УДК 004.8

А.А. Фильченков

АЛГОРИТМ ПОСТРОЕНИЯ МНОЖЕСТВА МИНИМАЛЬНЫХ ГРАФОВ СМЕЖНОСТИ ПРИ ПОМОЩИ КЛИК-СОБСТВЕННИКОВ ВЛАДЕНИЙ

*Фильченков А.А.***Алгоритм построения множества минимальных графов смежности при помощи клик-собственников владений.**
Аннотация.Существуетэффективный алгоритм построения множества минимальных графов смежности по заданному набору максимальных фрагментов (при помощи самоуправляемых клик), а также два улучшения, каждое из которых реализуется в отдельном алгоритме; однако нет алгоритма, который бы реализовал оба улучшения. Целью данной работы является создание такого алгоритма, который бы реализовывал одновременно ряд улучшений базового алгоритма, вследствие чего он был бы более эффективным, чем существующие.Такой алгоритм был предложен, его корректность доказана.
Ключевые слова:алгебраические байесовские сети, вторичная структура, машинное обучение, вероятностно-графические модели систем знаний, глобальная структура.

*Filchenkov A.A.***Minimal join graph set synthesis proprietor possession cliques algorithm.**
Abstract. The effective minimal join graph set synthesis algorithm (of self-managing cliques)from a given maximal knowledge pattern set and two its improvements, each of which is actualized in an algorithm, exists, however there is no algorithm that implements the two improvements. The goal of the work is to designing a new algorithm that actualizes all known improvements of the basic algorithm, and thus, is the most effective one. Such an algorithm implementing suggested improvement is designed and correctness of its work is proven.
Keywords: algebraic Bayesian networks, secondary structure, machine learning, probabilistic graphical models, global structure.

1. Введение.Алгебраические байесовские сети (АБС) принадлежат к классу логико-вероятностных графических моделей, являясь родственниками байесовских сетей доверия и скрытых марковских моделей. АБС представляют собой вероятностную графическую модель баз фрагментов знаний с неопределенностью, способную работать как скалярными, так и с интервальными оценками [1–2, 15, 28].АБС позволяют поддерживать непротиворечивость во фрагментах знаний, осуществлять логико-вероятностные априорный и апостериорный выводы. Алгоритмическая реализация последних опирается на вторичную структуру АБС [11–14, 16], которая характеризует связи между фрагментами знаний (ее первичной структурой) и представляется в виде графа смежности [7, 15, 17].

Так, например, пропагация свидетельства распространяется по ребрам графа смежности, поэтому крайне важная структура этого графа, потому что при «плохой» структуре свидетельство будет распространяться значительное количество времени, а в худшем случае алгоритмы логико-вероятностных выводов не могут быть применены — причиной этому являются циклы в графах смежности, которые могут быть обработаны далеко не всегда. Поэтому с описанной точки зрения лучшей вторичной структурой будет являться дерево смежности [10].

Однако дерево смежности не может быть построено над многими первичными структурами [22]. По этой причине особый интерес представляет изучение минимальных графов смежности, которые совпадают с деревьями смежности, если они возможны для данного набора фрагментов знаний, и являются наиболее «эффективными» в указанном смысле вторичными структурами в общем случае [4–8].

Однако, несмотря на свою значимость, множество минимальных графов смежности было изучено недостаточно.Основной работой, открывающей это направление исследований, является статья [22], в которой был сформулирован понятийный аппарат и были получены основные результаты, в том числе фундаментальная теорема о множестве минимальных графах смежности, являющаяся теоретическим фундаментом для всех существующих алгоритмов построения множества минимальных графов смежности, в том числе и для базовой схемы этого алгоритма, предложенной в той же работе. Ряд статей [3, 23–27] развивает достигнутые теоретические результаты. Следует выделить статьи по классификации владений клик ([27])и компаративному анализу клик ([23, 26]), которые легли в основу для улучшений алгоритма построения множества минимальных графов смежности, на которых основаны соответствующие алгоритмы ([19–21]). В статье [21] предложено улучшение за счет построения владения через сыновей клик, а в статье [20] — улучшение за счет учета основного множества ребер. Два указанных улучшения не являются взаимоисключающими, что открывает возможность создать новый алгоритм, который бы реализовывал оба указанных улучшения с целью ускорить построение множества минимальных графов смежности.

