УДК 004.8

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

РЕБРА ГРАФОВ СМЕЖНОСТИ В КОНТЕКСТЕ КОМПАРАТИВНОГО АНАЛИЗА КЛИК МИНИМАЛЬНЫХ ГРАФОВ СМЕЖНОСТИ АЛГЕБРАИЧЕСКИХ БАЙЕСОВСКИХ СЕТЕЙ

Фильченков А.А., Тулупьев А.Л., Сироткин А.В. **Ребра графов смежности в контексте компаративного анализа клик минимальных графов смежности алгебраических байесовских сетей.**

Аннотация. Алгебраические байесовские сети представляют собой логико-вероятностную графическую модель систем знаний с неопределенностью и позволяют работать в том числе с интервальными оценками вероятности. Существенной для их работы является вторичная структура, представляемая в виде графа смежности. Данная статья исследует ребра клик минимальных графов смежности для спецификации различных типов клик. В частности, было доказано, что у определенного класса клик, которые являются основными с точки зрения построения множества минимальных графов смежности, множество вершин совпадает с множеством концов особых ребер, вес которых совпадает с весом клики.

Ключевые слова: алгебраические байесовские сети, вторичная структура, машинное обучение, вероятностно-графические модели систем знаний.

Filchenkov A.A., Tulupyev A.L., Sirotkin A.V. **Join graph edges in context of algebraic Bayesian network minimal join graph cliques comparative analysis.**

Annotation. Algebraic Bayesian networks (ABN) are probabilistic-logic graphic models of knowledge systems with uncertainty and gives an advantage to deal with interval probability estimates. Secondary structure usually represented as an join graph is essential for ABN work. The article analyses edges of various minimal join graph cliques to specify different clique types. In particular, it is proven that vertex set of the class of cliques that are basic for minimal join graph set synthesis equals to set of end of specified edges, weight of those equals to the clique weight.

Keywords: algebraic Bayesian networks, secondary structure, machine learning, probabilistic graphical knowledge models.

1. Введение. Среди класса вероятностных графических моделей (ВГМ) алгебраические байесовские сети (АБС), представляющие собой логико-вероятностную графическую модель систем знаний с неопределенностью [1–2, 14, 24], выделяются тем, что позволяют работать с точечными и с интервальными оценками вероятностей. Как и родственные им байесовские сети доверия (БСД), АБС имеют возможность осуществлять логико-вероятностные априорный и апостериорный выводы (синтез согласованных оценок и пропагацию свидетельств соответственно).

Работа АБС в значительной степени зависит от выбора вторичной структуры [7, 14, 19], которая обычно представляется графом смежности над максимальными фрагментами знаний (ФЗ) — первичной структурой АБС. Так, скорость работы алгоритмов логико-вероятностных априорного и апостериорного выводов и поддержания непротиворечивости сети, а также сама возможность их применения определяются свойствами этого графа [10–13, 15, 17]. Лучшими в этом отношении являются деревья смежности [9], однако АБС, для которых можно построить ациклическую вторичную структуру, — это лишь частный случай АБС [20]. В общем же случае наиболее эффективные вторичные структуры содержатся в множестве минимальных по числу ребер графов смежности [4–8], которые одновременно являются нередуцируемыми [3].

К сожалению, несмотря на свою важность, минимальные графы смежности изучены недостаточно. Основополагающей работой, в которой формируется система терминов, характеризующих минимальные графы смежности, и содержащей результаты исследования этой глобальной структуры, является статья [20]. В ней же сформулирована базовая схема алгоритма синтеза семейства минимальных графов смежности. Ряд работ [21–24] развивали полученные результаты, в частности, статья [21] предлагает классификацию клик, которая позволяет учитывать особенности некоторых типов клик и обрабатывать их более эффективно, что положительно сказывается на времени работы алгоритмов, которые были сформулированы в статьях [18–19]. Это показывает, что дальнейший анализ клик позволит сформулировать дальнейшие улучшения для оптимизации работы алгоритмов синтеза множества минимальных графов смежности. В то же время алгоритмы, сформулированные в работах [18–19], опираются на утверждения, которые следуют из статьи [20], однако эти утверждения требуют формального доказательства.

