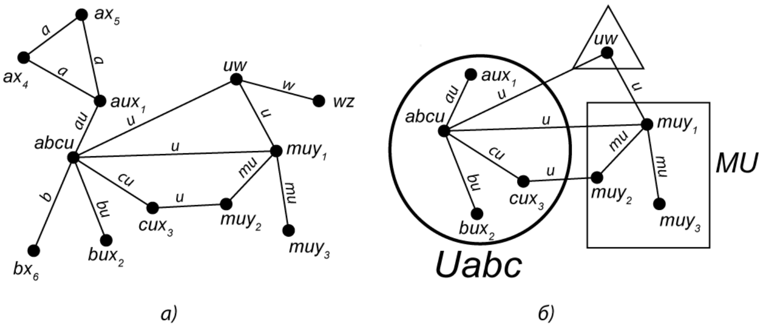

Рис. 6. Граф смежности.
На рис. 6.а изображен некоторый граф смежности над вершинами$abcu$, aux_1, bux_2, cux_3, ax_4, ax_5, bx_6, muy_1, muy_2, muy_3, uw, wz. На рис. 6.б изображено сужение этого графа на вес uс выделенными владениями.

Еще одним недостатком подобного определения было то, что оно недостаточно формально описывало процессы, происходящие с элементами графа: его вершинами и ребрам.

Наконец, данное определение сжатия не учитывало, число ребер между двумя владениями. Так, в старом определении подграф D, состоящий из двух владений, между каждой парой вершин, выбранных по одной из каждого владения, проведено ребро, и подграф E, состоящий из тех же двух владений, но с единственным ребром, соединяющим два этих владения, сжимались бы до одного и того же графа. Поскольку речь идет о минимальных подграфах, то, как будет показано в лемме 1, кратность ребер несущественна, поскольку она всегда равна единицы. Однако же подобное упрощение не позволяло нам обобщить операцию сжатия до сжатия на графах смежности вообще, а не только на минимальных графах смежности.

Учитывая все вышеизложенные замечания, мы предложим новое определение (а, точнее, новое семейство определений) сжатия. Число ребер, связывающих два владения, будет выражено в кратности ребра, связывающего две соответствующие вершины.

Определение 1. *Сжатие*σ_U компоненты связности $P_U^i \subseteq\subseteq G \not\downarrow U$ в вершину f_i — отображение на множестве подмножеств вершин, сопоставляющее множеству вершин P_U^i вершину f_i (рис. 7).

Определение 2.*Сжатие* σ_U множества ребер

$$E_{i,j} \subset (G \downarrow U) \backslash (G \not\downarrow U),$$

соединяющих владения P_U^iи P_U^j в ребро $e_{i,j}$ — отображение на множестве подмножеств ребер, сопоставляющее множеству ребер $E_{i,j}$ ребро $e_{i,j}$, соединяющее вершины$f_i = \sigma_U(P_U^i)$ и $f_j = \sigma_U(P_U^j)$ и имеющее кратность, равную $|E_{i,j}|$ (рис. 8).

Определение 3. *Сжатие* σ_U графа смежности G в граф K_U — отображение на множестве графов, сопоставляющий графу G, являющему графом смежности, граф K_U, вершинами которого являются владения сильного сужения $G \not\downarrow U$, а ребро между двумя вершинами f_1 и f_2 графа K_U существует, если существует ребро в графе G между вершинами, принадлежащими соответствующим f_1 и f_2 владениям P_U^1 и P_U^2. Кратность такого ребра (f_1, f_2) равна числу всех ребер, соединяющих вершины изP_U^1 и P_U^2 (рис. 9).

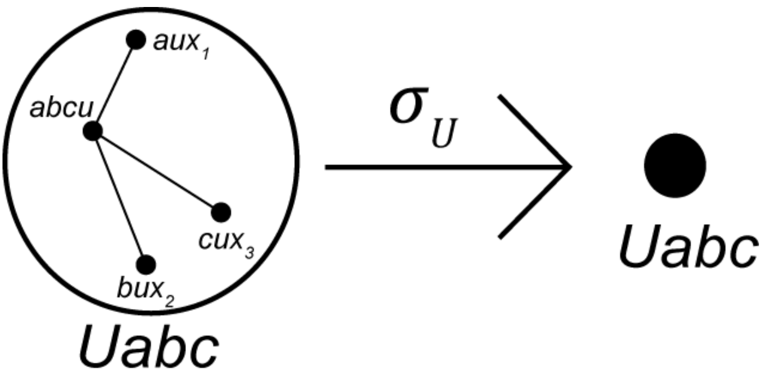

Рис. 7. Сжатие владения.
Сжатие братства $Uabc$ в вершину.