Цель данной работы — предложить алгоритм построения минимальных графов смежности, реализующий одновременно все известные пять улучшений для подобных алгоритмов.

2. Основные определения и обозначения.В данном разделе приводятся определения так, как они были введены в [24]. Данная статья содержится в этом же выпуске, поэтому определения приведены максимально сжато, опущен формализм, иллюстрации и пояснения.

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), $i \neq j, v_i,\ v_j \in V$.

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

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.
2) никакое слово полностью не содержит никакое другое слово.

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

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

Вес $W(V_i)$ *вершины* $V_i \in V(G)$ — множество атомов алфавита, вошедших в V_i. *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. *Вес* $W(H)$ *подграфа* $H \subseteq G$ — наибольшее по включению слово, которое входит в веса всех его вершин.

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.**

Граф смежности — магистрально связный граф МФЗ. *Минимальный граф смежности* — граф смежности, число ребер в котором минимально. Будем обозначать множество минимальных графов смежности через **MJG**. *Максимальный граф смежности* $G_{\max}$ — наибольший по числу ребер граф смежности.

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:

Любое слово, являющееся весом какого-либо ребра графа G, будем называть *значимым словом графа* G, а сужение графа G на такое слово — *значимым сужением*.

Клика U — значимое сужение максимального графа смежности на вес U. Любая клика является полным подграфом графа $G_{\max}$. Множество всехтаких клик будем обозначать как Clique.

Граф клик — направленный граф, вершинами которого являются клики из множества Clique. 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.

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. *Владение* P_U^i — множество вершин i-й компоненты связности сильного сужения $G_{\max} \not\downarrow U$.

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

Вассал V_U клики U — множество вершин, входящих в какого-либо сына клики U.

Два вассала V_U^i и V_U^j называются *братьями*, 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*, 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:

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

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

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

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

Сжатие σ_U компоненты связности $P_U^i \subseteq G \nleftarrow U$ в вершину f_i — отображение на множестве подмножеств вершин, сопоставляющее множеству вершин P_U^i вершину f_i.

Сжатие σ_U множества ребер

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

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 $|E_{i,j}|$.

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 $G \nleftarrow U$, 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.

Феод — f_i — вершина, получившаяся сжатием какого-то владения.

Курия веса U — K_U — ненаправленный граф с кратными ребрами, полученный сжатием значимого сужения $G \downarrow U$.

Оммаж H_U — курия K_U, являющаяся деревом, все ребра которой имеют кратность, равную единице. Любое сжатие минимального графа смежности является оммажем.

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

Пучок — граф, построенный на исходном наборе вершин, множество ребер которого равно объединению жил, выбранных по одной для каждого значимого слова.

Теорема (о множестве минимальных графов смежности) [22, 24]. Множество минимальных графов смежности совпадает с множеством пучков.

Следствие 1[22]. Число ребер в минимальных графах смежности одинаково.

Следствие 2[22]. Множество минимальных графов смежности совпадает с множеством пучков, которое равно декартовому произведению множеств жил каждой клики.

Следствие 3[22]. Мощность множества графов смежности равна произведению мощностей множеств жил каждой клики.

Следствие 4[22]. Согласно следствию 2, для того, чтобы построить множество графов смежности, достаточно для каждой клики построить множество соответствующей ей жил.

3. Классификация клик и алгоритм построения множества минимальных графов смежности при помощи самоуправляемых клик. *Собственное ребро клики* — ребро, принадлежащее клике, вес которого совпадает с весом клики.

