Generalized Rules of Probabilistic Independence

Janneke H. Bolt[1,2(✉)] and Linda C. van der Gaag[2]

[1] Department of Information and Computing Sciences, Utrecht University, Utrecht, Netherlands
j.h.bolt@uu.nl
[2] Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, Netherlands
{j.h.bolt,l.c.v.d.gaag}@tue.nl

Abstract. Probabilistic independence, as a fundamental concept of probability, enables probabilistic inference to become computationally feasible for increasing numbers of variables. By adding five more rules to an existing sound, yet incomplete, system of rules of independence, Studený completed it for the class of structural semi-graphoid independence relations over four variables. In this paper, we generalize Studený's rules to larger numbers of variables. We thereby contribute enhanced insights in the structural properties of probabilistic independence. In addition, we are further closing in on the class of probabilistic independence relations, as the class of relations closed under the generalized rules is a proper subclass of the class closed under the previously existing rules.

Keywords: Probabilistic independence · Rules of independence · Semi-graphoid independence relations · Structural semi-graphoid relations

1 Introduction

Probabilistic independence is a subject of intensive studies from both a mathematics and a computing-science perspective [1,2,10]. Pearl and his co-workers were among the first to formalize properties of independence in a system of qualitative rules [2], which characterizes the class of so-called semi-graphoid independence relations. Although the semi-graphoid rules of independence are probabilistically sound, they are not complete for probabilistic independence, as was shown by Studený [4]. While the independences of any discrete multivariate probability distribution adhere to the semi-graphoid rules, a set of independence statements that is closed under these rules, may lack statements that are probabilistically implied. As a consequence, the semi-graphoid rules allow independence relations for which there are no matching probability distributions.

For proving incompleteness of Pearl's system of rules, Studený formulated a new rule for probabilistically implied independence using a proof construct based

J. Vejnarová and N. Wilson (Eds.): ECSQARU 2021, LNAI 12897, pp. 590–602, 2021.
https://doi.org/10.1007/978-3-030-86772-0_42

on the concept of multiinformation. He further defined the class of structural semi-graphoid independence relations as the class of independence relations that are closed under all rules that can be found through such a construct [6–9], and presented a set of five rules that completes the existing rule system for structural relations involving four variables. For an unlimited number of variables, no finite rule system can fully characterize the class of structural semi-graphoid relations, which implies that there is also no finite complete set of rules for the class of probabilistic independence relations [5].

In this paper we generalize Studený's rules of independence to larger numbers of variables. By doing so, we uncover combinatorial structures in rules of probabilistic independence and enable further investigation of such structures. We moreover arrive at an enhanced description of the class of structural semi-graphoid independence relations and thereby close in on the class of probabilistic independence relations, as the class of relations closed under the generalized rules is a proper subclass of the class closed under the previously existing rule system.

2 Preliminaries

We review sets of independence rules and the classes of relations they govern.

2.1 Semi-graphoid Independence Relations

We consider a finite, non-empty set V of discrete random variables and use (possibly indexed) capital letters $A, B, C, \ldots$ to denote subsets of V. We will use concatenation to denote set union and will further abbreviate the union of sets in our figures by concatenating their indices, that is, we write $A_{123}B_{12}$ for $A_1A_2A_3B_1B_2$. A *triplet* over V now is a statement of the form $\langle A, B \mid C\,\rangle$, where $A, B, C \subseteq V$ are pairwise disjoint sets with $A, B \neq \varnothing$. A triplet $\langle A, B \mid C\,\rangle$ is taken to state that the sets of variables A and B are independent given the conditioning set C. Any set of triplets over V is called an *independence relation*.

Pearl introduced the class of so-called semi-graphoid independence relations by formulating four rules of independence [2], which are summarized by [3]:

$A1$: $\langle A, B \mid C\,\rangle \;\leftrightarrow\; \langle B, A \mid C\,\rangle$
$A2$: $\langle A, BC \mid D\rangle \;\leftrightarrow\; \langle A, B \mid CD\rangle \wedge \langle A, C \mid D\rangle$

These two rules are schemata in which the arguments A, B, C, D are to be instantiated to mutually disjoint subsets of V upon application. The rules $A1$, $A2$ are called the semi-graphoid rules of independence, and any independence relation that is closed under these rules is coined a *semi-graphoid independence relation*; in the sequel, we will use $\mathcal{A} = \{A1, A2\}$ to denote the system of semi-graphoid rules. The set $\mathcal{A}$ constitutes a *sound* inferential system for independence relative to the class of discrete multivariate probability distributions. The two rules in $\mathcal{A}$ do not constitute a *complete* system for independence in such probability distributions, however. The system's incompleteness was shown by Studený [4], who formulated the following additional rule:

$$\begin{aligned} A3: \; & \langle A, B \mid CD \rangle \wedge \langle A, B \mid \varnothing \rangle \wedge \langle C, D \mid A \rangle \wedge \langle C, D \mid B \rangle \leftrightarrow \\ & \langle A, B \mid C \rangle \wedge \langle A, B \mid D \rangle \wedge \langle C, D \mid AB \rangle \wedge \langle C, D \mid \varnothing \rangle \end{aligned}$$

For constructing $A3$ and other new rules of independence, he built on the notion of multiinformation, which we will review in the next section.

