*Annals of Mathematics and AI* **35** (2002) 287-294

# Lengths of semigraphoid inferences

František Matúš*

*Institute of Information Theory and Automation (Academy of Sciences of the Czech Republic) and Laboratory of Intelligent Systems (University of Economics, Prague)*

**Abstract.** A subset $\mathcal{L}$ of a semigraphoid $\mathcal{K}$ over $n$ elements is constructed in such a way that starting from $\mathcal{L}$ it is necessary to apply semigraphoid axioms recursively $2^{n-2}-1$ times to arrive at $\mathcal{K}$. This is first known example of exponentially long semigraphoid inference. A comparison is made between local and global inferences. Graphoids and their duals are discussed as well.

## 1. Introduction

A set $\mathcal{L}$ of ordered couples $(ij|K)$, where $ij \subseteq \widehat{n} = \{1, 2, \dots, n\}$, $n \geqslant 1$, is a two-element set and $K \subseteq \widehat{n}$ is disjoint from $ij$, is called *semigraphoid* if it satisfies the single axiom

$$\{(ij|kL), (ik|L)\} \subseteq \mathcal{L} \iff \{(ik|jL), (ij|L)\} \subseteq \mathcal{L}\,.$$

Given a set of couples $\mathcal{L}$ the smallest semigraphoid $\mathfrak{c}(\mathcal{L})$ containing this set can be constructed by means of the *inference sequence* starting at $\mathcal{L}^{(0)} = \mathcal{L}$ and defined by

$$\mathcal{L}^{(r+1)} = \mathcal{L}^{(r)} \cup \bigcup \left\{\{(ik|jL), (ij|L)\}\,;\ \{(ij|kL), (ik|L)\} \subseteq \mathcal{L}^{(r)}\right\}, \qquad r \geqslant 0\,.$$

Obviously, $\mathfrak{c}(\mathcal{L}) = \bigcup_{r\geqslant 0} \mathcal{L}^{(r)}$. The inference sequence is nondecreasing $\mathcal{L}^{(s)} \subseteq \mathcal{L}^{(s+1)}$, and once $\mathcal{L}^{(s)} = \mathcal{L}^{(s+1)}$ this equality is valid also for all $r \geqslant s$. The smallest $s \geqslant 0$ satisfying the equality will be denoted by $s_{\mathcal{L}}$, and called the *length of semigraphoid inference based on* $\mathcal{L}$.

A straightforward question is what is the maximum of $s_{\mathcal{L}}$ for $\mathcal{L} \subseteq \mathcal{R}(\widehat{n})$ where $\mathcal{R}(\widehat{n})$ is the set of all couples $(ij|K)$ with $ij \subseteq \widehat{n}$ and $K \subseteq \widehat{n}$ disjoint. The aim of this note is to construct sets of couples $\mathcal{L}$ having $s_{\mathcal{L}}$ exponentially large in $n$. To this end we identify semigraphoids as special subgraphs of a regular graph with colored edges and interpret

*This research was supported by grant A 1075801 of GA AV ČR and grant No. VS 96008 of Ministry of Education of Czech Republic.

*2000 Mathematics Subject Classification.* Primary 5C15; Secondary 8A02.

*Keywords:* semigraphoids, graphoids, inference length, regular graphs, edge-coloring.

*Permanent postal address:* Institute of Information Theory and Automation, Pod vodárenskou věží 4, 182 08 Prague, Czech Republic; matus@utia.cas.cz