По числу собственных ребер множество клик можно поделить на:

- *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:

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

Таблица 1. **Классификация клик**

По ребрам \ По детям	Бездетные	Родительские
Безреберные	C_0^-	C_0^+
Однореберные	C_1^-	C_1^+
Многореберные	C_n^-	C_n^+

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. Ее единственное ребро является *обязательным* — оно входит во все минимальные графы смежности [23].

Моноклика-1 — C_1^+ — клика, имеющая ровно одно собственное ребро и имеющая детей. Ребро это клики является *избыточным* — оно не входит ни в один из минимальных графов смежности [23].

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

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

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

Утверждение 1 [26]. Основное множество вершин стереоклики совпадает с множеством ее вершин.

Базовая схема алгоритма построения множества минимальных графов смежности [22].

Над множеством вершин V строится максимальный граф смежности $G_{\max}$. По этому графу строится граф клик Clique. Каждой клике сопоставляется множество вершин, которые в нее попали. На графе клик задается порядок.

По заданному порядку перебираются все клики из упорядоченного дерева клик. Для каждой клики при помощи алгоритма Прюфера перебираются все оммажи, поскольку оммаж является деревом — и для каждого оммажа перебираются все соответствующие ему жилы. Таким образом, для каждой клики перебираются все жилы, соответствующие ее весу. Все такие жилы записываются в массив жил для данной клики.

Согласно теореме о множестве минимальных графов смежности, нам достаточно построить все возможные пучки, чтобы получить все минимальные графы. Это можно сделать, перебрав всевозможные комбинации жил для каждой клики, объединяя такие жилы в единый граф.

В [19] было предложено три улучшения базового алгоритма, а также сам алгоритм построения множества минимальных графов смежности при помощи самоуправляемых клик.

Улучшение 1 (исключение незначимых сужений) [19]. Вместо того чтобы рассматривать все возможные клики, мы можем исключить из рассмотрения моноклики-0, так как они не создают никаких жил и перебор оставшихся клик будет сведен к перебору весов ребер максимального графа смежности.

Улучшение 2 (исключение клик с единственным владением) [19]. Вместо того чтобы строить жилы для моноклик-1 и моноклик-n, мы можем исключить их из множества клик, для которых строятся жилы, что ускорит вывод минимальных графов смежности, так как нужно будет перебирать меньшее число клик.

Улучшение 3 (априорный учет ребер однореберных бездетных клик) [19]. Вместо того чтобы строить жилы для биклик, мы можем исключить их из множества клик, для которых строятся жил, добавив их ребра в граф обязательных ребер, что ускорит вывод минимальных графов смежности, так как нужно будет перебирать меньшее число клик

Алгоритм построения минимальных графов смежности при помощи самоуправляемых клик (self-managingcliquesalgorithm) строит по заданному набору вершин V, соответствующих множеству главных конъюнктов максимальных ФЗ, все возможные минимальные графы смежности.

```
Require: V
Ensure: MJG = {⟨V, E_i⟩}
Weights = ∅
for all u, v ∈ V, (u ≠ v)&(W(u)⋂W(v) ≠ ∅) do
   Weighs ← W(u)⋂W(v)
end for
Cliques = ∅
for all v ∈ V do
   for all w ∈ Weights
      if w ⊂ W(v) do
         Cliques → C_w : W(C_w) = w
         C_w ⟵ u
         Cliques ⟵ C_w
      end if
   end for
end for
NecessaryEdges = ∅
StereoIndex = ∅
for all C_w ∈ Cliques do
   if |C_w| = 2 then
      C_w → v
      C_w → u: u ≠ v
      NecessaryEdges ⟵ (v, u)
   endif
   else
      P = компоненты связности строго сужения
      клики C_w
      if |P| > 1 then
         StereoIndex ⟵ w
         S_w = ∅
         for all t is_a дерево на P do
```

for all s **is_a** жила для t **do**
$S_w \longleftarrow s$
end for
end for
end if
end else
end for
$MJG = \emptyset$
forall $\{s_w^i\}$: $\{s_w^i\}$ —индексирующая последовательность жил стереоклик $w \in StereoIndex, s_w^i \in S_w$ **do**
$E = NecessaryEdges$
for all $s \in \{s_w^i\}$ **do**
$E = E \cup s.edges$
end for
$MJG \longleftarrow \langle V, E \rangle$
end for
return MJG