Цель данной работы — исследовать ребра графов смежности и получить их свойства, которые бы специфицировали клики и позволяли улучшить алгоритмы построения множества минимальных графов смежности.

2. Основные определения. *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:

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

Для внесения ясности определим: $\subseteq$ — (нестрогое) включение:

$$(A \subseteq B) \Leftrightarrow \big((\forall a \in A) a \in B\big).$$

В противоположность этому $\subset$ — строгое включение:

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

Алфавитом будем называть множество атомарных пропозициональных формул

$$A = \{x_1, \dots, x_n\}.$$

Словом тогда будем называть некое подмножество алфавита.

Первичная структура или множество главных конъюнктов максимальных ФЗ, вошедших в байесовскую сеть, — такое множество слов $V^* = \{V_i \subseteq A\}_{i=1}^{i=m}$, что

- оно не содержит несобственное подмножество алфавита:

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

- никакое слово полностью не содержит никакое другое слово:

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

Вторичная структура или граф максимальных ФЗ (МФЗ) — ненаправленный граф, вершины которого соответствуют элементам множества главных конъюнктов МФЗ, вошедших в АБС, а ребра возможны только между теми вершинами, что пересечение соответствующих МФЗ непусто.

С этого момента под словом «граф» мы будем понимать именно граф МФЗ, если не оговорено обратное.

Введем понятия веса для вершины, ребра и 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 $W(\{V_i, V_j\})$ ребра $\{V_i, V_j\} \in E(G)$, графа G определяется как пересечение весов тех вершин, которые соединены этим ребром:

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

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

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

Под *магистральным путем* $B: V_b \dashrightarrow V_e$ от вершины V_b до вершины V_e, пересечение весов которых непусто, будем понимать 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 как **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).

Дерево смежности — это граф смежности, являющийся деревом.

Минимальный граф смежности — граф смежности, число ребер в котором минимально (рис. 1).

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 [20] было доказано, что для заданного множества вершин существует и при этом единственный максимальный граф смежности.

3. Сужение, граф клик и сжатие. *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:

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

Значимое слово графа G — слово, являющееся весом какого-либо ребра графа G.

Значимое сужение $G \downarrow U$ — сужение графа G на значимое для графа G слово U.

Клика веса U — значимое сужение $G_{\max} \downarrow U$ (рис. 2).

В [20] было доказано, что если сужение G на произвольное непустое слово непусто, то оно является полным подграфом. Таким образом, введенная нами клика является кликой в общепринятом смысле слова, то есть полным подграфом. Будем обозначать множество всех клик графа $G_{\max}$ через Clique**.**

Граф клик — направленный граф, вершинами которого являются клики из множества Clique (рис. 3). 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.

$$G_{\text{Clique}} = \langle \text{Clique}, E_{\text{Clique}} \rangle, \text{ где } P, Q \in \text{Clique}, (P, Q) \in E \Leftrightarrow$$
$$\Leftrightarrow Q \subset P \; и \; \nexists R \in \text{Clique}: Q \subset R \subset P.$$

В [20] было доказано, что граф клик не содержит направленных циклов.

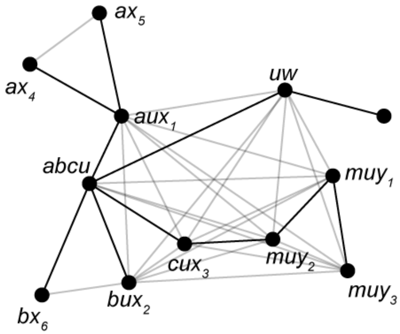

Рис. 1. Максимальный и минимальный графы смежности.
Графы смежности, построенные на вершинах $abcu$, ax_4, ax_5, aux_1, bux_2, cux_3, muy_1, muy_2, muy_3, uw, wz. Минимальный граф смежности состоит из черных ребер, максимальных граф смежности — из черных и серых.

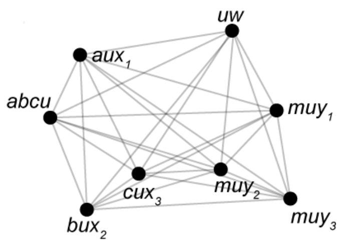

Рис. 2. Клика a.
Клика, образованная сужением максимального графа с рис. 1 на вес u.

