УДК 004.8

А.А. ФИЛЬЧЕНКОВ, А.Л. ТУЛУПЬЕВ

АНАЛИЗ ЦИКЛОВ В МИНИМАЛЬНЫХ ГРАФАХ СМЕЖНОСТИ АЛГЕБРАИЧЕСКИХ БАЙЕСОВСКИХ СЕТЕЙ

Фильченков А.А., Тулупьев А.Л. **Анализ циклов в минимальных графах смежности алгебраических байесовских сетей.**
Аннотация. Алгебраические байесовские сети (АБС) представляют собой логико-вероятностную графическую модель систем знаний с неопределенностью. Работа алгоритмов логико-вероятностного вывода АБС зависит от выбора вторичной структуры, обычно представляемой графом смежности. В частности, возможности применения указанных алгоритмов препятствуют циклы, содержащиеся в этих графах. Цель работы — исследовать циклы вторичной структуры и выявить необходимые и достаточные условия цикличности или ацикличности минимальных графов смежности. *Замкнутый сверху граф клик* определяется как граф клик с добавленным к нему корнем (*пракликой*), *полусиблинговые циклы* определены как циклы, состоящие из вассалов, *небратские полусиблинговые циклы* определены как полусиблинговые циклы, пересечение всех вассалов, входящих в которые, пусто. Сформулирована и доказана теорема о циклах, утверждающая, что необходимым и достаточным условием цикличности минимального графа смежности является существование небратских полусиблинговых циклов в какой-либо клике. Следствием из теоремы является то, что все минимальные графы смежности, построенные над данной первичной структурой АБС, являются либо циклическими, либо ациклическими одновременно.
Ключевые слова: алгебраические байесовские сети, четвертичная структура, машинное обучение, вероятностно-графические модели систем знаний, глобальная структура.

Filchenkov A.A., Tulupyev A.L. **The Algebraic Bayesian Network Minimal Join Graphs Cycles Analysis.**
Abstract. Algebraic Bayesian networks (ABN) are probabilistic-logical graphical models of knowledge systems with uncertainty. ABN probabilistic logical inference algorithms processing considerably depends on its secondary structure, which is usually represented as a join graph. In particular, the graphs cycles prevent the possibility of the mentioned algorithms application. The goal of the work is to analyze secondary structure cycles and to elucidate necessary and sufficient conditions of the minimal join graph cyclicity. The term of *clique graph closed from above* is defined as a clique graph with the added root (*praclique*) , *half-sibling cycles* are defined as cycles on vassals, non-fraternal half-sibling cycles are defined as such half-sibling cycles where intersection of all the vassals that belong to this cycle is empty. The first theorem on cycles that claims the necessary and sufficient condition of a minimal join graph cyclicity is existence of non-fraternal half-sibling cycles in any clique is formulated and proven. The consequence is that all minimal join graphs built under given algebraic Bayesian network primary structure are either cyclic or acyclic simultaneously.
Keywords: algebraic Bayesian networks, quaternary structure, machine learning, probabilistic graphical knowledge models, global structure.

1. Введение. Алгебраические байесовские сети (АБС) представляют собой логико-вероятностную графическую модель систем знаний с неопределенностью [13, 30]. Среди родственных им моделей можно также указать байесовские сети доверия (БСД) [32–34] и скрытые марковские модели (СММ) [28, 29, 31], от которых АБС отличает в первую очередь возможность работать как со скалярными (точечными), так и с интервальными оценками, заданными на пропозициональных формулах [4, 13]. Оперирование не просто вероятностям, но интервалами вероятностей, делают АБС применимыми в областях, в которых БСД не хватает точности, а работа со скалярными оценками вероятностей обусловливает возможность моделирования работы БСД [2, 3] и СММ [5, 8] на основе структуры АБС.

Выделяют первичную и вторичную структуры АБС [6, 13, 15]. Первичной структурой называют набор фрагментов знаний (максимальных фрагментов знаний), которые задают систему знаний с неопределенностью. Вторичная структура характеризует связи между этими фрагментами знаний, поэтому естественно представлять ее в виде графа (так называемого графа смежности). Алгоритмы пропагации поступившего в сеть свидетельства и синтеза согласованных оценок в значительной степени опираются на этот граф, поэтому выбор вторичной структуры критичен как для времени работы указанных алгоритмов, так и вообще их для осуществимости [7, 9–12, 14, 15].

С указанной точки зрения наличие циклов в графах смежности представляется серьезным затруднением для работы упомянутых алгоритмов. По этой причине большой интерес представляет изучение множества минимальных графов смежности, элементы которого имеют наименьшее число циклов среди всех графов смежности, построенных над данным набором фрагментов знаний [16, 22, 23]. Минимальность графа смежности не гарантирует отсутствия циклов [24–26]: для большого класса первичных структур минимальные графы смежности не будут являться деревьями смежности, т.е. будут иметь циклы. Поэтому актуальной является задача исследования циклов в графах смежности и выявления условий, необходимых для того, чтобы циклов в графах смежности не существовало.

Цель данной работы — выявление необходимых и достаточных условий цикличности или ацикличности минимальных графов смежности на основе исследований глобальной структуры АБС и, одновременно, формирование соответствующей системы терминов.

2. Основные определения и обозначения. В данном разделе приводятся определения так, как они были введены в [22].

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: $V(G') = V'$; $E(G') = E'$, где $G' = \langle V', E' \rangle$. Также V от множества ребер будет возвращать множество концов этих ребер: $V(E) = \{v | \exists e \in E, \exists u: e = (v, u)\}$.

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

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

$\subset$ — строгое включение:

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

Алфавитом будем называть множество атомарных пропозициональных формул $A = \{x_1, \ldots, x_n\}$, 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, вершины которого соответствуют элементам множества главных конъюнктов МФЗ, вошедших в АБС.