Листинг 1. Алгоритм построения минимального графа смежности при помощи самоуправляемых клик

Здесь и далее $S \leftarrow e$ обозначает добавление элемента e к множеству S, а $S \rightarrow e{:}\,Cond(e)$обозначает извлечение элемента, удовлетворяющего условиям $Cond$, из множества S в переменную e.

В работе [20, 21] было предложено два различных улучшения к приведенному выше алгоритму, и созданы алгоритмы, реализующие эти улучшения по отдельности.

Улучшение 4 (построение владений через сыновей) [21]. Вместо того, чтобы строить владения клик путем поиска компонент связности соответствующего строго сужения, мы можем строить владения как компоненты связности относительно родства вассалов.

Улучшение 5 (учет основного множества ребер) [20]. Вместо того, чтобы сначала строить множествовесов ребер, а потом сужать на каждый вес набор вершин, мы можем для каждого веса ребра построить основное множество вершин соответствующей ему клики. Благодаря этому мы можем сразу выделить клики с единственным собственным ребром (у них основное множество вершин состоит всего из двух вершин) и, проверив, кто из них является, а кто не является родителем, можем сразу построить биклики и исключить из рассмотрения моноклики-1.

Важно отметить, что улучшение 5 предполагает вычисление для каждой клики, является ли она родительской или нет, что уже включено в построение графа клик, который требуется для улучшения 4, поэтому алгоритм, который мы собирается создать, не будет механическим объединением двух указанных улучшений.

4. Алгоритм построения множества минимальных графов смежности при помощи клик-собственников владений.

Алгоритм построения минимальных графов смежности при помощи клик-собственников владений (proprietorpossessingcliquesalgorithm) строит по заданному набору вершин V, соответствующих множеству главных конъюнктов максимальных ФЗ, все возможные минимальные деревья смежности графа МФЗ.

Require: V

Ensure: $MJG = \{\langle V, E_i \rangle\}$

$Cliques = \emptyset$

for all $u, v \in V, (u \neq v) \& (W(u) \cap W(v) \neq \emptyset)$ **do**

  $w = W(u) \cap W(v)$

  $Cliques \rightarrow C_w : W(C_w) = w$

  $C_w \longleftarrow u$

  $C_w \longleftarrow v$

  $Cliques \longleftarrow C_w$

endfor

$ICl =$ множество индексов $Cliques$

$NecessaryEdges = \emptyset$

$StereoIndex = \emptyset$

for $w =$ последний по порядку вес из ICl **to** первый по порядку вес из ICl **do**

  $C = Cliques[w]$

  $Domains[w] = Cliques[w]$

  $C.Sons = \emptyset$

  for $w^* =$ последний по порядку вес из ICl **to** идущий по порядку после w вес из ICl **do**