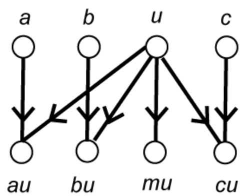

Рис. 3. Граф клик.
Граф клик для графа на рис. 1

Будем называть клики, в которые входят ребра из клики C в графе клик, *сыновьями клики* C.

Сильное сужение $G \downarrow U$ — значимое сужение $G \downarrow U$, из которого удалили все ребра веса U (рис. 4):

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

.

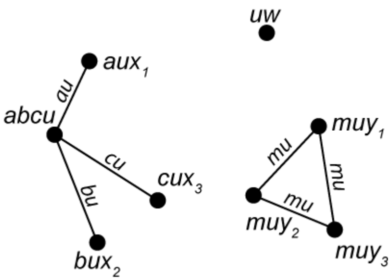

Рис. 4. Сильное сужение.
Сильное сужение графа с рис. 1 на вес u. Оно состоит из трех владений.

Сильное сужение графа $G_{\max} \downarrow U$ представляет собой компоненты связности, на которые разбивается сужение $G_{\max} \downarrow U$ удалением ребер веса U. Такие компоненты будем называть *владениями* (рис. 5).

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

Вассал V_U клики U — множество вершин, принадлежащих какому-либо одному сыну U по графу клик.

Два вассала V_U^i и V_U^j называются *братьями* (будем обозначать это через $\leftrightarrow$), если их пересечение непусто:

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

Два вассала V_U^i и V_U^j называются *полусиблингами* (будем обозначать это через $\leftrightarrow$), 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 $V_U^{w_2}, \dots, V_U^{w_{n-1}}$ — брат $V_U^{w_n}$, а $V_U^{w_n}$ — брат V_U^j:

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

Полусиблинговый путь между двумя полусиблинговыми вассалами V_U^i и V_U^j — такой набор вассалов $\{V_U^{w_1}, \dots, V_U^{w_n}\}$, 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):

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

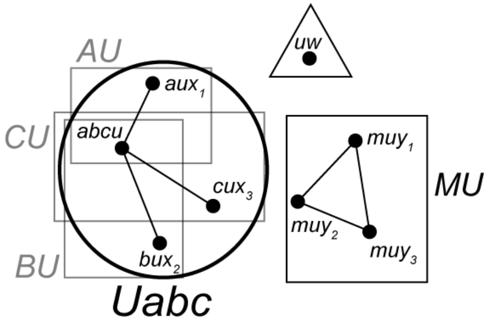

Рис. 5. Владения.

Треугольник — доменная вершина uw; прямоугольник — вассал MU, состоящий из вершин muy_1, muy_2 и muy_3; овал — братство $Uabc$, состоящее из выделенных полупрозрачными прямоугольниками вассалов AU, BU и CU; вассалы AU, BU и CU приходятся друг другу братьями.

Теорема (классификация владений) [23]. Любое владение любого сильного сужения $G \downarrow U$ является либо доменной вершиной, либо вассалом, либо братством U.

В [22] приводятся три определения *сжатия,* задающие эту операцию на различных множествах:

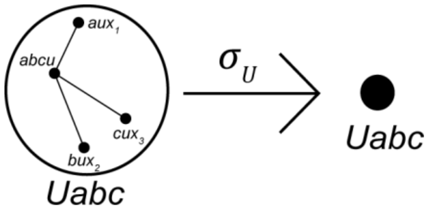

Рис. 6. Сжатие владения (сжатие братства $Uabc$ в вершину).

Сжатие σ_U компоненты связности $P_U^i \subseteq G \downarrow U$ в вершину f_i — отображение на множестве подмножеств вершин, сопоставляющее множеству вершин P_U^i вершину f_i (рис. 6);

Сжатие σ_U множества ребер

$$E_{i,j} \subset (G \downarrow U) \backslash (G \downarrow U),$$

соединяющих владения P_U^i и 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. 7);

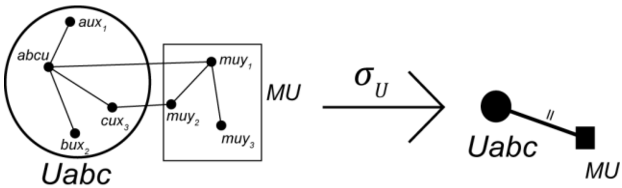

Рис. 7. Сжатие множества ребер.
Сжатие множества ребер, соединяющих владения $Uabc$ и MUв ребро кратности 2, соединяющее соответствующие вершины.