То же самое, но более формально:

Определение 3'. *Сжатие* $\sigma_U: G \to G, \sigma_U(G) = K_U,$ где $K_U = = \langle F_U, E_U, d(e) \rangle$, такое, что:

1) $F_U = \{f_i | f_i = \sigma_U(P_U^i), P_U^i \subseteq V(G \not\downarrow U)\}$;

2) $E_U = \{e_i | e_i = \sigma_U(E_{i,j}); E_{i,j} \subset (G \downarrow U) \backslash (G \not\downarrow U)\}$;

3) $d: E_U \to \mathbb{N}: d\left(\sigma_U(E_{i,j})\right) = |E_{i,j}|$.

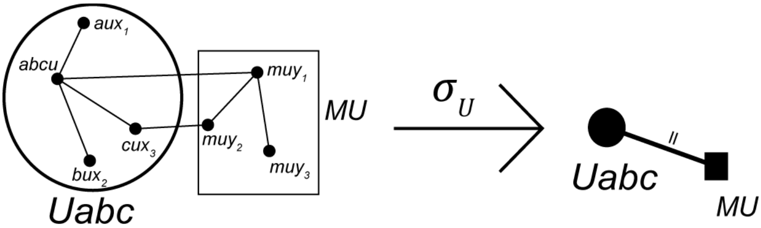

Рис. 8. Сжатие множества ребер.
Сжатие множества ребер, соединяющих владения $Uabc$ и MUв ребро кратности 2, соединяющее соответствующие вершины.

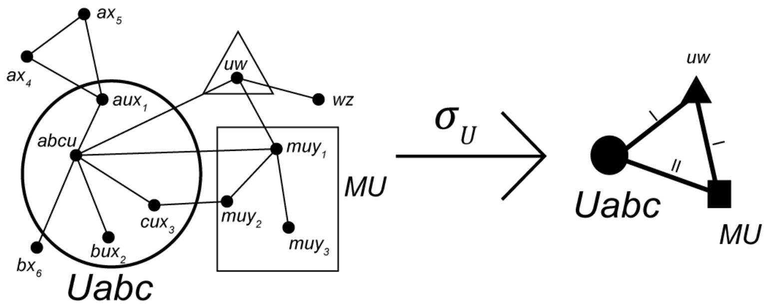

Рис. 9. Сжатие графа.
Сжатие графа, изображенного на рис. 6.

Нетрудно заметить, что новое сжатие и старое сжатие, примененные к минимальным графам смежности, дадут один и тот же результат. Данное положение будет формально обосновано ниже в утверждении 2.

Определение 4. *Феод* — f_i из определения 1 — вершина, получившаяся сжатием какого-то владения (рис. 7).

По сути, данное определение ничем не отличается от определения 3.8 из статьи [17], однако оно основано на другом понимании операции сжатия. Благодаря тому, что мы изменили операцию сжатия, став учитывать кратность ребер, мы можем ввести новый термин, который бы отвечал за получаемый в ходе сжатия результат.

Определение 5. *Курия веса* U — K_U из определения 3 — ненаправленный граф с кратными ребрами, полученный сжатием значимого сужения $G \downarrow U$ (рис. 9).

Учет кратности заставляет нас изменить и определение оммажа, который, по сути, остается ровно тем, чем он был в определении 3.9 из статьи [17].

Определение 6. *Оммаж*H_U — курия K_U, являющаяся деревом, все ребра которой имеют кратность, равную единице (рис. 10).

Поскольку мы решили идти по пути обобщения и ввели операцию сжатия не только на минимальных графах смежности, но и на графах смежности вообще, нам необходимо установить, как соотносятся между собой курии и оммажи. В работе [17] мы считали оммаж деревом, которое получается при сжатии минимального графа смежности. Теперь необходимо показать, что оммажи являются в определенном смысле «минимальными» куриями.