  $C^* = Cliques[w^*]$

  if $W(C) \subset W(C^*)$ **then**

  $C.Sons = C.Sons \setminus C^*.Sons$

  $C.Sons \leftarrow C^*$

$Domains[w] = Domains[w] \backslash Domains[w^*]$
endif
w =предыдущий по порядку вес изICl
end for
if$|Domains[w]| > 2$**then**
$IntersEdges = \emptyset$
for all$S_1, S_2 \in C.Sons, (S_1 \neq S_2)\&(S_1 \cap S_2) \neq \emptyset$ **do**
$IntersEdges \leftarrow (S_1, S_2)$
end for
$P =$ компоненты связности $\langle C.Sons, IntersEdges \rangle$
$P \longleftarrow Domains[w]$
if$|P| > 1$ **then**
$StereoIndex \leftarrow w$
$Sinews[w] = \emptyset$
for all t**is_a**дереванaP **do**
for all$s$**is_a**жиладляt**do**
$Sinews[w] \leftarrow s$
end for
end for
end if
end if
else
if$C.Sons = \emptyset$**then**
$Domains[w] \rightarrow v$
$Domains[w] \rightarrow u : u \neq v$
$NecessaryEdges \leftarrow (v, u)$
end if
endelse
w =предыдущий по порядку вес изICl
end for
$MJG = \emptyset$
for all$\{s_w^i\} : \{s_w^i\}$ – индексирующая последовательность жил стереоклик $w \in StereoIndex, s_w^i \in S_w$**do**
$E = NecessaryEdges$
for all $s \in \{s_w^i\}$ **do**

```
55:         E = E ∪ s.edges
56:      end for
57:      MJG ← ⟨V, E⟩
58:   end for
59:   return MJG
```

Листинг 2. Алгоритм построения минимального графа смежности при помощи клик-собственников владений.

В цикле (2–8) строится основное множество вершин для каждого значимого сужения.

В цикле (12–50) последовательно (в обратном лексикографическом порядке, заданном на весах) перебираются все клики, исключаются из рассмотрения моноклики, особым образом обрабатываются биклики, строится множество обязательных ребер и для всех стереоклик строится множество их жил.

В цикле (16–24) перебираются все клики, имеющий вес, следующий за весом рассматриваемой клики и строится множество сыновей данной клики, а также множество ее собственных вершин.

Условный оператор (18–22) в случае, если вес рассматриваемой клики C содержится в весе перебираемых в цикле (16–24) клик C^*, вычитает из множества сыновей клики C множество сыновей клики C^*, добавляет C^*в множество сыновей клики C и вычитает из множества собственных вершин клики C собственные вершины клики C^*.

Условный оператор (25–41) в случае, если мощность основного множества вершин клики больше двух, строит ее оммажи и жилы.

В цикле (27–29) строится множество ребер на сыновьях рассматриваемой клики: ребро между двумя сыновьями существует тогда, когда множество их вершин пересекается.

Условный оператор (32–40) в случае, если число владений у клики больше одного, строит множество всех возможных жил для данной клики.

В цикле (35–39) перебираются все деревья, построенные на феодах клики (т.е. оммажи). Для этого, как правило, используется алгоритм Прюфера.

В цикле (36–38) в массив жил для данной клики добавляются все жилы, соответствующие данному оммажу.

Условный оператор (42–48) вызывается в случае, когда мощность основного множества вершин клики равна двум. Он обрабатывает биклики и игнорирует моноклики-1.

Условный оператор (43–46) вызывается в случае, если у клики, у которой мощность основного множества вершин равна двум, нет сыновей (т.е. она биклика). Он добавляет две ее вершины к множеству обязательных ребер $NecessaryEdges$.

В цикле (52–58) алгоритм перебирает все кортежи, индексирующие жилы для каждой клики. Каждый кортеж индексирует набор жил, выбранных по одной для каждой клики. В цикле происходит объединение всех жил одного кортежа с ребрами из множества обязательных ребер. Граф, содержащий эти ребра, добавляется к множеству минимальных графов смежности.

Утверждение 2. Алгоритм построения множества минимальных графов смежности при помощи клик-собственниковвладений строит множество минимальных графов смежности.

Доказательство.В цикле (2–8) будет построено основное множество вершин для каждой клики, содержащей более двух вершин, поскольку к соответствующим множествам вершин для каждого ребра добавляются его концы.