Сжатие σ_U графа смежности G в граф K_U — отображение на множестве графов, сопоставляющий графу G, являющемуся графом смежности, граф K_U, вершинами которого являются владения сильного сужения $G \downarrow U$, 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 из P_U^1 и P_U^2 (рис. 8).*Феод* — f_i (из определения выше) — вершина, получившаяся путем сжатия какого-то владения (рис. 6).

Курия веса U — граф K_U из определения сжатия — ненаправленный граф с кратными ребрами, полученный путем сжатия значимого сужения $G \downarrow U$ (рис. 8).

Оммаж H_U — курия K_U, являющаяся деревом, все ребра которой имеют кратность, равную единице. В [22] было доказано, что любой минимальный граф смежности сжимается до оммажа (рис. 9).

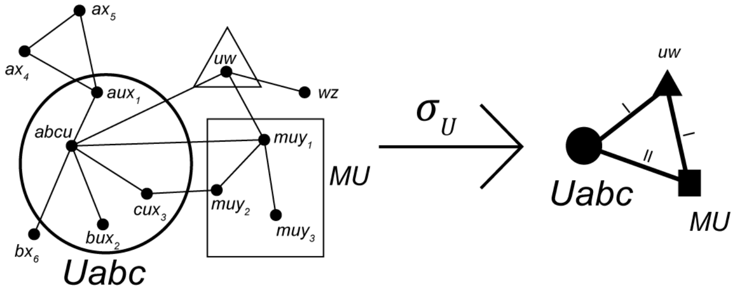

Рис. 8. Сжатие графа.

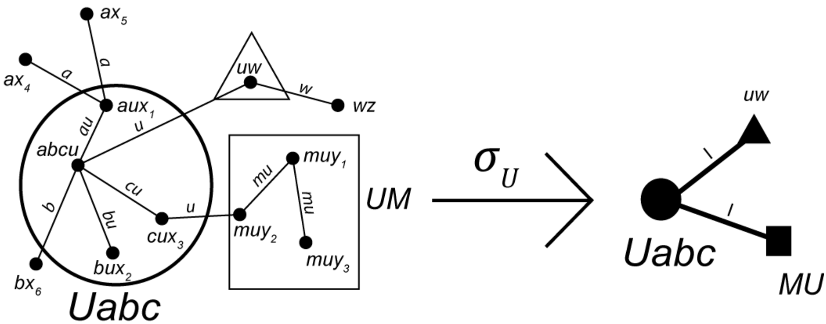

Рис. 9. Оммаж.
Сжатие минимального графа, построенного над теми же вершинами, что и на рис. 8, до оммажа.

4. Классификация клик. *Собственное ребро клики* — ребро, принадлежащее клике, вес которого совпадает с весом клики.

По числу собственных ребер множество клик можно поделить на (табл. 1):

Таблица 1. **Классификация клик**

Наличие ребер	Бездетные	Родительские
Безреберные	C_0^-	C_0^+
Однореберные	C_1^-	C_1^+
Многореберные	C_n^-	C_n^+

- *безреберные клики* — C_0 — клики (сужения), у которых нет собственных ребер;

- *однореберные клики* — C_1 — клики, у которых ровно одно собственное ребро;
- *многореберные клики* — C_n — клики, у которых более одного собственного ребра.

По наличию детей множество клик можно поделить на следующие:

- *бездетные* — C^- — клики, у которых нет детей;
- *родительские* — C^+ —клики, у которых есть дети.

По числу феодов, в которые сжимаются клики, они делятся на:

- *моноклики* — клики, сжимающиеся до одного феода;
- *стереоклики* — клики, сжимающиеся до более чем одного феода.

Псевдоклика — C_0^- — сужение, не имеющее собственных ребер и не имеющее детей.

Моноклика-0 — C_0^+ — сужение, не имеющее собственных ребер, но имеющее детей.

Биклика — C_1^- — клика, имеющее ровно одно собственное ребро и не имеющее детей.