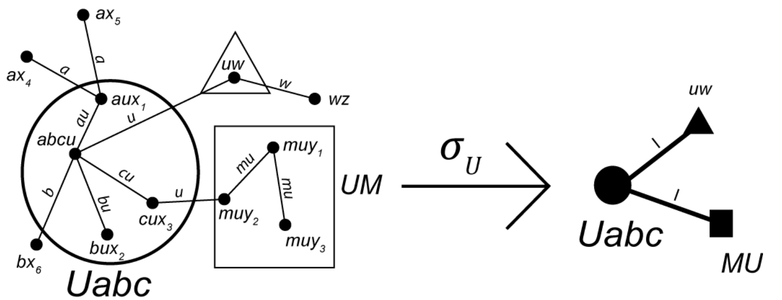

Рис. 10. Оммаж.
Сжатие минимального графа, построенного над теми же вершинами, что и на рис. 6 до оммажа.

Лемма 1. Пускай в какой-то курии веса U графа смежности G есть ребро кратности не меньше двух. Тогда удаление любого из ребер графа G, сужаемого до кратного ребра, сохранит его магистральную связность:

$$\exists e \in E\big(\sigma_U(G)\big): d(e) \geq 2 \Rightarrow \forall b \in E(G): \sigma_U(b) = e \;\; G^- =$$
$$= \langle V(G), E(G) \backslash \{b\} \rangle \in \mathbf{BCG}.$$

Доказательство. Пускай $b = (v_1, v_2), v_1 \in P_U^1, v_2 \in P_U^2$. Рассмотрим какое-нибудь ребро $a = (v_1^*, v_2^*)$: $\sigma_U(a) = e$ и $a \neq b. v_1^* \in P_U^1, v_2^* \in P_U^2$. Рассмотрим какой-нибудь магистральный путь, проходивший через это ребро:

$$W: p \dashrightarrow q; W = W_p^1 \cup (v_1, v_2) \cup W_2^q.$$

Его вес U^* был таков, что $U^* \subseteq U$, потому что вес самого удаленного ребра был U. Рассмотрим путь W^*:

$$W^* = W_p^1 \cup (W_1^*: v_1 \dashrightarrow v_1^*) \cup (v_1^*, v_2^*) \cup (W_2^*: v_2^* \dashrightarrow v_2) \cup W_2^q.$$

$W_1^* \subseteq P_U^1; W_2^* \subseteq P_U^2$,поэтому веса всех вершин, им принадлежащих, содержат вес U, поэтому $W^*: p \dashrightarrow q$.

Лемма 2. Пускай в какой-то курии веса U графа смежности G есть цикл. Тогда удаление любого из ребер графа G, сужаемого до ребра, принадлежащего этому циклу, сохранит магистральную связность графа:

$$\exists C \subseteq E\big(\sigma_U(G)\big): C \text{ — цикл} \Rightarrow \forall b \in E(G): \sigma_U(b) \in C \;\; G^- =$$
$$= \langle V(G), E(G) \backslash \{b\} \rangle \in \mathbf{BCG}.$$

Доказательство.$C = (f_1, \ldots, f_m)$, где $f_i = \sigma_U\big(P_U^i\big)$ — феоды. Предположим, что мы хотим удалить ребро $b = (v_1^m, v_m^1)$, где $v_1^m \in P_U^1$ и $v_m^1 \in P_U^m$. Рассмотрим какой-нибудь магистральный путь, проходивший через это ребро: $W = W_p^1 \cup (v_1^m, v_m^1) \cup W_m^q$. Его вес U^* был таков, что $U^* \subseteq U$, потому что вес самого удаленного ребра был U. Рассмотрим путь W_1^m, состоящий из следующих участков: магистральный путь от вершины $v_1^m \dashrightarrow v_1^2$ (внутри владения P_U^1, его вес содержит U), ребро (v_1^2, v_2^1) (его вес равен U), магистральный путь $v_2^1 \dashrightarrow v_2^3$, ребро$(v_2^3, v_3^2)$, … ребро (v_{m-1}^m, v_m^{m-1}), магистральный путь $v_m^{m-1} \dashrightarrow v_m^1$. Веса всех вершины предложенного пути будут содержать вес U, и, следовательно, U^*, поэтому путь $W^* = W_a^1 \cup W_1^m \cup W_m^b$ будет магистральным путем.

Благодаря доказанным леммам мы можем окончательно сформулировать и формально доказать утверждение о взаимосвязи между минимальным графом смежности и оммажем, которое было использовано в статье [17], но требует формального доказательства в новой системе терминов.