В цикле (16–24) будут построены все сыновья и доменные вершины рассматриваемой клики, потому что вес сыновей клики идет после веса самой клики в лексико-графическом порядке, ее доменные вершины — это ее вершины за исключением тех, которые входят в сыновей (строка 21), а сыновья определяются как клики, чей вес содержит вес рассматриваемой клики (условие в строке 18), и они не являются ее прямыми потомками (т.е. в первом колене, строка 19 исключает из множества сыновей всех «внуков», то есть сыновей предполагаемых сыновей).

В условном операторе (32–40) будет построено множество всех жил для данной клики, если у нее более двух владений (т.е. она является стереокликой), что следует из определения жилы и оммажа. Моноклики игнорируются согласно улучшениям 1–2.

Условный оператор (42–48) будет обрабатывать биклики, поскольку только они удовлетворяют поставленным условиям.

В цикле (12–50) будут рассмотрены все клики, для стереоклик будет построено множество жил, собственные ребра биклик будут добавлены в граф обязательных ребер, а моноклики будут проигнорированы.

Цикл (52–58) перебирает все наборы жилы по одной для каждой клики стереоклики и объединяет их с множеством обязательных ребер. Согласно теореме о минимальных графах смежности, мы перебираемтаким образом все возможные множества ребер минимальных графов смежности.

Хотя мы склонны ожидать, что алгоритм построения множества минимальных графов смежности при помощи клик-собственников владений будет лучшим по времени работы среди уже существующих, однако отметим, что это требует формального доказательства, поэтому вопрос о времени работы этого алгоритма остается открытым.

5. Заключение. Был рассмотрена базовая схема построения множества минимальных графов смежности, а также 5 улучшений, предложенных в работах [19–21]. Так, алгоритм построения множества минимальных графов смежности при помощи самоуправляемых клик реализует первые три улучшения к базовому алгоритму [19], алгоритм построения множества минимальных графов смежности при помощи клик владений реализует четвертое улучшение к алгоритму самоуправляемых клик [21], а алгоритм самоуправляемых клик владений — пятое улучшение к алгоритму самоуправляемых клик [20]. Таким образом, существующие пять улучшений никогда не применялись в одном алгоритме.

Предложен алгоритм построения множества минимальных графов смежности, реализующий все пять указанных улучшений. Доказана корректность алгоритма.

Дальнейший анализ глобальной структуры алгебраических байесовских сетей позволит обособить и структурировать новые элементы, специальная работа с которым ляжет в основу дальнейших улучшений алгоритма построения множества минимальных графов смежности.Кроме того, предложенные в работах [19–21] и в данной статье алгоритмы требуют сравнения по времени работы, что может быть осуществлено либо теоретической оценкой времени их работы (что, в свою очередь, потребует дальнейших теоретических изысканий в теории глобальной структуры АБС), либо в ослабленной форме — экспериментальным путем, что, однако, также представляется достаточно трудоемким, поскольку вопрос о репрезентативной выборке для всех возможных первичных структур также остается открытым вопросом: разнообразие и сложность алгебраических байесовских сетей пока не позволяют делать какие-либо предположения о множестве таких структур в целом.

Литература

1. *Городецкий В.И.* Алгебраические байесовские сети — новая парадигма 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. *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 // 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.
7. *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).
8. *Тулупьев А.Л.* Алгебраические байесовские сети: реализация логико-вероятностного вывода в комплексе java-программ // Труды СПИИРАН. СПб.: Наука, 2009. Вып. 8. С. 191–232.
9. *Тулупьев А.Л.* Алгебраические байесовские сети: система операций глобального логико-вероятностного вывода // Информационно-измерительные и управляющие системы. 2010. № 11. С 65–72.
10. *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.
11. *Тулупьев А.Л.* Байесовские сети: логико-вероятностный вывод в циклах. СПб.: Изд-во С.-Петербургского ун-та, 2008. 140 с. (Элементы мягких вычислений.)
12. *Тулупьев А.Л.* Непротиворечивость оценок вероятностей в алгебраических байесовских сетях. Вестник СПбГУ. Сер. 10. 2009. Вып. 3. С. 144–151.
13. *Тулупьев А.Л.* Преобразование ациклических байесовских сетей доверия в алгебраические байесовские сети // Известия высших учебных заведений: Приборостроение. 2009. № 3. С. 21–23.
14. *Тулупьев А.Л.* Согласованность данных и оценка вероятности альтернатив в цикле стохастических предпочтений // Известия высших учебных заведений: Приборостроение. 2009. № 7. С. 3–8.
15. *Тулупьев А.Л., Николенко С.И., Сироткин А.В.* Байесовские сети: логико-вероятностный подход. СПб.: Наука, 2006. 607 с.