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 $W(\{V_i, V_j\})$, ребра $\{V_i, V_j\} \in E(G)$, графа G определяется как пересечение весов 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: $W(H) = \bigcap_{V \in H} W(V)$.

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 $\forall\, V_i \in B \quad W(V_b) \cap W(V_e) \subset W(V_i)$. Граф *магистрально связен*, если между каждой парой несовпадающих вершин, веса которых содержат общие элементы, существует магистральный путь.

Благодаря введенным понятиям *граф смежности* определяется как магистрально связный граф МФЗ.

Минимальный граф смежности — граф смежности, число ребер которого минимально. Известно [1], что минимальные графы смежности и только они являются *нередуцируемыми графами смежности*, то есть такими графами смежности, что любые графы, полученные удалением из них любого числа ребер, не будут являться графами смежности. Множество минимальных графов смежности будем обозначать как **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. Для заданного множества вершин существует и при этом единственный максимальный граф смежности [22].

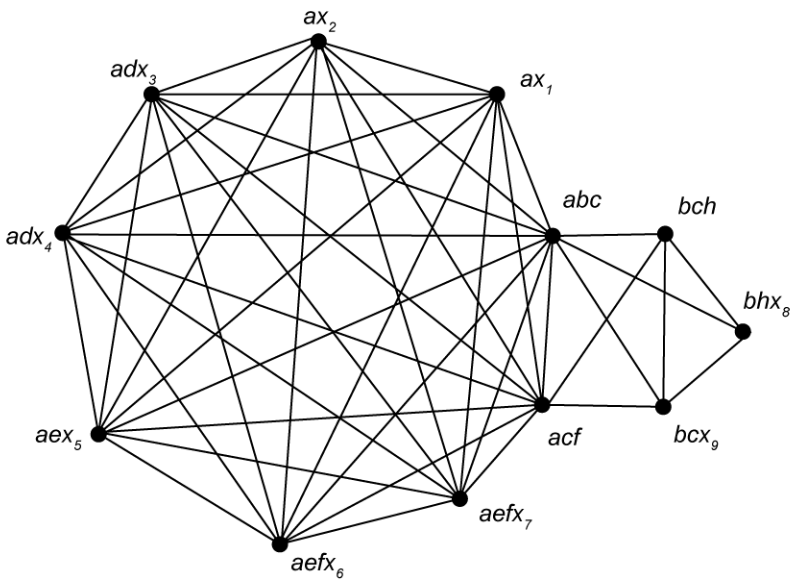

Рис. 1. Максимальный граф смежности.
Максимальный граф смежности на вершинах ax_1, ax_2, adx_3, adx_4, aex_5, $aefx_6$, $aefx_7$, acf, abc, bch, bhx_8, bcx_9.

3. Сужения и клики. В данном разделе приводятся определения так, как они были введены в [27].

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 клик будем обозначать как Clique.

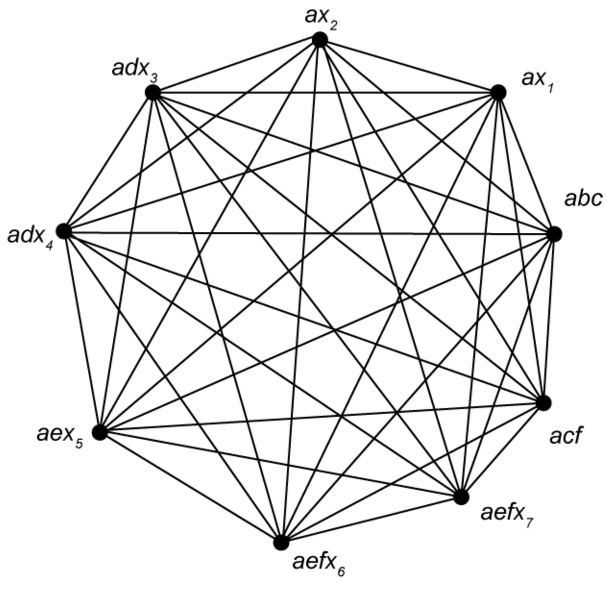

Рис. 2. Клика a.

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

Доказано [22], 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). 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 $\langle \mathrm{Clique}, E_{\mathrm{Clique}} \rangle$, где

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

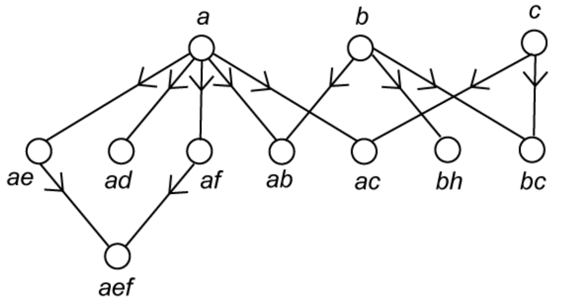

Рис. 3. Граф клик.
Граф клик, построенный для графа на рис. 1.

4. Владения. В данном разделе приводятся определения так, как они были введены в [27].

Сильное сужение $G \downarrow\!\!\!\!\!- U$ — значимое сужение $G \downarrow U$, из которого удалили все ребра веса U:

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

Сильное сужение графа $G_{\max} \downarrow\!\!\!\!\!- U$ представляет собой компоненты связности, на которые разбивается сужение $G_{\max} \downarrow U$ удалением ребер веса U (рис. 4). Такие компоненты называются *владениями.*

Теорема о классификации владений [27]. Любое владение любого сильного сужения $G \downarrow\!\!\!\!\!- U$ является либо доменной вершиной, либо вассалом, либо братством U.

Доменной вершиной называется вершина, принадлежащая клике U и не принадлежащая ни одному из ее сыновей (рис. 5).