Утверждение 2. Для любого минимального графа смежности M и любого значимого слова U курия$K_U = \sigma_U(M)$ является оммажем.

Доказательство. Если курия не является оммажем, то в ней есть кратные ребра или циклы. Но тогда, по лемме 4.1 или по лемме 4.2 соответственно, мы можем удалить из графа M ребро так, чтобы магистральная связность сохранилась. Так как число ребер уменьшилось, то граф M не был минимальным.

Теперь мы несколько по-другому определим жилу. В работе [17] жила понималась как граф, тогда как в настоящей работе мы определим ее как множество ребер. Это нужно для более удобного понимая пучков.

Определение 7.Жила S_U — множество ребер графа смежности G, такое, что $E_u = \{\sigma_U(e) | e \in S_U\}$ является множеством ребер оммажа сжатия σ_U (рис. 11).

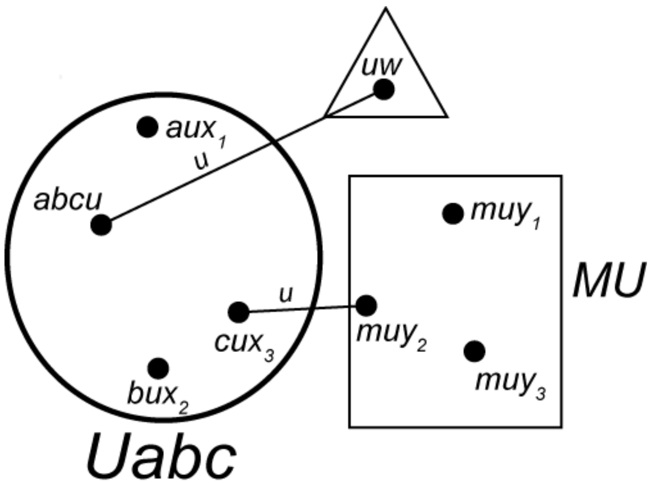

Рис. 11. Жила.
Жила веса U, соответствующая оммажу на рис. 10.

Замечание 1. Жила является элементом множества $E(\sigma_U^{-1}(H_u))$.

В работе [17] утверждение 3.2 говорило, что ребра любой жилы любого значимого слова Uимеют вес U; следующее за ним утверждение 3.3 состояло в том, что для любого минимального графа смежности Mв произвольном сужении на вес U множество ребер веса Uявляется жилой. Мы можем объединить эти утверждения в одно, а также улучшить доказательства, используя достижения работы [18] (в частности, теорему о классификации владений).

Утверждение 3. В жилу S_U входят те и только те ребра минимального графа смежности M, вес которых равен U.

Доказательство. Докажем, что вес любого ребра жилы равен U. Действительно, если вес какого-то ребра отличен от U, то он входит в $M \downarrow U$, поэтому будет соединять вершины одного владения, и, следовательно, не сможет быть ребром, соединяющим вершины из разных владений.

Теперь докажем, что в минимальном графе смежности множество ребер веса U сжимается до оммажа. Из утверждения 2 следует, что все ребра веса U, соединяющие вершины различных владений, сжимаются до оммажа. Докажем, что в минимальном графе смежности не существует ребер веса U, соединяющих вершины внутри какого-нибудь владения P_U^i. Предположим, что такое ребро b нашлось. Из теоремы о классификации владений следует, что владение является либо доменной вершиной, либо вассалом, либо братством.

P_U^i, очевидно, не может быть доменной вершиной. P_U^i не может проходить внутри вассала, иначе бы вес b совпадал с весом сына клики U, соответствующем вассалу. Значит, P_U^i соединяет вершины двух вассалов V_1 и V_2, содержащихся в одном братстве. Однако тогда граф M не будет минимальным, потому что при удалении ребра b братство останется связным, так как все вассалы, лежащие в полусиблинговом пути от V_1 и V_2 магистрально связны по утверждению 1, а, значит,V_1 и V_2 магистрально связны посредством пути, проходящего через всех этих вассалов.

Пояснение: Таким образом, мы получили, что в минимальном графе все ребра одного веса образуют дерево на владениях соответствующего веса.

Замечание 2. Множеству жил может соответствовать один и тот же оммаж.

В соответствии с переопределением понятия жилы мы изменим понятие пучка по сравнению с тем, как оно было дано в [17].