16. *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.
17. *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.
18. *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.
19. *Фильченков А.А.* Алгоритм построения множества минимальных графов смежности при помощи самоуправляемых клик // Труды СПИИРАН. 2010. Вып. 1 (12). С. 119–133.
20. *Фильченков А.А.* Алгоритм построения множества минимальных графов смежности при помощи самоуправляемых клик-собственников // Труды СПИИРАН. 2010. Вып. 3 (14) [в печати].
21. *Фильченков А.А.* Алгоритм построения множества минимальных графов смежности при помощи клик владений // Труды СПИИРАН. 2010. Вып. 1 (13). [в печати]
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. 2 (13). [в печати]
24. *Фильченков А.А., Тулупьев А.Л., Сироткин А.В.* Мощность множества минимальных графов смежности // Труды СПИИРАН. 2010. Вып. 4 (15). [в печати]
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. 2 (14). [в печати]
27. *Фильченков А.А., Тулупьев А.Л., Сироткин А.В.* Структурный анализ клик минимальных графов смежности // Вестник Тверского государственного университета. Сер. Прикладная математика. 2010. [в печати].
28. *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 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 — 11. 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. Научныйруководитель — 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 — 11. 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. Scientificadvisor — A.L. Tulupyev.

РекомендованоТиМПИСПИИРАН, зав. лаб. А.Л. Тулупьев, д.ф.-м.н., доцент.

Работа поступила в редакцию 07.12.2010.

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. **Алгоритм построения множества минимальных графов смежности при помощи клик-собственников владений.**

Существует несколько известных алгоритмов построения множества минимальных графов смежности, реализующих ряд улучшений базового алгоритма (они перечислены ниже). Однако не существует алгоритма, который бы реализовал все пять представленных здесь улучшений.

Цель данной работы — предложить алгоритм построения минимальных графов смежности, реализующий одновременно все известные улучшения для подобных алгоритмов.

Представлены определения и свойства объектов из теории глобальной структуры алгебраических байесовских сетей, в частности, классификация владений клик и классификация клик графов смежности. Приведена базовая схема алгоритма построения множества минимальных графов смежности, пять ее улучшений и известный улучшенный алгоритм построения множества минимальных графов смежности (при помощи самоуправляемых клик).

Базовый алгоритм состоит из следующих шагов. Над множеством вершин строится максимальный граф смежности. По этому графу строится граф клик. Каждой клике сопоставляется множество вершин, которые в нее попали. На графе клик задается порядок. По заданному порядку перебираются все клики из упорядоченного дерева клик. Для каждой клики при помощи алгоритма Прюфера перебираются все оммажи, поскольку оммаж является деревом — и для каждого оммажа перебираются все соответствующие ему жилы. Таким образом, для каждой клики перебираются все жилы, соответствующие ее весу. Все такие жилы записываются в массив жил для данной клики. Согласно теореме о множестве минимальных графов смежности, нам достаточно построить все возможные пучки, чтобы получить все минимальные графы. Это можно сделать, перебрав всевозможные комбинации жил для каждой клики, объединяя такие жилы в единый граф.