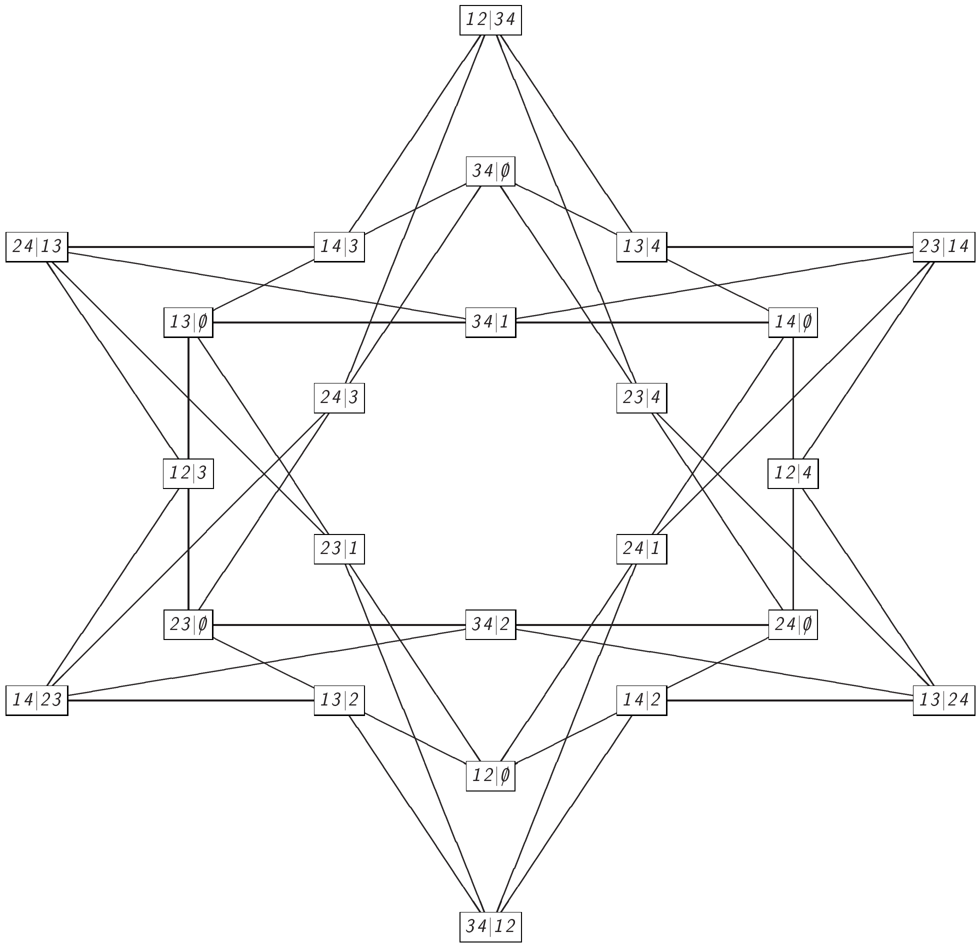
inference as addition of edges, cf. Section 2. An open problem about exponential behaviour of inference is formalized in Section 4. A quicker 'global' inference, see Section 5, is shown to provide an acceleration only at most $n$ times. Finally, inferences in pseudographoids, graphoids and their duals are discussed.

A few conventions and notations follow. Elements of $\widehat{n}$ are not distinguished from singletons and the sign for union of subsets of $\widehat{n}$ is omitted. To avoid parentheses the union has the highest priority among the set operations. The power set of $\widehat{n}$ is $2^{\widehat{n}}$. The cardinality of a set $I$ is denoted by $|I|$. Given $\mathcal{L} \subseteq \mathcal{R}(\widehat{n})$ let $E_{\mathcal{L}} = \{ij;\, \exists K \text{ s.t.} (ij|K) \in \mathcal{L}\}$ and $\mathcal{L}^{ij} = \{K;\ (ij|K) \in \mathcal{L}\}$ for $ij \subseteq \widehat{n}$. Note that $E_{\mathcal{L}} = E_{\mathfrak{c}(\mathcal{L})}$.

## 2. Semigraphoids and colored graphs

A simple undirected graph $\mathcal{G}_n$ with colored edges is constructed for every $n \geqslant 3$ as follows. The vertex set of $\mathcal{G}_n$ is $\mathcal{R}(\widehat{n})$ and its edges are $\{(ij|kL), (ik|L)\}$ where $ijk$ is a three-element subset of $\widehat{n}$ disjoint from $L \subseteq \widehat{n}$. In other words, the neighbours of a vertex $(ij|K)$ are $(ik|K-k)$ and $(jk|K-k)$ for $k \in K$ and $(ik|jK)$ and $(jk|iK)$ for $k \in \widehat{n} - ijK$. Hence, each vertex is incident with $2(n-2)$ edges and the graph $\mathcal{G}_n$ is regular. It has $n(n-1)2^{n-3}$ vertices and only $n(n-1)(n-2)2^{n-3}$ edges. All edges are colored, each color being used twice to match exactly one pair of edges $\{(ij|kL), (ik|L)\}$ and $\{(ik|jL), (ij|L)\}$. The number of colors is thus $n(n-1)(n-2)2^{n-4}$.

For example, $\mathcal{G}_3$ is the cycle of length six having three pairs of the opposite edges colored by three distinct colors. The graph $\mathcal{G}_4$ is depicted below omitting 24 colors.