2.2 Multiinformation and Rules of Independence

A well-known measure for the amount of information shared by two (sets of) variables A, B given a third (set of) variables C in the context of a discrete multivariate probability distribution Pr, is the conditional mutual information (see for example [12]), which is defined as:

$$I(A; B \mid C \mid\mid \Pr) \;=\; \sum_{abc} \Pr(abc) \cdot \log \frac{\Pr(ab \mid c)}{\Pr(a \mid c) \cdot \Pr(b \mid c)}$$

where abc ranges over all possible value combinations for the variables in ABC with $\Pr(a \,|\, c), \Pr(b \,|\, c) \neq 0$. We have that $I(A; B \,|\, C \mid\mid \Pr) \geq 0$ for any A, B, C and Pr, and note that the mutual-information measure is related to independence through the following property: $I(A; B \,|\, C \mid\mid \Pr) = 0$ iff the triplet $\langle A, B \,|\, C \rangle$ is a valid independence statement in Pr. In the sequel, we omit Pr from the notation as long as no ambiguity arises and take the sets A, B, C to be mutually disjoint.

For studying rules of independence, Studený exploited the notion of *multiinformation* [4,11], which is a function M: $2^V \to [0, \infty)$ over all subsets of variables V with $I(A; B \,|\, C) = M(ABC) + M(C) - M(AC) - M(BC)$; for details, we refer to [11]. The multiinformation function thereby has the following properties:

- $M(ABC) + M(C) - M(AC) - M(BC) \geq 0$;
- $M(ABC) + M(C) - M(AC) - M(BC) = 0$ iff $\langle A, B \mid C \rangle$.

The relation between the conditional mutual-information measure and the notion of multiinformation now enables elegant soundness proofs for rules of independence: a rule is sound if all its multiinformation terms 'cancel out', that is, if the multiinformation terms of its set of premise triplets equal the multiinformation terms of its set of consequent triplets. In the sequel, we will refer to this type of proof as a *multiinformation proof construct*.

2.3 Structural Semi-graphoid Independence Relations

Building on the multiinformation concept, Studený introduced the class of *structural semi-graphoid independence relations* [6–9] where, roughly stated, a structural semi-graphoid relation is an independence relation that is closed under all possible rules whose soundness derives from a multiinformation proof construct. In addition to rule $A3$ stated above, Studený formulated four more rules for probabilistic independence found through such proof constructs [6]. We state these rules here in their original form, with their original numbering:

$$\begin{aligned}
A4:\ &\langle A,B \mid CD\rangle \wedge \langle A,D \mid B\rangle \wedge \langle C,D \mid A\rangle \wedge \langle B,C \mid \varnothing\rangle \quad \leftrightarrow\\
&\langle A,B \mid D\rangle \wedge \langle A,D \mid BC\rangle \wedge \langle C,D \mid \varnothing\rangle \wedge \langle B,C \mid A\rangle\\
A5:\ &\langle A,C \mid D\rangle \wedge \langle B,D \mid C\rangle \wedge \langle B,C \mid A\rangle \wedge \langle A,D \mid B\rangle \quad \leftrightarrow\\
&\langle A,C \mid B\rangle \wedge \langle B,D \mid A\rangle \wedge \langle B,C \mid D\rangle \wedge \langle A,D \mid C\rangle\\
A6:\ &\langle A,B \mid C\rangle \wedge \langle A,C \mid D\rangle \wedge \langle A,D \mid B\rangle \quad \leftrightarrow\\
&\langle A,B \mid D\rangle \wedge \langle A,C \mid B\rangle \wedge \langle A,D \mid C\rangle\\
A7:\ &\langle A,B \mid CD\rangle \wedge \langle C,D \mid AB\rangle \wedge \langle A,C \mid \varnothing\rangle \wedge \langle B,D \mid \varnothing\rangle \quad \leftrightarrow\\
&\langle A,B \mid \varnothing\rangle \wedge \langle C,D \mid \varnothing\rangle \wedge \langle A,C \mid BD\rangle \wedge \langle B,D \mid AC\rangle
\end{aligned}$$

In the sequel, we will use $\mathcal{S}$ to denote the rule system $\{A1,\ldots,A7\}$. We note that, while all triplets in the rules $A1$, $A2$ respectively, share a fixed same conditioning (sub-)set of variables, the rules $A3,\ldots,A7$ do not. To equally accommodate additional conditioning variables, the latter five rules can each be enhanced by adding an extra set of variables X to the conditioning parts of its triplets. In phrasing the generalizations of these rules, we will omit such additional sets to conform to the literature.