Моноклика-1 — C_1^+ — клика, имеющая ровно одно собственное ребро и имеющая детей.

Бездетная поликлика (возможно также название *бездетная стереоклика*) — C_n^- —клика, имеющая более одного собственного ребра, но не имеющая детей.

Родительская поликлика — C_n^+ —клика, имеющая более одного собственного ребра и имеющая детей.

Моноклика-n — mC_n^+ — клика, имеющая более одного собственного ребра и имеющая детей, но состоящая ровно из одного феода.

Родительская стереоклика — pC_n^+ —клика, имеющая более одного собственного ребра, имеющая детей, и состоящая более чем из одного феода.

Полученные в [21] результаты удобно расположить в таблице (табл. 2).

Таблица 2. **Характеристики различных клик**

Сужение	Обозначение	Является ли кликой	Есть ли дети	Число собственных ребер	Число вершин	Число феодов	Число жил
Псевдоклика	C_0^-	Нет	Нет	0	1	0	0
Моноклика-0	C_0^+	Нет	Да	0	> 2	1	1
Биклика	C_1^-	Да	Нет	1	2	2	1
Моноклика-1	C_1^+	Да	Да	1	> 2	1	1
Бездетная стереоклика	C_n^-	Да	Нет	> 1	> 2	> 2	> 1
Родительская поликлика	C_n^+	Да	Да	> 1	> 2	$\forall$	$\forall$
Моноклика-n	mC_n^+	Да	Да	> 1	> 2	1	1
Родительская стереоклика	pC_n^+	Да	Да	> 1	> 2	> 1	> 1

5. Анализ ребер.

Определение 1. *Обязательное ребро* — ребро, которое входит в любой минимальный граф смежности.

Работа алгоритмов построения множества минимальных графов смежности при помощи самостоятельных клик [18] и при помощи клик владений [19] опирается на тот факт, что обязательными ребрами являются ребра биклики, указания на что содержатся в [20], однако, поскольку в работе [21] мы создали новую классификацию клик, это утверждение требует должно быть сформулировано и доказано в новых условиях.

Утверждение 1. Обязательными ребрами являются те и только те ребра, которые являются собственными ребрами биклики.

Доказательство.

1) $\Rightarrow$. Ребра биклик являются обязательными, так как они являются единственными магистральными путями, связывающими пары вершины биклик, поэтому ребра биклики должны присутствовать в любом графе смежности, в том числе магистральном.

2) $\Leftarrow$. Теперь рассмотрим произвольное собственное ребро клики, состоящей из более чем двух вершин. Если хотя бы один из концов выбранного ребра не является доменным, то, заменив выбранное ребро на ребро, соединяющее один конец и вершину из владения, к которому

принадлежит второй конец, мы сохраним магистральную связность. Если же обе вершины являются доменными, и в клике кроме них есть еще какие-то вершины, выходящие в хотя бы еще один феод, то будет существовать множество оммажей, не содержащих выбранного ребра, так как для произвольного набора из более чем двух ребер существует дерево, в котором заранее выбранные вершины не являются смежными.

Таким образом, обязательными ребрами являются те и только те ребра, которые являются собственными ребрами биклик.

Определение 2. *Избыточное ребро* — ребро, которое не входит ни в какой минимальный граф смежности.

Определение 3. *Сводное ребро* — ребро, соединяющее две вершины из одного братства, не принадлежащие одному и тому же вассалу.

***Утверждение* 2.** В минимальном графе смежности любой магистральный путь между двумя вершинам одного братства не содержит сводных ребер.

Доказательство. Рассмотрим две вершины (a и b) одного братства, принадлежащие, соответственно, вассалам V_U^a и V_U^b. Существует полусиблинговый путь от V_U^a до V_U^b: $V_U^{w_1}, \dots, V_U^{w_n}$. Все вассалы являются магистрально связными, поэтому магистральный путь от a до b является объединением магистральных путей между точками пересечения братьев, и ребра, которые он содержит, целиком лежат внутри вассалов, поэтому он не содержит сводных ребер.

Следствие 1. Сводные ребра являются избыточными.

Следствие 2. Единственное собственное ребро моноклики-1 является избыточным.