For a three-element subset $I$ of $\widehat{n}$ disjoint from $L \subseteq \widehat{n}$ the subgraph of $\mathcal{G}_n$ induced by $\{(ij|KL) \in \mathcal{R}(\widehat{n});\ ijK \subseteq I\}$ is a cycle of length six isomorphic to $\mathcal{G}_3$. In the above figure there are eight cycles of this type, and keeping them in mind it is easy to recover all monochromatic pairs of edges.

The graph $\mathcal{G}_n$ is $n-1$ partite since if $\mathcal{R}(\widehat{n})$ is partitioned into $\mathcal{R}_r(\widehat{n})$ according to the cardinality $r = 0, \ldots, n-2$ of $K$ in a couple $(ij|K) \in \mathcal{R}(\widehat{n})$ then each edge has endpoints in $\mathcal{R}_r(\widehat{n})$ and $\mathcal{R}_{r+1}(\widehat{n})$ for some $r = 0, \ldots, n-3$.

It is not difficult to show by induction that $\mathcal{G}_n$ is connected. In fact, using the mapping $\tau_n \colon \mathcal{R}(\widehat{n}) \to \mathcal{R}(\widehat{n+1})$ defined by $(ij|K) \mapsto (ij|K(n+1))$ it is easy to see that the subgraph of $\mathcal{G}_{n+1}$ induced by $\mathcal{R}(\widehat{n}) \cup \tau_n \mathcal{R}(\widehat{n})$ consists of two disconnected subgraphs isomorphic to $\mathcal{G}_n$. Then each of the remaining vertices $(ij|K)$ of $\mathcal{G}_{n+1}$ has $n+1$ in the set $ij$. Such a couple $(ij|K)$ has $|K|$ neighbours in $\mathcal{R}(\widehat{n})$ and $n-1-|K|$ neighbours in $\tau_n \mathcal{R}(\widehat{n})$. Thus the connectedness of $\mathcal{G}_n$ implies the connectedness of $\mathcal{G}_{n+1}$.

A set of couples $\mathcal{L} \subseteq \mathcal{R}(\widehat{n})$ is a semigraphoid if and only if the *subgraph of* $\mathcal{G}_n$ *induced by the vertex set* $\mathcal{L}$ *contains no color exactly once.* In other words, if the subgraph contains an edge it must contain also the second edge of the same color. For example, anticliques (subgraphs without edges, notably the union of $\mathcal{R}_r(\widehat{n})$ over $0 \leqslant r \leqslant n-2$ odd) of $\mathcal{G}_n$ are semigraphoids.

The inference step $\mathcal{L}^{(r)} \mapsto \mathcal{L}^{(r+1)}$ amounts addition of those edges that have their mates contained in $\mathcal{L}^{(r)}$. The reader is invited to play the inference game for the set of couples $\mathcal{L} = \{(13|2), (13|\emptyset), (14|3), (24|1), (23|4)\}$ in the above figure.

## 3. A sequence of semigraphoids

For $n \geqslant 3$ the set $\mathcal{K}_n \subseteq \mathcal{R}(\widehat{n})$ is defined recursively as

$$\mathcal{K}_3 = \big\{(12|3), (13|\emptyset), (13|2), (12|\emptyset)\big\}$$
$$\mathcal{K}_4 = \mathcal{K}_3 \cup \tau_3 \mathcal{K}_3 \cup \big\{(24|1), (24|\emptyset)\big\}$$
$$\mathcal{K}_{n+1} = \mathcal{K}_n \cup \tau_n \mathcal{K}_n \cup \big\{(2(n+1)|13n), (2(n+1)|3n)\big\}$$

where the inductive definition works only for $n \geqslant 4$. We are going to prove that $\mathcal{K}_n$ is a semigraphoid.

First, let us record for later purposes that the cardinality of $\mathcal{K}_n$ is $3 \cdot 2^{n-2} - 2$ and $E_{\mathcal{K}_n} = \{12, 13\} \cup \{2j;\ 4 \leqslant j \leqslant n\}$. An easy induction provides

$$\mathcal{K}_n^{12} = 2^{\widehat{n} \setminus 12}, \quad \mathcal{K}_n^{13} = 2^{\widehat{n} \setminus 13}, \quad \mathcal{K}_n^{24} = 2^{\widehat{n} \setminus 234}, \text{ and}$$
$$\mathcal{K}_n^{2j} = \{K \subseteq \widehat{n};\ 3(j-1) \subseteq K \subseteq 13(j-1)(j+1)\ldots n\}, \quad 5 \leqslant j \leqslant n\,.$$

The key point is that the subgraph of $\mathcal{G}_n$ induced by $\mathcal{K}_n$ looks like


where the symbols $1, 2, \ldots, c$ label distinct colors of edges. There are $2^{n-2}$ connected components, two of them being isolated edges, and each of the remaining ones consisting of two incident edges. The isolated edges are $\{(12|3), (13|\emptyset)\}$ and $\{(12|3n), (13|n)\}$ for $n \geqslant 4$.