Any structural semi-graphoid independence relation is closed under the rule system $\mathcal{S}$ by definition. As any semi-graphoid relation is closed under the system of rules $\mathcal{A}$, and $\mathcal{A}$ does not imply the additional rules in $\mathcal{S}$, we have that the class of structural semi-graphoid independence relations is a proper subclass of the class of semi-graphoid relations. The system $\mathcal{S}$ was shown to fully characterize the class of structural semi-graphoid relations over at most four variables [6], that is, the system is both sound and complete for this class. Although sound, the system is not complete for the class of structural independence relations over more than four variables. Studený proved in fact that there is no finite axiomatization of probabilistic independence [5,11], by providing the following rule for all $n \geq 2$ (reformulated):

$$\begin{aligned}
&\langle A,B_0 \mid B_1\rangle \wedge \ldots \wedge \langle A,B_{n-1} \mid B_n\rangle \wedge \langle A,B_n \mid B_0\rangle \quad \leftrightarrow\\
&\langle A,B_0 \mid B_n\rangle \wedge \langle A,B_1 \mid B_0\rangle \wedge \ldots \wedge \langle A,B_n \mid B_{n-1}\rangle
\end{aligned} \tag{1}$$

We note that this rule is a generalization of rule $A6$ in the system S. From the structure of this rule, it is readily seen that any complete system of rules for a fixed number of k variables, will not be complete for $k+1$ variables. The hierarchy of independence relations is depicted in Fig. 1(a).

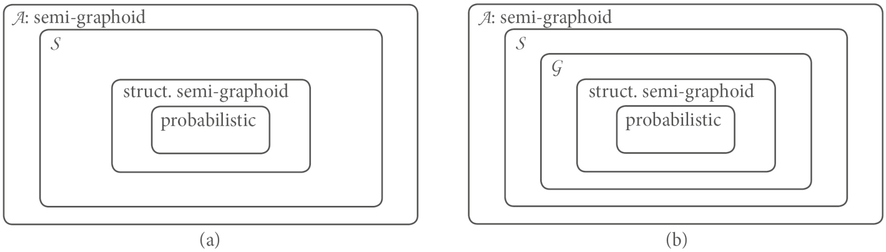

Fig. 1. The existing hierarchy of semi-graphoid, structural semi-graphoid, and probabilistic independence relations (a) and the hierarchy extended with our results (b).

3 Generalizing Inference Rules

For the class of structural semi-graphoid relations, the new rules $A3$–$A7$ of independence have been formulated, involving four sets of variables each (disregarding any fixed additional conditioning context in all triplets). We now reconsider these rules and generalize them to larger numbers of variable sets. By doing so, we arrive at an enhanced characterization of the class of structural semi-graphoid relations and at further insights in the combinatorial structures of independence rules. Because of space restrictions, we cannot detail full soundness proofs for the new rules; instead, we provide brief proof sketches in the appendix.

We begin with formulating our generalization of rule $A5$. The idea underlying its generalization is analogous to the idea of rule $A6$. Informally stated, the rule takes a sequence of sets of variables and moves it (up to symmetry) over the three argument positions of the triplets involved.

Proposition 1. *Let $A_0, \ldots, A_n, n \geq 2$, be non-empty, mutually disjoint sets of variables. Then,*

$$\bigwedge_{i \in \{0,\ldots,n\}} \langle A_i, A_{\mu(i+1)} \mid A_{\mu(i+2)} \rangle \;\; \leftrightarrow \bigwedge_{i \in \{0,\ldots,n\}} \langle A_i, A_{\mu(i+1)} \mid A_{\mu(i-1)} \rangle$$

with $\mu(x) \; := \; x \bmod (n+1)$.