Определение 8. *Пучок* — граф, построенный на исходном наборе вершин, множество ребер которого равно объединению жил, выбранных по одной для каждого значимого слова (рис. 12).

Каждый пучок однозначно задается набором жил по одной для каждой клики. Для данного набора вершин V все пучки имеют одинаковое число ребер. Два этих утверждения были сформулированы и доказаны в работе [17] (теорема 1 и лемма 3.1 соответственно).

Бездетным называется значимое слово, такое, что у порожденной этим словом клики нет сыновей в графе клик.

Благодаря всем введениям доказательство теоремы о множестве минимальных графов смежности, сформулированной и обоснованной в работе [17], может быть сокращено.

Теорема (о множестве минимальных графов смежности). Множество минимальных графов смежности совпадает с множеством пучков.

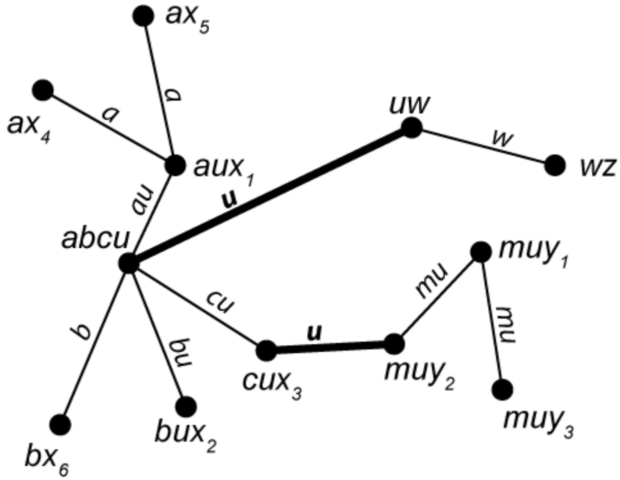

Рис. 12. Пучок.
Пучок, в котором выделена жила из рис. 11.

Доказательство. Любой граф смежности M является пучком, так как множество его ребер совпадает с объединением соответствующих жил всех значимых весов, т.к. по утверждению 2 следует, что для любого сужения минимального графа смежности на значимое слово U множество ребер этого сужения является жилой.

Докажем, что любой пучок является графом смежности. Будем доказывать, что любое значимое сужение пучка магистрально связно. Проведем индукцию на частично упорядоченном множестве, которым является граф клик. В данном случае порядком будет выступать отношение отцовства.

База: Сужение пучка на бездетные слова магистрально связно, потому как оно представляет собой жилу только на доменных вершинах — то есть дерево.

Переход: Рассмотрим некое значимое сужение $P \downarrow U$, про которое мы знаем, что все сужения $P \downarrow S_i$ на сыновей U магистрально связны. Значит, магистрально связными являются и все владения сужения $P \downarrow U$. Рассмотрим жилу пучка, выбранную для U. Она соединяет все владения $P \downarrow U$ и все ее ребра имеют вес U, поэтому $P \downarrow U$ магистрально связно.

Таким образом, любой пучок оказывается магистрально связным для любого значимого сужения, а значит, любой пучок магистрально связен — то есть он являет графом смежности.

Поскольку любой минимальный граф смежности является пучком, а любой пучок является графом смежности и число ребер у всех пучков равно, значит, множества графов смежности и пучков совпадают.

Из теоремы в работе [17] получено 4 следствия, повторим их:

Следствие 1. Число ребер в минимальных графах смежности одинаково.

Следствие 2. Множество минимальных графов смежности совпадает с множеством пучков, которое равно декартовому произведению множеств жил каждой клики.

Следствие 3. Мощность множества графов смежности равна произведению мощностей множеств жил каждой клики.

Следствие 4. Согласно следствию 2, для того, чтобы построить множество графов смежности, достаточно для каждой клики построить множество соответствующей ей жил.

Теперь мы можем добавить еще одно:

Следствие 5. Если любая курия графа смежности является оммажем, то граф является минимальным графом смежности.

Заключение.В статье мы развили и использовали систему терминов и утверждений работы [17], основанные на минимальных графах смежности, в применении к обычным (не ограниченным требованием минимальности) графам смежности. Все рассмотренные понятия и утверждения удалось обобщить на более широкую область и сделать их доступными для исследования произвольной вторичной структуры алгебраической байесовской сети.