The above assertions will be proved by induction. The case $n = 3$ is trivial and the case $n = 4$ can be verified visually from the graph $\mathcal{G}_4$. Let us consider $n \geqslant 4$. Using the mapping $\tau_n$ it is easy to see that the subgraph of $\mathcal{G}_{n+1}$ induced by ${}^{\circ}\mathcal{K}_n = \mathcal{K}_n \cup \tau_n\mathcal{K}_n$ consists of two disconnected copies of $\mathcal{K}_n$ and has therefore four isolated edges

$$\big\{(12|3),(13|\emptyset)\big\} \quad \big\{(12|3n),(13|n)\big\}$$
$$\big\{(12|3(n+1)),(13|(n+1))\big\} \quad \big\{(12|3n(n+1)),(13|n(n+1))\big\}\,.$$

Since $\mathcal{K}_{n+1}$ has two more vertices than ${}^{\circ}\mathcal{K}_n$, $(2(n+1)|13n)$ and $(2(n+1)|3n)$, it suffices to analyse their neighbours.

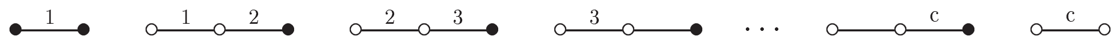
If $(ij|K)$ in ${}^{\circ}\mathcal{K}_n$ is incident with $(2(n+1)|3n)$ then, choosing notation appropriately, $i = 2$. Also, $(ij|K)$ is either $(23|n)$ or $(2n|3)$ or $(2j|3n(n+1))$ for $j \in 145\ldots(n-1)$. But $\mathcal{K}_n^{23}$ is empty and $3 \notin {}^{\circ}\mathcal{K}_n^{2n}$ (this needs a separate verification for $n = 4$ and $n \geqslant 5$). The couple $(2(n+1)|3n)$ is incident with $(2j|3n(n+1))$ for $j = 1$. For $j \in 45\ldots(n-1)$ the set $3n(n+1)$ is not in ${}^{\circ}\mathcal{K}_n^{2j}$ because $3n$ is not in $\mathcal{K}_n^{2j}$ ($3 \in 3n$ for $j = 4$ and $j - 1 \notin 3n$ for $j \geqslant 5$). Addition of the vertex $(2(n+1)|3n)$ to ${}^{\circ}\mathcal{K}_n$ brings thus only one additional edge leading to $(12|3n(n+1))$.

If $(ij|K)$ in ${}^{\circ}\mathcal{K}_n$ is incident with $(2(n+1)|13n)$ then again $i = 2$. Also, $(ij|K)$ is either $(23|1n)$ or $(2n|13)$ or $(21|3n)$ or $(2j|13n(n+1))$ for $j \in 45\ldots(n-1)$. But $\mathcal{K}_n^{23}$ is empty and $13 \notin {}^{\circ}\mathcal{K}_n^{2n}$ (this needs a separate verification for $n = 4$ and $n \geqslant 5$, too). The couple $(2(n+1)|3n)$ is incident with $(12|3n)$. For $j \in 4\ldots(n-1)$ the set $13n(n+1)$ is not in ${}^{\circ}\mathcal{K}_n^{2j}$ because $13n$ is not in $\mathcal{K}_n^{2j}$ by analogous arguments as before. Addition of the vertex $(2(n+1)|13n)$ to ${}^{\circ}\mathcal{K}_n$ brings thus only one additional edge leading to $(12|3n)$.

Since the two edges of $\mathcal{K}_{n+1}$ not in ${}^{\circ}\mathcal{K}_n$ are of the same color one can see that $\mathcal{K}_{n+1}$ has the same structure of connected components as $\mathcal{K}_n$ and the induction step is over.