Определение 4. *Основное множество вершин клики* — множество вершин, являющихся концами ее собственных ребер.

Утверждение 3. Основное множество вершин стереоклики совпадает с множеством ее вершин.

Доказательство. С одной стороны, концы собственного ребра стереоклики принадлежат стереоклике. С другой стороны, так как стереоклика состоит хотя бы из двух владений, то для любой вершины одного владения можно выбрать произвольную вершину из любого другого владения: такая пара вершин соединена собственным ребром. Отсюда следует, что два упомянутых множества совпадают.

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

Следствие 4. Мощность множества минимальных графов смежности равна произведению мощностей множеств жил каждой стереоклики.

5. Заключение. Множество минимальных графов смежности играет важную роль в исследовании АБС, и данная статья продолжает изучение этого множества. В работе были продолжены исследования клик минимальных графов смежности, начатые в [21]. Анализу были подвергнуты ребра в кликах.

Так, было доказано, что *обязательными ребрами* — ребрами, содержащимися во всех графах смежности, построенных над данным набором ФЗ, являются ребра биклик и только они. Это формально подтверждает корректность алгоритмов, сформулированных в [18–19].

Было показано, что *избыточными ребрами* (т.е. ребрами, не входящими ни в один из минимальных графов смежности, построенных над данным набором вершин) в минимальных графах смежности являются ребра, которые соединяют вершины в братствах, но не лежат на магистральном пути между ними.

Было доказано, что множество вершин стереоклики совпадает с *множеством ее основных вершин* — т. е. с вершинами, являющимися концами ее собственных ребер. Последние два утверждения позволят улучшить построение множества вершин для каждой клики в алгоритме построения множества минимальных графов смежности, потому что именно стереоклики рассматриваются как основное множество клик, для которых нужно строить множество вершин, им принадлежащих. Остальные клики либо игнорируются совсем, либо обрабатываются особым образом [18–19].

Полученные результаты развивают теорию АБС и, в частности, теорию минимальных графов смежности АБС, являясь основой для разработки улучшенных алгоритмов синтеза множества минимальных графов смежности, которые будут учитывать специфику различных типов клик, что позволит ускорить их работу и вывод результатов.

Дальнейший анализ различных элементов клик может быть продуктивным с точки зрения спецификации различных типов клик, что в свою очередь будет являться теоретической основой для улучшений алгоритмов синтеза множества минимальных графов смежности, а также с точки зрения построения наиболее «эффективных» минимальных графов смежности с точки зрения основных алгоритмов АБС.