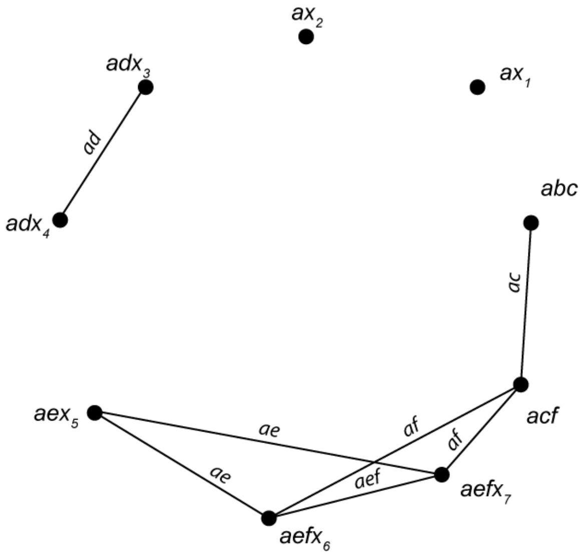

Рис. 4. Сильное сужение.
Сильное сужение графа с рис. 1 на вес a. Оно состоит из 4–х владений.

Вассалом называется 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 $\{V_U^{w_1}, \ldots, V_U^{w_n}\}$, 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 (рис. 5):

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

Полусиблинговый путь между двумя родственными вассалами V_U^i и V_U^j — такой набор вассалов $\{V_U^{w_1}, \ldots, V_U^{w_n}\}$ из определения полусиблингов, что $V_U^b \leftrightarrow V_U^{w_1}$, $V_U^{w_n} \leftrightarrow V_U^e$ и $\forall i < n\ V_U^{w_i} \leftrightarrow V_U^{w_{i+1}}$ (рис. 5).

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

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

Если построить граф, вершинами которого будут вассалы клики U, а ребра между двумя вассалами будут проведены, если их пересечение непусто, то братством будет являться компонента связности такого графа.

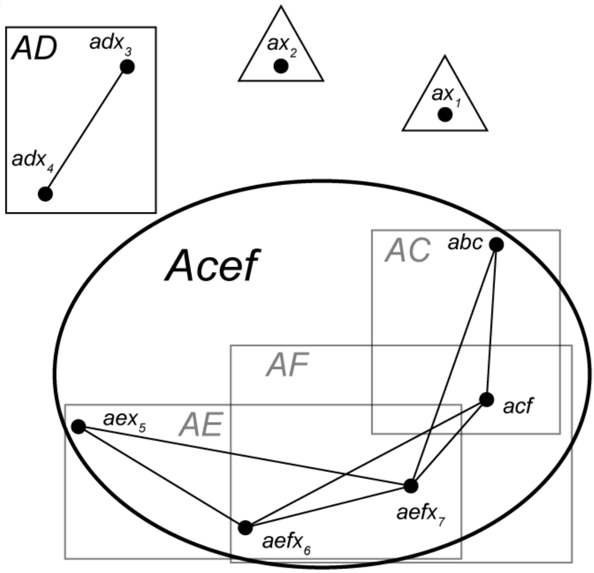

Рис. 5. Владения.

Треугольниками обозначены доменные вершины ax_1 и ax_2; прямоугольником обозначен вассал AD, состоящий из вершин adx_3 и adx_4; овалом обозначено братство $Acef$, состоящее из выделенных полупрозрачными прямоугольниками вассалов AC, AE и AF; вассалы AC и AF, точно так же, как и вассалы AE и AF, приходятся друг другу братьями; вассалы AC и AE являются полусиблингами. Они соединены полусиблинговым тем (AE, AF, AC).

5. Теорема о циклах. Прежде, чем приступать к изучению циклов, нам необходимо ввести дополнительные понятия, которыми мы будем оперировать.

Определение 1. *Праклика* PraC — ненаправленный граф, являющийся сужением максимального графа смежности на пустой вес (т. е. фактически совпадающий с ним).

То же самое, более формально:

Определение 1'. *Праклика* PraC:

$$PraC = G_{max} \downarrow U_{\emptyset} = G_{max},$$

где $U_{\emptyset}$ — пустой вес.

Хотя праклика и не является кликой в том смысле, в котором мы ее определили (она не является сужением на значимый вес и, более того, не является полным подграфом), удобно рассматривать ее в качестве таковой, добавляя корнем в граф клик. Тогда все «настоящие» клики будут являться чьими-либо сыновьями. В частности, это обеспечит связность графа клик (оба этих утверждения будут доказаны ниже).

Определение 2. *Расширенное вверх множество клик* $\mathrm{Clique}^{\uparrow}$ — множество клик, к которому добавлена праклика.

То же самое, более формально:

Определение 2'. Расширенное вверх множество клик $\mathrm{Clique}^{\uparrow}$:

$$Clique^{\uparrow} = Clique \cup \{PraC\}$$

Определение 3. *Замкнутый сверху граф клик*[1] — граф клик, к которому в качестве вершины дерева добавлена праклика.

Определение 3'. Замкнутый сверху граф клик G_{Clique}:

$$\overline{G}^{\uparrow}_{Clique} = \langle Clique^{\uparrow}, E_{Clique^{\uparrow}} \rangle,$$

где $\forall P, Q \in \mathrm{Clique}^{\uparrow}, (P, Q) \in E_{\mathrm{Clique}^{\uparrow}} \Leftrightarrow Q \subset P$ и $\nexists R \in \mathrm{Clique}: Q \subset R \subset$ $\subset P$.