## 4. Long semigraphoid inference

Let $\mathcal{L}_n$ be union of the left edge of $\mathcal{K}_n$ and the set of right vertices in all three-element components of $\mathcal{K}_n$




The cardinality of $\mathcal{L}_n$ is $2^{n-2}$. Obviously, $\mathcal{L}_n^{(r+1)}$ has two more elements (one more edge) than $\mathcal{L}_n^{(r)}$ for $0 \leqslant r \leqslant 2^{n-2} - 1$ and $\mathcal{L}_n^{(2^{n-2}-1)} = \mathcal{L}_n^{(2^{n-2})} = \mathcal{K}_n$. The length of semigraphoid inference based on $\mathcal{L}_n$ is thus $2^{n-2} - 1$, growing exponentially in $n \geqslant 3$.

Open problem. *Do there exist a polynomial $P$ and a sequence $\mathcal{L}_n \subseteq \mathcal{R}(\widehat{n})$ such that the cardinalities of $\mathcal{L}_n$ are bounded by $P(n)$ and the length of inference based on $\mathcal{L}_n$ grows exponentially in $n$?*

## 5. Global semigraphoid inference

Where $I$, $J$, $K$ are three disjoint subsets of $\widehat{n}$, let

$$(I,J|K)_\star = \{(ij|L) \in \mathcal{R}(\widehat{n})\,;\ i \in I, j \in J, \text{ and } K \subseteq L \subseteq IJK - ij\}\,.$$

Note that $(I,J|K)_\star = (J,I|K)_\star$ is a semigraphoid which is empty if and only if $I$ or $J$ empty. These semigraphoids are called in [3] *elementary.* The inclusion $(I,J|K)_\star \subseteq \mathcal{L}$ is interpreted as conditional independence of $I$ and $J$ given $K$ in $\mathcal{L}$. By induction on the cardinality of $IJ$, one has

$$(I,J|KL)_\star \cup (I,K|L)_\star \subseteq \mathcal{L} \iff (I,JK|L)_\star \subseteq \mathcal{L}$$

for arbitrary disjoint $I, J, K, L \subseteq \widehat{n}$, cf. Lemma 3 of [1], p. 747. We refer to this property of semigraphoids as to the *extended semigraphoid axiom.*

Based on this axiom one can construct the 'global' inference sequence by $\mathcal{L}^{(0)\star} = \mathcal{L}$ and

$$\mathcal{L}^{(r+1)\star} = \mathcal{L}^{(r)\star} \cup \bigcup \left\{(I, JK|L)_\star\,;\ (I, J|KL)_\star \cup (I, K|L)_\star \subseteq \mathcal{L}^{(r)\star}\right\}, \qquad r \geqslant 0\,.$$

Obviously, $\mathfrak{c}(\mathcal{L}) = \bigcup_{r\geqslant 0} \mathcal{L}^{(r)\star}$. The *length of global semigraphoid inference*, denoted by $s^\star_{\mathcal{L}}$, is the minimal $r$ satisfying $\mathcal{L}^{(r)\star} = \mathcal{L}^{(r+1)\star}$. From $\mathcal{L}^{(r)} \subseteq \mathcal{L}^{(r)\star}$ for all $r \geqslant 0$ it is clear that $s^\star_{\mathcal{L}} \leqslant s_{\mathcal{L}}$. However, for the sets of couples $\mathcal{L}_n$ examined in Section 4 the global inference sequence coincides with the (local) inference sequence and one has the equality $s^\star_{\mathcal{L}_n} = s_{\mathcal{L}_n}$.

## 6. Inference in elementary semigraphoids

Let $(I, J|K)_\star$ be a nonempty elementary semigraphoid and let $i_1, \ldots, i_p$ $(j_1, \ldots, j_q)$ be an ordering of all elements of the set $I$ $(J)$. For

$$\mathcal{M}_{p,q} = \{(i_u j_v|\tilde{\imath}_u \tilde{\jmath}_v K)\,;\ 1 \leqslant u \leqslant p,\ 1 \leqslant v \leqslant q\}$$

where $\tilde{\imath}_u = i_1 \ldots i_{u-1}$ and $\tilde{\jmath}_v = j_1 \ldots j_{v-1}$ it is clear that $\mathfrak{c}(\mathcal{M}_{p,q}) \subseteq (I, J|K)_\star$. Using double induction on $p$ and $q$ and the extended semigraphoid axiom, it is not difficult to prove the opposite inclusion, so that $\mathfrak{c}(\mathcal{M}_{p,q}) = (I, J|K)_\star$. More laborious seems to be to write down explicitly the local inference sequence based on $\mathcal{M}_{p,q}$. For our purposes the length of inference $s_{\mathcal{M}_{p,q}}$ will be bounded from above by $p + q$.