Литература

1. *Городецкий В.И.* Алгебраические байесовские сети — новая парадигма экспертно-вычислительных систем // Юбилейный сб. тр. институтов Отделения информатики, вычислительной техники и автоматизации РАН. Т. 2. М., 1993, С. 120–141.
2. *Городецкий В.И., Тулупьев А.Л.* Формирование непротиворечивых баз знаний с неопределенностью // Изв. 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 // Материалы XI Санкт-Петербургской междунар. конф. «Региональная информатика-2008 (РИ-2008). Санкт-Петербург, 22–24 октября, 2008 г. СПб., 2009. С. 68–76.
5. *Сироткин А.В.* Модели, алгоритмы и вычислительная сложность синтеза согласованных оценок истинности в алгебраических байесовских сетях // Информационно-измерительные и управляющие системы. 2009. № 11. С. 32–37.
6. *Тотмянина С.А., Павельчук А.В., Тулупьев А.Л.* Алгебраические байесовские сети: структуры данных в СУБД и Java- коде // Докл. науч.-практич. конф. студентов, аспирантов, молодых ученых и специалистов «Интегрированные модели, мягкие вычисления, вероятностные системы и комплексы программ в искусственном интеллекте. Коломна, 26–27 мая 2009 г. Т. 2. М., 2009. С. 123–131.
7. *Тулупьев А.Л.* Алгебраические байесовские сети: глобальный логико-вероятностный вывод в деревьях смежности: Учеб. пособие. СПб.: СПбГУ; ООО «Издательство «Анатолия», 2007. 40 с.
8. *Тулупьев А.Л.* Алгебраические байесовские сети: реализация логико-вероятностного вывода в комплексе java-программ // Тр. СПИИРАН. 2009. Вып. 8. С. 191–232.
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. Т. 1, № 1. С. 57–93.
10. *Тулупьев А.Л.* Байесовские сети: логико-вероятностный вывод в циклах. СПб.: Изд-во СПбГУ, 2008. 140 с.
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. *Тулупьев А.Л., Столяров Д.М., Ментюков М.В.* Представление локальной и глобальной структуры алгебраической байесовской сети в Java-приложениях. // Тр. СПИИРАН. 2007. Вып. 5. СПб.: Наука, 2007. С. 71–99.
18. *Фильченков А.А,* Алгоритм построения множества минимальных графов смежности при помощи клик владений// Тр. СПИИРАН. 2010. Вып. 2 (13). [в печати]
19. *Фильченков А.А.* Алгоритм построения множества минимальных графов смежности при помощи самоуправляемых клик // Тр. СПИИРАН. 2010. Вып. 1 (12). С. 119–133.
20. *Фильченков А.А., Тулупьев А.Л.* Структурный анализ систем минимальных графов смежности Тр. СПИИРАН. 2009. Вып. 11. С. 104–127.
21. *Фильченков А.А., Тулупьев А.Л., Сироткин А.В.* Компаративный анализа клик минимальных графов смежности // Тр. СПИИРАН. 2010. Вып. 1 (13). [в печати]
22. *Фильченков А.А., Тулупьев А.Л., Сироткин А.В.* Особенности анализа вторичной структуры алгебраической байесовской сети // Тр. СПИИРАН. 2010. Вып. 1 (12). С. 97–118.
23. *Фильченков А.А., Тулупьев А.Л., Сироткин А.В.* Структурный анализ клик минимальных графов смежности // Вестник Тверского государственного ун-та. Сер. Прикладная математика. 2010.
24. *Gorodetsky V.I., Drozdgin V.V., Jusupov R.M.* Application of Attributed Grammar and Algorithmic Sensitivity Model for Knowledge Representation and Estimation // Artificial Intelligence and Information, Control Systems of ROBOTSA. Amsterdam: Elsevier Science Publishers B. V., 1984. P. 232–237.

Фильченков Андрей Александрович — аспирант кафедры информатики математико-механического факультета Санкт-Петербургского государственного университета (СПбГУ), м. н. с. лаборатории теоретических и междисциплинарных проблем информатики СПИИРАН. Область научных интересов: автоматическое обучение вероятностных графических моделей. Число научных публикаций — 9. aaafil@mail.ru, СПИИРАН, 14-я линия В.О., д. 39, Санкт-Петербург, 199178, РФ; р.т. +7(812)328-3337, факс +7(812)328-4450. Научный руководитель — д-р физ.-мат. наук, доцент А.Л. Тулупьев.

Filchenkov Andrey Alexandrovich — 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 — 9. aaafil@mail.ru, SPIIRAS, 14th line V.O., 39, St. Petersburg, 199178, Russia; office phone +7(812)328-3337, fax +7(812)328-4450. Scientific advisor — PhD in Computer Science, Dr. of Sc. Associate Professor A.L. Tulupyev.

Тулупьев Александр Львович — д-р физ.-мат. наук, доцент; заведующий лабораторией теоретических и междисциплинарных проблем информатики СПИИРАН, доцент кафедры информатики математико-механического факультета Санкт-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 — 220. ALT@iias.spb.su, www.tulupyev.spb.ru; СПИИРАН, 14-я линия В.О., д. 39, 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 — 220. ALT@iias.spb.su, www.tulupyev.spb.ru; SPIIRAS, 14th line V.O., 39, St. Petersburg, 199178, Russia; office phone +7(812)328-3337, fax +7(812)328-4450.

Сироткин Александр Владимирович — м. н. с. лаборатории теоретических и междисциплинарных проблем информатики СПИИРАН. Область научных интересов: алгебраические байесовские сети: вычислительные аспекты логико-вероятностного вывода в условиях неопределенности. Число научных публикаций — 50. avs@iias,spb.su, www.tulupyev.spb.ru; СПИИРАН, 14-я линия В.О., д. 39, 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 — 50. avs@iias,spb.su, www.tulupyev.spb.ru; SPIIRAS, 39, 14th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)328-3337, fax +7(812)328-4450.