Утверждение 1. Любая клика в замкнутом сверху графе клик является чьим-либо сыном:

$$\forall C \in Clique \ \exists C^{*} \in Clique^{\uparrow}: (C^{*}, C) \in E_{Clique^{\uparrow}}.$$

Доказательство. Рассмотрим произвольную клику C. Либо у нее есть родитель из множества Clique, либо у нее такого родителя нет, но тогда ее родителем в замкнутом сверху графе клик будет праклика, потому вес любой клики содержит вес праклики (пустой вес) и у C по предположению нет родителей.

---

[1] В данном случае замыкание осуществляется относительно операции сужения на вес: теперь сужение даже не пустой вес элемента из замкнутого сверху графа клик будет оставаться элементом замкнутого сверху графа клик. Однако замкнутый сверху граф клик остается «открытым снизу», поскольку сужение на вес конкретной вершины не будет являться элементом замкнутого сверху графа клик, и тем более графа клик. Однако в данной статье нам достаточно ограничиться рассмотрением замкнутого сверху графа клик, а не строить полностью замкнутый граф клик.

Следует обратить внимание, что праклика не является чьим-нибудь сыном, но она не является и кликой.

Утверждение 2. Ненаправленный граф, который соответствует замкнутому сверху графу клик (такой граф получается вследствие утраты ориентации всеми ребрами замкнутого сверху графа клик) является связным.

Доказательство. Так как вес любой клики содержит вес праклики, то любая клика является потомком праклики, а потому соединена с ней путем. Таким образом, в ненаправленном графе, соответствующем замкнутому сверху графу клик, любые две вершины будут соединены путем, который получается объединением путей, соединяющих обе вершины с пракликой.

Утверждение 3. Замкнутый сверху граф клик не содержит направленных циклов.

Доказательство. Не отличается от доказательства аналогичного утверждения для графа клик [22]. От клики C можно добраться до клики D, если вес клики D содержит вес клики C. Поэтому если от клики C можно добраться до клики D и от клики D можно добраться до клики C, то вес клики C одновременно содержит и включается в вес клики D, что противоречит тому, что для каждого веса по определению существует ровно одна клика.

Определение 4 [22]. *Обязательное ребро* — ребро, которое входит в любой минимальный граф смежности, построенный над заданным набором вершин (обозначим множество таких ребер как NecessaryEdges).

То же самое, более формально:

Определение 4'. Обязательное ребро $e \in E(G)$:

$$\forall M \in \boldsymbol{MJG}\; e \in E(M).$$

Определение 5. *Обязательный цикл* — цикл, образованный обязательными ребрами.

Очевидно, что условие существования обязательного цикла во вторичной структуре является достаточным для того, чтобы все минимальные графы смежности, построенные над заданным набором вершин, не являлись ациклическими. Однако это условие не является необходимым для существования циклов в графе смежности.

Пример 1. Рассмотрим набор вершин ad, ab, bcd и cdx (рис. 6.*а*). Обязательных ребер всего три: $(ad, ab), (ab, bcd), (bcd, cdx)$ (рис. 6.*б*). Как видно, они не образуют обязательного цикла. Однако же оба минимальных графа смежности, построенные на данном наборе вершин (рис. 6.*в–г*) будут содержать циклы.

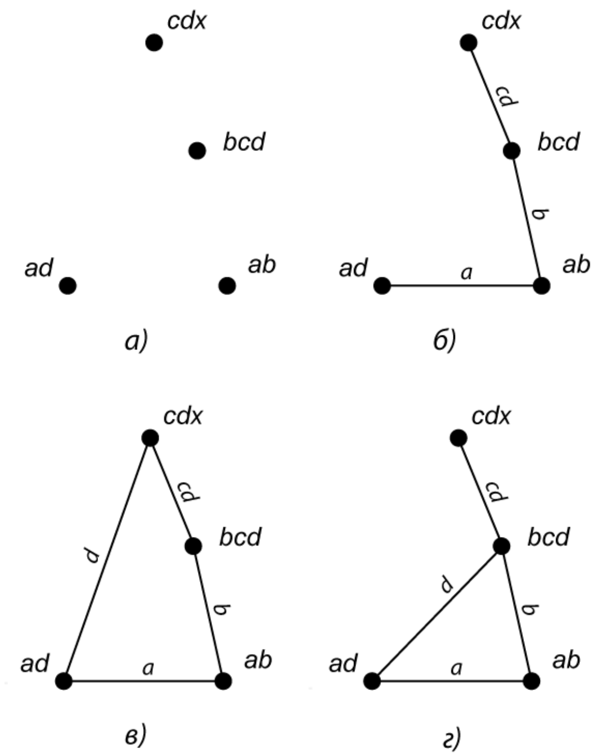

Рис. 6. Циклическая вторичная структура без обязательных циклов.
На рис. 6.*а* изображен набор вершин; на рис. 6.*б* — обязательные ребра для этого набора вершин; на рис. 6.*в* и рис. 6.*г* — минимальные графы смежности, построенные над данным набором вершин и содержащие циклы.

Для поиска необходимых и достаточных условий обратимся к рассмотрению владений.

Определение 6. *Полусиблинговый цикл* — цикл, состоящий из вассалов, в братстве какой-либо клики.

То же самое более формально:

Определение 6'. Полусиблинговый цикл — последовательность вассалов $S_1, S_2, \dots, S_n$, такая, что $\forall i, 1 < i < n-1$, $S_i \leftrightarrow S_{i+1}$; $S_n \leftrightarrow S_1$ и $n > 2$.

Определение 7. *Братский цикл* — полусиблинговый цикл $S_1, S_2, \dots, S_n$, такой, что пересечение всех этих вассалов непусто (рис. 7)

То же самое более формально:

Определение 7'. Братский цикл — полусиблинговый цикл $S_1, S_2, \dots, S_n$, такой, что $\bigcap_{i=1}^{n} S_i \neq \emptyset$.