Так, понятие *сжатие* было переработано по сравнению с [17] и рассматривается в применении к трем объектам: владениям, множеству ребер и графам (в последнее основывается на первых двух). Были изменены или дополнены в целях улучшения понимания и удобства примененияпонятия *жилы, пучка, феода* и *оммажа*. Последнее теперь понимается как частный случай курии — результата применения сжатия к графу.

Так, была доказана эквивалентность магистральной связности каждого сужения графа тому, что он является графом смежности, сформулированы и доказаны две леммы, на основе которых было обосновано утверждение о том произвольное сжатие любого минимального графа смежности является оммажем — деревом с ребрами кратности один.

Полученные результаты позволили усовершенствовать доказательства ряда утверждений из статьи [17], в частности, теоремы о множестве минимальных графов смежности.

На основе сформированного теоретического аппарата и доказанных утверждений открывается путь для дальнейших исследований графов смежности — представлений вторичной структуры алгебраической байесовской сети. В частности, это подводит нас к сравнению множеств минимальных графов смежности и нередуцируемых графов смежности, эквивалентность которых, доказанная в [3], может быть обоснована и без привлечения теории матроидов.

Литература

1. *Городецкий В.И.* Алгебраическиебайесовскиесети — новаяпарадигма экспертно-вычислительных систем // Юбилейный сборник трудов институтов Отделения информатики, вычислительной техники и автоматизации РАН. Т. 2. М.: РАН, 1993, С. 120–141.
2. *Городецкий В.И., Тулупьев А.Л.* Формирование непротиворечивых баз знаний с неопределенностью // Известия РАН. Сер. Теория и системы управления. 1997. №5. С. 33–42.
3. *Опарин В.В., Фильченков А.А., Тулупьев А.Л., Сироткин А.В.* Матроидное представление семейства графов смежности над набором фрагментов знаний // Научно-технический вестник Санкт-Петербургского государственного университета информационных технологий, механики и оптики. 2010. Вып. 4. С. 73–76.
4. *Павельчук А.В., Тулупьев А.Л., Тотмянина С.А.* Подход к объектно-ориентированному представлению данных алгебраических байесовских сетей в java-коде и реляционных СУБД // Региональная информатика-2008 (РИ-2008). XI Санкт-Петербургская международная конференция. Санкт-Петербург, 22–24 октября, 2008 г.: Материалы конференции / СПОИСУ. СПб.: 2009. С. 68–76.
5. *Тотмянина С.А., Павельчук А.В., Тулупьев А.Л.* Алгебраические байесовские сети: структуры данных в СУБД и Java- коде // Интегрированные модели, мягкие вычисления, вероятностные системы и комплексы программ в искусственном интеллекте. Научно-практическая конференция студентов, аспирантов, молодых ученых и специалистов (Коломна, 26–27 мая 2009 г.). Научные доклады. В 2-х т. Т. 2. М.: Физматлит, 2009, С. 123–131.
6. *Тулупьев А.Л.* Алгебраические байесовские сети: глобальный логико-вероятностный вывод в деревьях смежности: Учеб.пособие. СПб.: СПбГУ; ООО Издательство «Анатолия», 2007. 40 с. (Сер.Элементы мягких вычислений).
7. *Тулупьев А.Л.* Алгебраические байесовские сети: реализация логико-вероятностного вывода в комплексе java-программ // Труды СПИИРАН. СПб.: Наука, 2009. Вып. 8. С. 191–232.