To this end one observes that $(i_u, J|\tilde{\imath}_u K)_\star$ is contained in $\mathcal{M}^{(r)}_{p,q}$ for $r = s_{\mathcal{M}_{1,q}}$. Let a couple $(kl|L)$ belong to $(I, J|K)_\star$ where $k \in I$, $l \in J$ and $L = I_1 J_1 K$ for $I_1 = L \cap I \subseteq I - k$ and $J_1 = L \cap J \subseteq J - l$. Then from $(i_u l|\tilde{\imath}_u J_1 K) \in \mathcal{M}^{(r)}_{p,q}$, $1 \leqslant u \leqslant p$, one can deduce that $(I, l|J_1 K)_\star \subseteq (\mathcal{M}^{(r)}_{p,q})^{(t)}$ for $t = s_{\mathcal{M}_{p,1}}$. Thus $(kl|I_1 J_1 K) = (kl|L) \in \mathcal{M}^{(t)}_{p,q}$ for $t$ equal to $s_{\mathcal{M}_{p,1}} + s_{\mathcal{M}_{1,q}}$. As a consequence, $s_{\mathcal{M}_{p,q}} \leqslant s_{\mathcal{M}_{p,1}} + s_{\mathcal{M}_{1,q}}$ and, by symmetry, it suffices to prove that $s_{\mathcal{M}_{1,q}} \leqslant q$. This can be done by induction on $q$ proving in parallel that

$$\mathcal{M}^{(q-1)}_{1,q} \supseteq \{(i_1 j_1|LK)\,; L \subseteq j_2 \ldots j_q\} \cup \{(i_1 j_q|LK)\,; L \subseteq j_1 \ldots j_{q-1}\}\,.$$

Obviously, for $q \leqslant 3$ the inclusion is valid and $s_{\mathcal{M}_{1,q}} \leqslant q$; note that $s_{\mathcal{M}_{1,3}} = 3$. For $q \geqslant 3$ the induction step is first performed for the inclusion. One can start by deriving

$$\mathcal{M}^{(q-1)}_{1,q+1} \supseteq \{(i_1 j_1|LK)\,; L \subseteq j_2 \ldots j_q\} \cup \{(i_1 j_{q+1}|L j_1 K)\,; L \subseteq j_2 \ldots j_q\}\,.$$

For $M \subseteq j_2 \ldots j_q$ the couples $(i_1 j_{q+1}|j_1 MK)$ and $(i_1 j_1|MK)$ are in $\mathcal{M}^{(q-1)}_{1,q+1}$ and therefore $(i_1 j_{q+1}|MK)$ and $(i_1 j_1|j_{q+1} MK)$ are in $\mathcal{M}^{(q)}_{1,q+1}$. Thus $\mathcal{M}^{(q)}_{1,q+1}$ contains all $(i_1 j_1|LK)$ for $L \subseteq j_2 \ldots j_{q+1}$ and all couples $(i_1 j_{q+1}|LK)$ for $L \subseteq j_1 \ldots j_q$. Starting the induction step for the inequality $s_{\mathcal{M}_{1,q}} \leqslant q$ one has $\mathcal{M}^{(q)}_{1,q+1} \supseteq \mathcal{M}^{(q)}_{1,q} = (i_1, j_1 \ldots j_q|K)_\star$. To arrive at $\mathcal{M}^{(q+1)}_{1,q+1} = (i_1, j_1 \ldots j_{q+1}|K)_\star$ one has to verify $(i_1 l|MK) \in \mathcal{M}^{(q+1)}_{1,q+1}$ under the assumption $j_{q+1} \in lM \subseteq j_1 \ldots j_{q+1}$. If $l = j_{q+1}$ then $(i_1 j_{q+1}|MK)$ is in $\mathcal{M}^{(q)}_{1,q+1}$ by the inclusion being proved a short while ago. If $j_{q+1} \in M$ the same inclusion implies that the couple $(i_1 j_{q+1}|l(M - j_{q+1})K)$ is in $\mathcal{M}^{(q)}_{1,q+1}$ what combined with $(i_1 l|(M - j_{q+1})K) \in \mathcal{M}^{(q)}_{1,q}$ provides $(i_1 l|MK) \in \mathcal{M}^{(q+1)}_{1,q+1}$. This means $(i_1, j_1 \ldots j_{q+1}|K)_\star = \mathcal{M}^{(q+1)}_{1,q+1}$ and thus $s_{\mathcal{M}_{1,q+1}} \leqslant q + 1$.