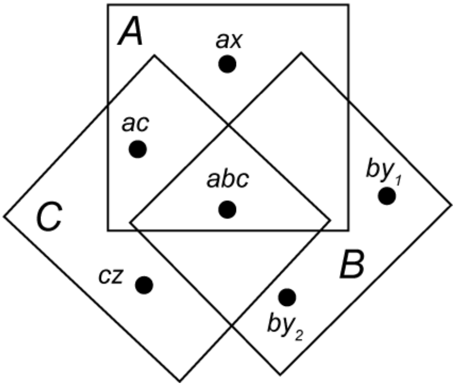

Рис. 7. Братский цикл.

Братский цикл, состоящий из вассалов A, B и C, которые пересекаются по вершине abc.

Определение 8. *Небратский цикл* — полусиблинговый цикл, не являющийся братским (рис. 8).

То же самое более формально:

Определение 8'. *Небратский цикл* — полусиблинговый цикл $S_1, S_2, \dots, S_n$, такой, что $\bigcap_{i=1}^{n} S_i = \emptyset$.

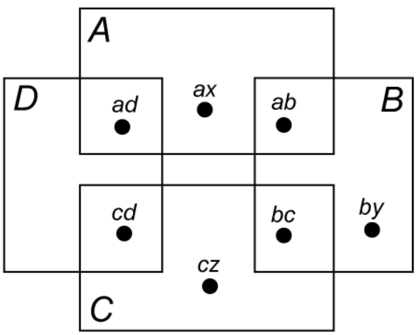

Рис. 8. Небратский цикл.
Небратский цикл, состоящий из вассалов A, B, C и D.

Замечание 1. В братском цикле все вассалы, в него входящие, приходятся друг другу братьями, тогда как в небратском цикле это условие может как выполняться, так и не выполняться.

Для дальнейших исследований нам потребуется лемма о независимом магистральном пути, которая первоначально сформулирована в терминах тораксов [21], однако мы решили не воспроизводить в данной статье весь обширный понятийный аппарат, связанный с ними, а ограничиться переформулировкой леммы на языке терминов, используемых в данной статье.

Лемма (о независимом пути) [21]. Для любого несвязного подграфа G любого значимого сужения G^* любого графа смежности существует пара вершин v и u из G, таких, что существует магистральный путь между v и u, не пересекающийся по ребрам с G.

Теорема (о циклах минимальных графов смежности). Любой минимальный граф смежности, построенный над заданным набором вершин, является ациклическим тогда и только тогда, когда не существует полусиблинговых небратских циклов.

Доказательство. Докажем сначала в одну, а потом в другую сторону.

$\Rightarrow$. Пусть в минимальном графе смежности нет циклов. Докажем, что для этого графа не существует полусиблинговых циклов.

Пускай существует полусиблинговый цикл, состоящий из вассалов $V_1, V_2, \ldots, V_n$. Для того, чтобы граф был магистрально связным, все вассалы должны быть магистрально связны. Таким образом, от любой вершины из V_1 до любой вершины из $V_1 \backslash V_n$ будет существовать как путь, проходящий только по ребрам внутри этих двух вассалов, так и путь, проходящий последовательно по ребрам вассалов $V_1, V_2, \ldots, V_n$, то есть в графе есть цикл.

$\Leftarrow$. Пускай в минимальном графе смежности есть циклы. Докажем, что в нем есть полусиблинговые циклы. Рассмотрим произвольный цикл C_0, состоящий из ребер e_1, $e_2, \ldots, e_n$ с весами $U_1, U_2, \ldots, U_n$ соответственно. Найдем вес U_0, который содержат все ребра (он может быть и пуст). Поскольку мы рассматриваем именно минимальный граф смежности, в цикле не может оказаться ребер с весом, в точности равным U_0 (иначе любой магистральный путь, проходящий через такое ребро, можно заменить на путь, проходящий по противоположной дуге цикла, все ребра которого содержат этот же вес — магистральная связность сохранится, то есть ребро можно удалить, а это значит, что граф не был минимальным).

Все ребра рассматриваемого цикла содержатся в клике веса U_0, и, поскольку ни одно из этих ребер не имеет веса, в точности равного U_0, то каждое ребро этого цикла содержится хотя бы в одном из сыновей клики (или праклики) с весом U_0.

Распределим ребра цикла между сыновьями U_0, то есть каждому ребру поставим в соответствие произвольного сына U_0, содержащего данное ребро. Таким образом мы получим цикл $(S_{p_1}, S_{p_2}, \dots, S_{p_m})$ из сыновей клики U_0, который, возможно, содержит самопересечения. Этот цикл описывает то, по каким сыновьям проходит первоначальный цикл. Объединим в цикле всех подряд идущих сыновей: $(\dots, S_{fp}, S_i, S_i, S_i, \dots, S_i, S_{lp}, \dots) \to (\dots, S_{fp}, S_i, S_{lp}, \dots)$, где $fp \neq i$ и $lp \neq i$ (передвигаясь по ребрам одного и того же сына, мы не выходим за его пределы). Полученный цикл назовем C_0^S.

К сожалению, пока что мы не получили искомый полусиблинговый цикл, поскольку цикл сыновей с самопересечениями вовсе необязательно будет являться полусиблинговым циклом: это может быть просто цепь из сыновей, по которой прошли сначала в одну, а потом в другую сторону.