8. *Тулупьев А.Л.* Алгебраические байесовские сети: система операций глобального логико-вероятностного вывода // Информационно-измерительные и управляющие системы. 2010. № 11. С 65–72.
9. *Тулупьев А.Л.* Ациклические алгебраические байесовские сети: логико-вероятностный вывод // Нечеткие системы и мягкие вычисления: Научный журнал Российской ассоциации нечетких систем и мягких вычислений. 2006. Том 1, № 1. С. 57–93.
10. *Тулупьев А.Л.* Байесовские сети: логико-вероятностный вывод в циклах. СПб.: Изд-во С.-Петербургского ун-та, 2008. 140 с. (Элементы мягких вычислений.)
11. *Тулупьев А.Л.* Непротиворечивость оценок вероятностей в алгебраических байесовских сетях. Вестник СПбГУ. Сер. 10. 2009. Вып. 3. С. 144–151.
12. *Тулупьев А.Л.* Преобразование ациклических байесовских сетей доверия в алгебраические байесовские сети // Известия высших учебных заведений: Приборостроение. 2009. № 3. С. 21–23.
13. *Тулупьев А.Л.* Согласованность данных и оценка вероятности альтернатив в цикле стохастических предпочтений // Известия высших учебных заведений: Приборостроение. 2009. № 7. С. 3–8.
14. *Тулупьев А.Л., Николенко С.И., Сироткин А.В.* Байесовские сети: логико-вероятностный подход. СПб.: Наука, 2006. 607 с.
15. *Тулупьев А.Л., Сироткин А.В.* Алгебраические байесовские сети: принцип декомпозиции и логико-вероятностный вывод в условиях неопределенности // Информационно-измерительные и управляющие системы. 2008. № 10. т. 6. С. 85–87.
16. *Тулупьев А.Л., Сироткин А.В., Николенко С.И.* Байесовские сети доверия: логико-вероятностный вывод в ациклических направленных графах. СПб.:Изд-во С.-Петерб. ун-та, 2009, 400 с.
17. *Фильченков А.А., Тулупьев А.Л.* Структурный анализ систем минимальных графов смежности Труды СПИИРАН. 2009. Вып. 11. СПб. Наука, 2009. С. 104–127.
18. *Фильченков А.А., Тулупьев А.Л., Сироткин А.В.*Структурный анализ клик минимальных графов смежности // ВестникТверского государственного университета. Сер. Прикладнаяматематика. 2010. [впечати].
19. *Gorodetsky V.I.,Drozdgin V.V., Jusupov R.M.*ApplicationofAttributedGrammarandAlgorithmicSensitivityModelforKnowledgeRepresentation and Estimation // Artificial Intelligence and Information, Control Systems of ROBOTSA. Amsterdam: ElsevierSciencePublishersB. V., 1984, P. 232–237.

Фильченков Андрей Александрович — аспирант кафедры информатики математико-механического факультета С.-Петербургского государственного университета (СПбГУ), младший научный сотрудник лаборатории теоретических и междисциплинарных проблем информатики СПИИРАН. Область научных интересов: автоматическое обучение вероятностных графических моделей. Число научных публикаций — 6. aaafil@mail.ru, СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-3337, факс +7(812)328-4450. Научный руководитель — А.Л. Тулупьев.

Filchenkov AndreyAlexandrovich — PhD student of Computer Science Department, SPbGU, junior researcher, Theoretical and Interdisciplinary Computer Science Laboratory, SPIIRAS. Research area: machine learning of probabilistic graphical models. The number of publications — 6. aaafil@mail.ru, SPIIRAS, 14-th line V.O., 39, St. Petersburg, 199178, Russia; office phone +7(812)328-3337, fax +7(812)328-4450. Scientific advisor — A.L. Tulupyev.

Тулупьев Александр Львович — д.ф.-м.н., доцент; заведующий лабораторией теоретических и междисциплинарных проблем информатики СПИИРАН, доцент кафедры информатики математико-механического факультета С.-Петербургского государственного университета (СПбГУ). Область научных интересов: представление и обработка данных и знаний с неопределенностью, применение методов математики и информатики в социокультурных исследованиях, применение методов биостатистики и математического моделирования в эпидемиологии, технология разработки программных комплексов с СУБД. Число научных публикаций — 210. ALT@iias.spb.su, www.tulupyev.spb.ru; СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-3337, факс +7(812)328-4450.

Tulupyev Alexander Lvovich — PhD in Computer Science, Dr. of Sc.. Associate Professor; Head of Theoretical and Interdisciplinary Computer Science Laboratory, SPIIRAS, Associate Professor of Computer Science Department, SPbSU. Research area: uncertain data and knowledge representation and processing, mathematics and computer science applications in socio-cultural studies, biostatistics, simulation, and mathematical modeling applications in epidemiology, data intensive software systems development technology. Number of publications — 210. ALT@iias.spb.su, www.tulupyev.spb.ru; SPIIRAS, 14-th line V.O., 39, St. Petersburg, 199178, Russia; office phone +7(812)328-3337, fax +7(812)328-4450.