Рекомендовано ТиМПИ СПИИРАН, зав. лаб. А.Л. Тулупьев, д.ф.-м.н., доцент.

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», а также грантом правительства Санкт-Петербурга для победителей конкурса грантов Санкт-Петербурга для студентов, аспирантов, молодых ученых, молодых кандидатов наук 2010 г., диплом ПСП№10697.

РЕФЕРАТ

Фильченков А.А, Тулупьев А.Л., Сироткин А.В. **Ребра графов смежности в контексте компаративного анализа клик минимальных графов смежности алгебраических байесовских сетей.**

В статье даются основы теории глобальной структуры алгебраической байесовской сети (АБС), в том числе определения таких основных понятий как *сужения*, *клики*, *владения*, *сжатия*. Приводится известная классификация клик (по числу собственных ребер — ребер того же веса, что и клика, и по наличию детей).

Цель работы — исследовать ребра графов смежности и получить их свойства, которые бы специфицировали клики и позволяли улучшить алгоритмы построения множества минимальных графов смежности.

Обязательные ребра определяются как ребра, которые входят в каждый минимальный граф смежности над данным набором фрагментов знаний, а *избыточные* — как ребра, не входящие ни в один из них. Доказано, что обязательными являются те и только те ребра, которые входят в *биклику* (клику, имеющую только две вершины), а избыточными являются сводные ребра — собственные ребра особого владения (*братства*).

Вводится понятие основного множества вершин клики как множества вершин, являющихся концами собственных ребер этой клики. Доказано, что для *стереоклик* — клик, у которых число владений больше одного, множество вершин совпадает с основным множеством вершин. Доказанные положения влекут ряд следствий: мощность множества графов смежности равна произведению мощностей множеств жил каждой поликлики, мощность множества графов смежности равна произведению мощностей множеств жил каждой стереоклики.

Полученные результаты развивают теорию АБС и, в частности, теорию минимальных графов смежности АБС, являясь основой для разработки улучшенных алгоритмов синтеза множества минимальных графов смежности, которые будут учитывать специфику различных типов клик, что позволит ускорить их работу и вывод результатов.

Дальнейший анализ различных элементов клик может быть продуктивным с точки зрения спецификации различных типов клик, что в свою очередь будет являться теоретической основой для улучшений алгоритмов синтеза множества минимальных графов смежности, а также с точки зрения построения наиболее «эффективных» минимальных графов смежности с точки зрения основных алгоритмов АБС.

SUMMARY

Filchenkov A.A, Tulupyev A.L., Sirotkin A.V. **Join graph edges in context of algebraic Bayesian network minimal join graph cliques comparative analysis.**

The basics of algebraic Bayesian network global structure and, in particular, definitions of such basic concepts as *narrowings*, *cliques*, *possessions*, *compression* are presented and adjusted to the paper purposes. The known clique classification (on the number of *proper edges* — edges with the same weight that the clique has, and on existence of its children).

The goal of this work is to research join graph edges and to obtain their properties that will specify cliques and allow to improve minimal join graph algorithm synthesis.

Necessary edges are defined as the edges that belong to each minimal join graph built over the knowledge patterns set, and *excess* ones as the edges that do not belong to any of these graphs. It is proven that the necessary edges are those and only those edges that belong to a *biclique* (a clique that has only two vertices), and that the excess edges are *stepedges* — proper edges of a special possession (*brotherhood*).

Main vertex set of a clique is defined as a set of vertices that belong to the proper edges of the clique. It is proven that number of vertices of a *stereoclique* ─ such a clique that has two or more possessions, equals to it main vertex set. The proven statements condition the conclusions: minimal join graphs set cardinal number equals to the product of each polyclique sinew set cardinal number and minimal join graphs set cardinal number equals to the product of each stereoclique sinew set cardinal number

The results develop algebraic Bayesian network theory and algebraic Bayesian network global structure theory in particular. New improved minimal join graph set synthesis algorithms that would pay attention to the different clique types features (and thus working faster) could be suggested on the base of the represented classification.

Further analysis of different components of cliques may be fruitful in the question of different clique types specification that in its turn may be theoretical basis for improved minimal join graph synthesis algorithm and also in the question of designing the most "efficient" minimal join graphs.