Будем исключать самопересечения. Для этого воспользуемся леммой о независимом пути. Рассмотрим все вхождения одного и того же сына S_i в цикл C_0^S, предположив, что их больше одного. Эти вхождения соответствуют в первоначальном цикле цепям (то есть деревьям со степенью вершины не больше 2) — дугам цикла. Выберем все эти цепи и заменим их на полные графы. При этом, во-первых, объединение таких подграфов не будет связным (потому что полные подграфы не пересекаются по вершинам с другими полными подграфами, и, следовательно, не связны, причем таких подграфов по предположению больше одного), а во-вторых, множество ребер этих подграфов содержит вес $V(S_i)$, поэтому является подмножеством ребер соответствующего значимого сужения на этот вес. Все это позволяет нам применить указанную лемму, согласно которой существуют две такие вершины в этих полных подграфах, между которыми существует магистральный путь, не пересекающийся ни с одним из этих подграфов. То есть мы можем найти магистральный путь, который соединяет два вхождения одного и того же сына, и при этом не пересекается ни с каким из вхождений этого сына в цикл. Добавим также, что сам этот магистральный путь лежит в сыне, так как соединяет две вершины, принадлежащие сыну.

Заменим одну из дуг исходного цикла на найденный магистральный путь — в новом цикле из сыновей станет хотя бы на одно самопересечение меньше, потому что новых вхождений сына не добавилось, тогда как два старых были заменены на одно (при этом, возможно, еще какие-то оказались отброшенными).

Повторяя описанные действия, мы теоретически можем достигнуть двух результатов: либо получить цикл сыновей без самопересечений длины не меньше 3 (и тогда это будет искомый полусиблинговый цикл), либо получить цикл, состоящий ровно из двух сыновей.

Второй случай будет соответствовать циклу, в котором одна дуга лежит внутри одного сына S_1, а оставшаяся дуга — внутри другого сына S_2. Рассмотрим две вершины этого цикла, которые являются концами ребер с разными весами. Обе они попадают в пересечение двух сыновей, а, значит, магистрально связны путем M с весом $W(S_1) \cap W(S_2)$. Как минимум одна из двух дуг, на которые цикл разбивается этими точками, не совпадает с этим путем. Из этой дуги можно выкинуть произвольное ребро, сохраняя магистральную связность через путь M, а, значит, граф не был минимальным.

Таким образом, если в минимальном графе смежности есть цикл, то тогда существует и полусиблинговый цикл.

Пример 2. Рассмотрим цикл $(bc, abt), (abt, act), (act, bc)$ в графе на рис. 9. Все ребра этого цикла являются обязательными и лежат в сужениях на веса b, at и c соответственно. С точки зрения замкнутого сверху графа клик, клика at является сыном клики a, которая вместе с кликами b и c является сыном праклики.

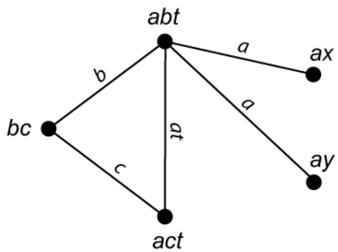

Рис.9. Минимальный граф смежности с циклом.

Мы можем построить соответствующий ей полусиблинговый небратский цикл (рис. 10).

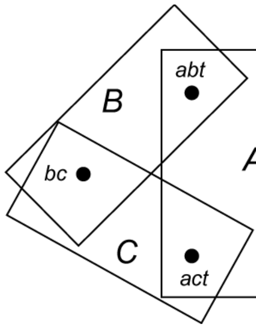

Рис. 10. Полусиблинговый небратский цикл.
Полусиблинговый небратский цикл, порождающий цикл на рис. 9.

Пример 3. Рассмотрим полусиблинговый небратский цикл (рис. 11). Все минимальные графы, которые могут быть построены на соответствующем наборе вершин, изображены на рис. 6.*в*–*г*.

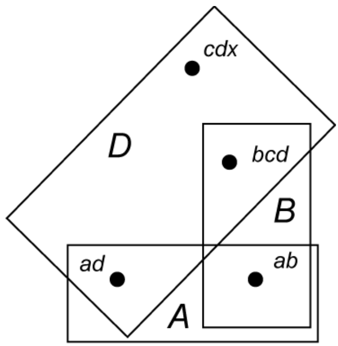

Рис. 11. Полусиблинговый небратский цикл.
Полусиблинговый небратский цикл, которому соответствуют циклические минимальные графы смежности на рис. 6.*в*–*г*.

Пример 4. Еще один немаловажный пример — ациклический граф смежности, который строится по полусиблинговому братскому циклу. Построим такой минимальный граф смежности для уже данного нами в качестве иллюстрации на рис. 7 братского цикла (рис. 12).

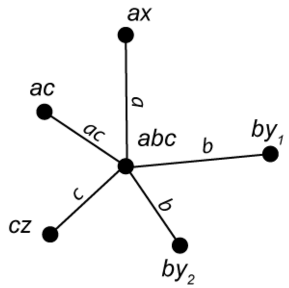

Рис. 12. Ациклический граф смежности.
Ациклический граф смежности, построенный на наборе вершин, изображенных на рис. 7.

Следствие 1. Все графы смежности, построенные над данным набором вершин, являются циклическими или ациклическим одновременно.

Следствие 1 позволяет оценивать набор вершин на цикличность его минимальной вторичной структуры еще до построения этой самой структуры. Информация об этом может быть получена путем построения и исследования замкнутого сверху графа клик — третичной структуры, синтез которой в некоторых случаях предваряет построение минимальных графов смежности [15, 22, 23, 26].

6. Заключение. В работе исследуются циклы в минимальных графах смежности и выявляются условия их существования.

Введено понятие *замкнутого сверху графа клик*, который представляет собой граф клик, в который в качестве вершины добавили *праклику* — сужение максимального графа смежности на пустой вес (что фактически совпадает с самим максимальным графом смежности).

Кроме этого введено понятие *полусиблинговых циклов*, которые представляют собой последовательность вассалов, в которой каждые два соседа, а также первый и последний вассалы являются братьями. Полусиблинговые циклы делятся на *братские циклы* — полусиблинговые циклы, у которых пересечение всех вассалов непусто, и *небратские полусиблинговые циклы*, у которых это пересечение пусто.