Сироткин Александр Владимирович — младший научный сотрудник лаборатории теоретических и междисциплинарных проблем информатики СПИИРАН. Область научных интересов: алгебраические байесовские сети: вычислительные аспекты логико-вероятностного вывода в условиях неопределенности. Число научных публикаций — 40. avs@iias,spb.su, www.tulupyev.spb.ru; СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-3337, факс +7(812)328-4450.

Sirotkin Alexander Vladimirovich — junior researcher, Theoretical and Interdisciplinary Computer Science Laboratory, SPIIRAS. Research interests: algebraic Bayesian networks, algorithms of probabilistic-logic inference under uncertainty. The number of publications — 40.avs@iias,spb.su, www.tulupyev.spb.ru; SPIIRAS, 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)328-3337, fax +7(812)328-4450.

Поддержка исследования. Работа выполнена при финансовой поддержке РФФИ, проект № **09-01-00861**-а «Методология построения интеллектуальных систем поддержки принятия решений на основе баз фрагментов знаний с вероятностной неопределенностью».

РЕФЕРАТ

Фильченков А.А., Тулупьев А.Л., Сироткин А.В. **Особенности анализа вторичной структуры алгебраической байесовской сети.**

Цель работы — обобщить терминологию, применяемую для структурного анализа минимальных графов смежности, представляющих вторичную структуру алгебраической байесовской сети, а также некоторые результаты, полученные для этих графов, на произвольные графы смежности, представляющие эту структуру.

Рассмотрена система терминов, употребляемая при исследовании и описании свойств минимальных графов смежности, и на ее основе построена расширенная система терминов, позволяющая работать с графами смежности, на которые не наложено ограничение минимальности по числу ребер. Так, была расширена операция сжатия графа смежности, которая преобразует граф смежности в кратный граф, превращая компоненты связности особых подграфов максимального графа смежности — владения — в вершины, а множества ребер, соединяющие эти владения — в кратные ребра.

На основе введенного понятия сжатия были переопределены понятия феода и оммажа. Последний определялся через введенное понятия курии — упомянутого выше кратного графа, как дерево с кратностью ребер, равной единице.

На основе введенных понятий были обобщены и доказаны утверждения относительно графов смежности без ограничения на их минимальность. Кроме того, было доказано, что оммаж является в определенном смысле минимальной курией, на основе чего было установлено, что минимальный граф смежности сжимается до оммажа.

Помимо этого были упрощено доказательство теоремы о множестве минимальных графов смежности.

Полученные результаты развиваютпонятия, лежащие на стыке теории графов смежности и теории алгебраических байесовских сетей, а также развивают последнюю, формируя теоретическую базу для дальнейшего изучения вторичной структуры, которую представляют в виде графа смежности, и свойства которой обуславливают эффективность и даже саму возможность логико-вероятностных априорного и апостериорного выводов и поддержания непротиворечивости всех видов — основных достоинств байесовской сети.

Полученные результаты раскрывают определенную взаимосвязь минимальности и нередуцируемости графов смежности.

SUMMARY

Filchenkov A.A., Tulupyev A.L., Sirotkin A.V. **Algebraic Bayesian networks secondary structure analysis features.**

The goal of this work is to generalize terminology applied for minimal join graph representing algebraic Bayesian networks secondary structure analysis and also some results obtained for this graphs to arbitrary join graphs representing secondary structure.

Term system, applied in minimal join graphs research and description is considered and on its base an extended term system is build that allow to deal with join graphs not limited with the property of its edges number minimality. So that, a join graph compression operator that transforms a join graph into multiply graph by turning maximal join graph special subgraph connection components (possessions) into vertexes, and edge set connecting this vertexes into multiply edges.

Terms of feud and homage are redefined on the base of new term of compression. The homage is defined by a term of curia — the multiplied graph mentioned above, as a tree with edge multiplicity equals to 1.

Assertions about join graphs not limited by its minimality were generalized and proved. Homage is proven to be a minimal curia in some meanings and by that is proven that compressed minimal join graph is an homage.

Also minimal join graph set theorem proof was simplified.

Obtained results develops the terms belonging to the graph theory and the algebraic Bayesian network theory, and develops the last one by designing a theoretical basement for further researches of secondary structure that is usually presented as a join graph and that its properties condition efficiency and even the ability to maintain a priori and a posteriori probabilistic-logic inference and reconciliation that are referred to be algebraic Bayesian network main virtues.

Obtained results explore minimal (irreducible) and minimum join graphs interconnection.