Though the elementary semigraphoid $(I, J|K)_\star$ can have a large cardinality, namely $p\,q\,2^{p+q-2}$, it can be reached from a set of couples $\mathcal{M}_{p,q}$ of the small cardinality $p\,q$ within at most $p + q$ steps of inference.

This observation applies also to the largest semigraphoid $\mathcal{R}(\widehat{n})$. When, for simplicity, $n = 2^m$ and $1 \leqslant t \leqslant m$ the set $\widehat{n}$ can be partitioned into $I_t$ and $J_t$ according to the $t$-th digit from right in the dyadic expansion of $i \in \widehat{n}$. Then $\mathcal{R}(\widehat{n})$, having the cardinality $n(n-1)2^{n-3}$, is covered by $(I_1, J_1|\emptyset)_\star \cup \ldots \cup (I_m, J_m|\emptyset)_\star$, and is reachable from a set of couples of cardinality not more than $\frac{n^2}{4}\log_2 n$ within at most $n$ steps of inference.

## 7. Comparison of semigraphoid inferences

For union of two elementary semigraphoids $\mathcal{L} = (I, J|KL)_\star \cup (I, K|L)_\star \subseteq \mathcal{R}(\widehat{n})$ the length of global inference $s^\star_{\mathcal{L}}$ is clearly equal to one unless some of the sets $I$, $J$ and $K$ is empty. The length of local inference $s_{\mathcal{L}}$ is at most $n$. In fact, ordering elements of each of the sets $I$, $J$ and $K$ with cardinalities $p$, $q$ and $r$, respectively, one can have $\mathfrak{c}(\mathcal{M}_{p,q}) = (I, J|KL)_\star$ and $\mathfrak{c}(\mathcal{M}_{p,r}) = (I, K|L)_\star$ similarly as in the previous section. Then

$$(I, JK|L)_\star = \mathfrak{c}(\mathcal{L}) = \mathfrak{c}(\mathfrak{c}(\mathcal{M}_{p,q}) \cup \mathfrak{c}(\mathcal{M}_{p,r})) = \mathfrak{c}(\mathcal{M}_{p,q} \cup \mathcal{M}_{p,r})$$

where the union $\mathcal{K} = \mathcal{M}_{p,q} \cup \mathcal{M}_{p,r}$ is of the form $\mathcal{M}_{p,q+r}$ for $(I, JK|L)_\star$ with $I$ and $JK$ properly ordered. Therefore, $s_{\mathcal{L}} \leqslant s_{\mathcal{K}} \leqslant p + q + r \leqslant n$.

As a consequence, for all $\mathcal{L} \subseteq \mathcal{R}(\widehat{n})$ one has $\mathcal{L}^{(1)_\star} \subseteq \mathcal{L}^{(n)}$ and by induction $\mathcal{L}^{(r)_\star} \subseteq \mathcal{L}^{(rn)}$, $r \geqslant 0$. Finally, $s_{\mathcal{L}} \leqslant n\, s^\star_{\mathcal{L}}$.

A version of the open problem presented in Section 4 can be formulated replacing the local inference by the global one. This 'global' version is equivalent to the original one because $\frac{1}{n}\, s_{\mathcal{L}} \leqslant s^\star_{\mathcal{L}} \leqslant s_{\mathcal{L}}$. Another version of the open problem emerges when we admit $\mathcal{L}$ to be union of at most $P(n)$ elementary semigraphoids. This is again equivalent to the original version because each elementary semigraphoid can be replaced by a set of couples of the type $\mathcal{M}_{p,q}$, which is of size at most $\frac{n^2}{4}$, whereby the length of inference increases at most $n$ times.

Originally, semigraphoids were defined as ternary relations $\mathbb{I}$ on a set of ordered triples, say $(I, K, J)$, $I, J, K \subseteq \widehat{n}$ disjoint, satisfying four axioms, see [4]. Given such a relation $\mathbb{I}$, the set $\mathcal{L}$ of all couples $(ij|K)$ such that $\mathbb{I}(i, K, j)$ and $\mathbb{I}(j, K, i)$ are true is a semigraphoid in our sense and, vice versa, given a semigraphoid $\mathcal{L} \subseteq \mathcal{R}(\widehat{n})$, the ternary relation $\mathbb{I}$ given by "$\mathbb{I}(I, K, J)$ is true if and only if the elementary semigraphoid $(I, J|K)_\star$ is contained in $\mathcal{L}$" is a semigraphoid in the sense of [4]. We remark that problems on inference in the original framework are covered by the preceding discussion up to technical trivialities.