На основе введенных понятий сформулирована и доказана первая теорема о циклах, утверждающая, что необходимым и достаточным условиями ацикличности минимальных графов смежности является отсутствие небратских полусиблинговых циклов у каждой клики. Следствие этой теоремы утверждает, что все минимальные графы смежности, построенные над одним и тем же набором вершин, являются ациклическими или циклическими одновременно.

В статье активно применялась классификация владений клик. Накопленные знания об этих объектах заставляют нас выделить их в отдельную структуру алгебраической байесовской сети — *четвертичную* (наименование *третичной структуры* закрепилось за графом клик, также достаточно подробно рассмотренным в этой статье). Четвертичная структура, в отличие от первичной, вторичной и третичной, не едина для конкретной сети, но существует для каждой клики. Данная статья достаточно убедительно показывает преимущества исследования этой структуры, начало которому положила работа [27].

Полученные результаты позволяют выявлять циклы в минимальных графах смежности без их построения. Это особенно важно с точки зрения оценки первичной структуры относительно возможности построения над ней АБС: для того, чтобы оценить, является ли она циклической, не нужно строить все вторичные структуры, возможные над данной, или даже одну такую структуру. Вместо этого можно ограничиться построением третичной структуры и искать небратские полусиблинговые циклы в четвертичных структурах.

Полученные теоретические результаты требуют алгоритмизации. Кроме того, сохраняется интерес к критериям выявления циклов во вторичной структуре за счет анализа первичной или третичной структур — т.е. тех структур, построение которых требует меньших временных затрат по сравнению с четвертичными (в рамках существующих алгоритмов построения множества минимальных графов смежности [17–20]). Одновременно с этим интерес представляет усовершенствование построения четвертичной структуры (пока только как составной части алгоритмов построения множества минимальных графов смежности [17, 18]).

Литература

1. *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.
2. *Момзикова М.П., Великодная О.И., Пинский М.Я., Сироткин А.В., Тулупьев А.Л., Фильченков А.А.* Оценка вероятности наблюдаемой последовательности в бинарных линейных по структуре скрытых марковских моделях с помощью апостериорного вывода в алгебраических байесовских сетях // Труды СПИИРАН. 2010. Вып. 2 (13). С. 122–142.

3. *Момзикова М.П., Великодная О.И., Пинский М.Я., Сироткин А.В., Тулупьев А.Л., Фильченков А.А.* Представление бинарных линейных по структуре скрытых марковских моделей в виде алгебраических байесовских сетей // Труды СПИИРАН. 2010. Вып. 1 (12). С. 134-150.
4. *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.
5. *Сироткин А.В.* Представление байесовской сети доверия с бинарными переменными в виде алгебраической байесовской сети // Интегрированные модели, мягкие вычисления, вероятностные системы и комплексы программ в искусственном интеллекте. VI–я научно-практическая конференция (Коломна, 16–19 мая 2011 г.). Сборник научных трудов. В 2-х т. Т. 2. М.: Физматлит, 2011, С. 992–1000.
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.* 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.
8. *Тулупьев А.Л.* Байесовские сети доверия и алгебраические байесовские сети: сравнительный анализ выразительной мощности // Информационные технологии и интеллектуальные методы. 1997. Вып. № 2. С. 121–147.
9. *Тулупьев А.Л.* Байесовские сети: логико-вероятностный вывод в циклах. СПб.: Изд-во С.-Петербургского ун-та, 2008. 140 с. (Элементы мягких вычислений).
10. *Тулупьев А.Л.* Непротиворечивость оценок вероятностей в алгебраических байесовских сетях. Вестник СПбГУ. Сер. 10. 2009. Вып. 3. С. 144–151.
11. *Тулупьев А.Л.* Преобразование ациклических байесовских сетей доверия в алгебраические байесовские сети // Известия высших учебных заведений: Приборостроение. 2009. № 3. С. 21–23.
12. *Тулупьев А.Л.* Согласованность данных и оценка вероятности альтернатив в цикле стохастических предпочтений // Известия высших учебных заведений: Приборостроение. 2009. № 7. С. 3–8.
13. *Тулупьев А.Л., Николенко С.И., Сироткин А.В.* Байесовские сети: логико-вероятностный подход. СПб.: Наука, 2006. 607 с.
14. *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.
15. *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.
16. *Тулупьев А.Л., Столяров Д.М., Ментюков М.В.* Представление локальной и глобальной структуры алгебраической байесовской сети в Java-приложениях // Труды СПИИРАН. 2007. Вып. 5. СПб.: Наука, 2007. С. 71–99.
17. *Фильченков А.А.* Алгоритм построения множества минимальных графов смежности при помощи самоуправляемых клик // Труды СПИИРАН. 2010. Вып. 1 (12). С. 119–133.
18. *Фильченков А.А.* Алгоритм построения множества минимальных графов смежности при помощи самоуправляемых клик-собственников // Труды СПИИРАН. 2010. Вып. 3 (14) С. 150–169.