Первое улучшение состояло в том, чтобы рассматривать при построении графа клик только клики, то есть только те подграфы максимального графа смежности, вес которых совпадает с весом какого-либо из ребер максимального графа смежности.

Второе улучшение состояло в том, чтобы рассматривать при переборе клик из графа клик только те клики, которые имеют более одной компоненты связности при удалении из них ребер веса, равного весу клики (т.е. имеющих ровно одно владение).

Третье улучшение состояло в том, чтобы клики, состоящие ровно из двух вершин, не использовать для построения особых множеств ребер — жил, а выделять из них единственное ребро, которое добавлять к специальному графу — графу обязательных ребер.

Четвертое улучшение состояло в том, чтобы строить владения путем поиска компонент связности не на множестве вершин клики, а на множестве вассалов, к которым далее добавляются доменные вершины.

Пятое улучшение состоит в том, чтобы строить множество вершин стереоклик не прямым перебором, а путем построения основного множества вершин — множества концов ребер с весом, совпадающим с весом клики.

Предложен алгоритм построения множества минимальных графов смежности, реализующий все описанные улучшения, доказана его корректность.

Отметим, что дальнейший анализ глобальной структуры алгебраических байесовских сетей позволит обособить и структурировать новые элементы, специальная работа с которым ляжет в основу дальнейших улучшений алгоритма построения множества минимальных графов смежности. Кроме того, существующие алгоритмы требуют сравнения по времени работы, что может быть осуществлено либо чистой оценкой времени их, либо экспериментальным путем, что, однако, также представляется достаточно трудоемким.

SUMMARY

Filchenkov A.A. **Minimal join graph set synthesis proprietor possession cliques algorithm**

There are some know minimal join graph set synthesis algorithm that realize number of known the basic algorithm improvements (they are enumerated below). But there is no such an algorithm that implements all the five improvements.

The goal of this work is to suggest the algorithm that actualizes all the improvements.

Definitions and properties of objects of algebraic Bayesian network global structure theory, classification of clique possessions and classification of join graph cliques in particular, are overviewed. Minimal join graph set synthesis algorithm basic scheme, four its improvement and the known minimal join graph set synthesis (self-managed cliques) algorithm are represented.

Basic algorithm has he following steps.Maximal join graph should be built over a given vertex set. A clique graph should be built by that graph. Every clique should be collated to the set of vertexes that belong to the clique. Cliques order should be specified. With the order all the cliques should be sorted out. All homages should be sorted out for every clique by the means of the Pruefer algorithm, because an homage is a tree. Than all sinews should be sorted out for every homage. So that, all the sinews would be sorted out for each clique. All such sinews should be reordered to sinew array for the clique. According to the minimal join graph set theorem, it's enough to design all branches in the purpose to design all minimal join graphs. It can be done by sorting out all sinew combinations chosen one for each clique and uniting them into the one graph.

On the base of observed classification three basic algorithm improvements that base some cliques specifies, are suggested.

The first improvement is to use only the cliques (e.g. only the maximal join graph subgraphs), the weight of which equals to weigh of any maximal join graph edge.

The second improvement is to use only the cliques that have more than one connection components having lost all the edges of the same weight. (e.g. the cliques that has more than one possession).

The third improvement is to use cliques that have only two vertices not in designing sinews, but to add their only edge to a special graph — necessary edges graph.

The fourth improvement is to design possessions by searching connection components not in clique vertex set, but in vassal set to which domain vertex are added aftermath.

The fifth improvement is to design stereoclique vertex set not by entire enumeration but by designing main vertices set — such a set of vertices that belong to edges with the weight that the clique has, is suggested.

A new minimal join graph set synthesis algorithm that actualize all the improvement is designed and its robustness is proven.

We should notice that the further algebraic Bayesian network global structure analysis might be useful in the purpose to create new efficient minimal join graph set designing algorithms. Moreover, the existing algorithms have to be compared, that can be done or by theoretical estimation either by experiments, that also seems to be quite laborious.