## 8. Inferences in pseudographoids and graphoids

*Pseudographoids* can be defined (compare [4], p. 97, with [2], p. 102 and Proposition 1), as sets $\mathcal{L}$ satisfying the implication $\Rightarrow$ in

$$(*) \qquad \{(ij|kL), (ik|jL)\} \subseteq \mathcal{L} \iff \{(ij|L), (ik|L)\} \subseteq \mathcal{L}\,.$$

Their duals $\mathcal{L}^\urcorner = \big\{(ij|K) \in \mathcal{R}(\widehat{n})\ ;\ (ij|N - ijK) \in \mathcal{L}\big\}$ obey the opposite implication $\Leftarrow$. The equivalence $(*)$ is for example true for conditional independences among the variables in any nondegenerate Gaussian random vector. *Graphoids* are semigraphoids that are pseudographoids.

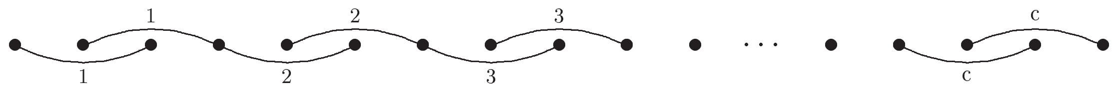
Since the elementary semigraphoids satisfy $(*)$ one can define not only a local but also a global inference for pseudographoids, with lengths $t_{\mathcal{L}} \geqslant t^\star_{\mathcal{L}}$, the latter one using the *extended pseudographoid axiom*

$$(I, J|KL)_\star \cup (I, K|JL)_\star \subseteq \mathcal{L} \iff (I, JK|L)_\star \subseteq \mathcal{L}\,,$$

see the proof of Proposition 1 in [2], p. 105. The inference in pseudographoids is however short; it is left to the reader to show the elementary inequality $t_{\mathcal{L}} \leqslant n-2$ for $\mathcal{L} \subseteq \mathcal{R}(\widehat{n})$. Also, having the set $\{(ij|(I-i)(J-j)K)\,;\, i \in I\,,\, j \in J\}$ on the place of $\mathcal{M}_{p,q}$ it is obvious that its pseudographoid closure is $(I, J|K)_\star$ and that the corresponding length of local pseudographoid inference is $|IJ|-2$ for $I$ and $J$ nonempty.

In graphoids, their duals or in semigraphoids satisfying $(*)$, however, presence of the semigraphoid axiom makes the inference again difficult. Within these classes, various versions of the open problem from Section 4 (see also the end of Section 7) can be easily shown to be equivalent. A technical discussion of these versions within graphoids would readily provide the main result of [5]. This research report, treating graphoids as ternary relations, asks for the first time a question on exponential behaviour of inference in graphoids that is similar to the open problem of Section 4.

A colored graph can be constructed starting from $(*)$ in an analogous way it was done from the semigraphoid axiom and all related considerations go through. Then it is easy to see that its subgraph induced by $\mathcal{K}_n$ has the form




and thus $\mathcal{K}_n$ is a pseudographoid and the dual of a pseudographoid at the same time. For $\mathcal{L}_n$ the corresponding inference, using both the semigraphoid axiom and $(*)$, is the same as within semigraphoids and therefore the corresponding local and global lengths do not decrease, remaining exponential in $n$.

## References


[1] F. Matúš (1992) On equivalence of Markov properties over undirected graphs. *J. Applied Probab.* **29** 745–749.

[2] F. Matúš (1997) Conditional independence structures examined via minors. *The Annals of Math. and Artificial Intelligence* **21** 99–128.

[3] F. Matúš (2001) Towards classification of semigraphoids. (submitted to *Discrete Mathematics*)

[4] J. Pearl (1988) *Probabilistic Reasoning in Intelligent Systems.* Morgan Kaufman: San Mateo, California.

[5] S. Ur and A. Paz (1991) Some aspects of the membership problem for graphoids. Technical Report 677, Technion – Israel Institute of Technology, Haifa, Israel.