19. *Фильченков А.А.* Алгоритм построения множества минимальных графов смежности при помощи клик владений // Труды СПИИРАН. 2010. Вып. 1 (13). С. 119–133.
20. *Фильченков А.А.* Алгоритм построения множества минимальных графов смежности при помощи клик-собственников владений // Труды СПИИРАН. 2010. Вып. 4 (15). С. 193–212.
21. *Фильченков А.А., Тулупьев А.Л.* Понятие торакса в применении к исследованию графов смежности алгебраических байесовских сетей // Труды СПИИРАН. 2011. Вып. 16. С. 186–205.
22. *Фильченков А.А., Тулупьев А.Л.* Структурный анализ систем минимальных графов смежности Труды СПИИРАН. 2009. Вып. 11. С. 104–127.
23. *Фильченков А.А., Тулупьев А.Л., Сироткин А.В.* Компаративный анализ клик минимальных графов смежности алгебраических байесовских сетей // Труды СПИИРАН. 2010. Вып. 13. С. 87–105.
24. *Фильченков А.А., Тулупьев А.Л., Сироткин А.В.* Мощность множества минимальных графов смежности // Труды СПИИРАН. 2010. Вып. 4 (15). С. 136–161.
25. *Фильченков А.А., Тулупьев А.Л., Сироткин А.В.* Особенности анализа вторичной структуры алгебраической байесовской сети // Труды СПИИРАН. 2010. Вып. 1 (12). С. 97–118.
26. *Фильченков А.А., Тулупьев А.Л., Сироткин А.В.* Ребра графов смежности в контексте компаративного анализа клик минимальных графов смежности алгебраических байесовских сетей // Труды СПИИРАН. 2010. Вып. 14. С. 132–149.
27. *Фильченков А.А., Тулупьев А.Л., Сироткин А.В.* Структурный анализ клик минимальных графов смежности // Вестник Тверского государственного университета. Сер. Прикладная математика. 2011. Вып. 2.
28. *Baum L.E., Petrie T., Soules G., Weiss N.* A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains // Ann. Math. Stat. 41. 1970. P. 164–171.
29. *Cowell R.G., Dawid A.P., Lauritzen S.L., Spiegelhalter D.J.* Probabilistic Networks and Expert Systems. NY.: Springer-Verlag, 1997. 370 p.
30. *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.
31. *Huang X., Jack M., Ariki Y.* Hidden Markov Models for Speech Recognition., NY.: Columbia University Press. 1990. 276 p.
32. *Pearl J.* Distributed Revision of Composite Beliefs. // Artificial Intelligence, vol. 33. 1987. P. 173–215.
33. *Pearl J.* Fusion, propagation, and structuring in belief networks. // Artificial Intelligence, vol. 29, 1986. P. 241–288.
34. *Pearl J.* Probabilistic Reasoning in Intelligent Systems. Morgan Kaufmann, 1988. 552 p.

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 — 23. 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 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 — 23. 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., профессор; 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 — 250. 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. 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 — 250. 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.

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

Работа поступила в редакцию 01.07.2011.

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 гранта 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. **Анализ циклов в минимальных графах смежности алгебраических байесовских сетей.**

Работа ряда алгоритмов алгебраических байесовских сетей (АБС) опирается на ее вторичную структуру, представляемую графом смежности, и особенную трудность для применения этих алгоритмов создает наличие циклов во вторичной структуре, поэтому важно уметь определять, существуют ли циклы в графах смежности, представляющих вторичные структуры над данной первичной структурой без непосредственного их построения этих графов.

В работе расширена терминологическая основа исследований минимальных графов смежности благодаря добавлению понятия *праклика*, соответствующего сужению максимального графа смежности на пустой вес, а также производных от этого понятий, в частности, понятия *замкнутого сверху графа клик*, образованного добавлением праклики к графу клик. Про этот граф доказано, что любая клика в нем является чьим-то сыном, а также то, что в нем нет направленных циклов и что соответствующий ему неориентированный граф связен.

Продемонстрировано, что отсутствие циклов, образованных обязательными ребрами, не является достаточным условием отсутствия циклов в графе.

Далее, введены понятия *братского* и *небратского полусиблинговых циклов*, которые представляют собой циклы из сыновей клики и различаются тем, что пересечение всех элементов первого непусто, тогда как второго — пусто.

Сформулирована теорема о циклах в минимальных графах смежности, утверждающая, что необходимым и обязательным условием цикличности минимального графа смежности является существование небратского полусиблингового цикла. Теорема доказана с опорой на лемму о независимом пути.

Из теоремы следует, что все минимальные графы смежности, построенные над данной первичной структурой, являются циклическими или ациклическими одновременно.

Полученный результат позволит оценивать первичную структуру с точки зрения возможности построения над ней АБС без построения и анализа не только всех вторичных структур, возможных над данной, но даже и одной такой структуры.

ABSTRACT

Filchenkov A.A., Tulupyev A.L. **The Algebraic Bayesian Network Minimal Join Graphs Cycles Analysis.**

Functioning of a number of algebraic Bayesian networks (ABN) algorithms is based on their secondary structure represented as a join graph. Existence of cycles in the secondary structure especially complicates these algorithms application, so it is important to determine whether there are cycles in the join graph representing secondary structure on given primary structure without direct synthesis of these graphs.

We extended the terminological basis of minimal join graph research by adding the concept of praclique, corresponding to a narrowing of the maximum join graph by the empty weight, and also by adding the derivatives of this concept, in particular, by adding the concept of clique graph closed from above resulted by adding the praclique to the clique graph. It is proven that any clique in this graph is a son of some element of this graph. Also it is proven that there are no directed cycles in the graph and that the corresponding undirected graph is connected.

It is shown that the absence of loops of necessary edges is not a sufficient condition for absence of cycles in the graph.

Further, the terms of brotherly half-sibling cycle and unbrotherly half-sibling cycle are suggested, which represent the cycles of the clique sons of cliques, and differ in that the intersection of all elements of the first is nonempty, while the same intersection of the least is not empty.

The theorem on minimal join graph cycles which states that a necessary and sufficient condition for the minimal join graph cyclisity is the existence of an unbrotherly half-sibling cycle is formulated. The theorem is proven with application of the lemma of independent paths.

The theorem implies that all minimal join graphs on given primary structure are cyclic or acyclic simultaneously.

This obtained result allows to estimate the primary structure in terms of the possibility to synthesize an ABN without synthesizing and analyzing not only all possible secondary structures on the primary structure, but even without synthesizing even one such structure.