The property stated in the proposition is taken as the independence rule $G5$. Note that in the case of $n = 2$, the rule reduces to a tautology. $G5$ further embeds rule $A5$ as a special case with $n = 3$, as is seen by setting $A_0 \leftarrow C$, $A_1 \leftarrow A$, $A_2 \leftarrow D$ and $A_3 \leftarrow B$. As an example of the generalization, rule $G5$ is now detailed for $n = 4$:

$$\begin{aligned}&\langle A_0, A_1 \mid A_2 \rangle \wedge \langle A_1, A_2 \mid A_3 \rangle \wedge \langle A_2, A_3 \mid A_4 \rangle \wedge \langle A_3, A_4 \mid A_0 \rangle \wedge \langle A_4, A_0 \mid A_1 \rangle \;\; \leftrightarrow \\ &\langle A_0, A_1 \mid A_4 \rangle \wedge \langle A_1, A_2 \mid A_0 \rangle \wedge \langle A_2, A_3 \mid A_1 \rangle \wedge \langle A_3, A_4 \mid A_2 \rangle \wedge \langle A_4, A_0 \mid A_3 \rangle\end{aligned}$$

Note that the rule's consequent cannot be derived from its premise using the existing rule system $\mathcal{S}$ considered thus far. For $n + 1$ sets of variables, rule $G5$

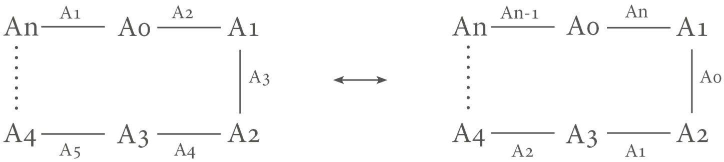

Fig. 2. The combinatorial structure of rule $G5$, for $n \geq 2$.

includes $n+1$ premise triplets and $n+1$ consequent triplets. When proceeding from $n+1$ to $n+2$ sets of variables therefore, both the number of premise triplets and the number of consequent triplets increase by one. Throughout this section, the combinatorial structures of the generalized rules are illustrated by schematic visualizations. The structure of rule $G5$ is shown in Fig. 2. Each labeled line $X \overset{Z}{\text{---}} Y$ in the figure depicts a triplet $\langle X, Y \mid Z \rangle$, with the line's label corresponding with the triplet's conditioning set. The left-hand side of the figure summarizes the joint premise of the rule and the right-hand side its joint consequent. The figure is readily seen to reflect the rule's sliding structure.

Before addressing our generalization of rule $A6$ from $\mathcal{S}$, we recall that Studený already generalized this rule from pertaining to four sets of variables, to an arbitrary number of variable sets; we recall this rule as Eq. (1) from Sect. 2.3. Where Eq. (1) included just a single variable set B_i in a triplet's conditioning part, our generalization has conditioning parts including multiple such sets per triplet. More specifically, the rule takes a sequence of variable sets $B_0, \ldots, B_n$, just like rule $G5$ above, and moves it sliding over a triplet, yet now over just its second and third argument positions, with the first argument fixed to an unrelated set A. The number of sets that are included in the conditioning parts of the rule's triplets, is governed by a parameter k.

Proposition 2. *Let $A, B_0, \ldots, B_n$, $n \geq 0$, be non-empty, mutually disjoint sets of variables. Then, for all $k \in [0, n]$,*

$$\bigwedge_{i \in \{0,\ldots,n\}} \langle A, B_i \mid \mathbf{B}^k_{\mathbf{i}+} \rangle \quad \leftrightarrow \quad \bigwedge_{i \in \{0,\ldots,n\}} \langle A, B_i \mid \mathbf{B}^k_{\mathbf{i}-} \rangle$$

where $\mathbf{B}^k_{\mathbf{i}+} = B_{\mu(i+1)} \cdots B_{\mu(i+k)}$, $\mathbf{B}^k_{\mathbf{i}-} = B_{\mu(i-k)} \cdots B_{\mu(i-1)}$, taking $\mathbf{B}^k_{\mathbf{i}+}, \mathbf{B}^k_{\mathbf{i}-} := \varnothing$ for $k = 0$, and where $\mu(x) := x \bmod (n+1)$ as before.

The property stated in the proposition is taken as the independence rule $G6$. Note that for $k = 0$, $k = n$, $n = 0$ and $n = 1$ the rule reduces to a tautology. $G6$ further embeds rule $A6$ as a special case with $n = 2$, $k = 1$, as is seen by setting $B_0 \leftarrow B$, $B_1 \leftarrow C$ and $B_2 \leftarrow D$. Equation (1) is embedded as the cases with $k = 1$ and any $n \geq 2$. As an example of the generalization, rule $G6$ is detailed for $n = 3$ and $k = 2$ below:

$$\langle A, B_0 \mid B_1 B_2 \rangle \wedge \langle A, B_1 \mid B_2 B_3 \rangle \wedge \langle A, B_2 \mid B_3 B_0 \rangle \wedge \langle A, B_3 \mid B_0 B_1 \rangle \;\leftrightarrow$$
$$\langle A, B_0 \mid B_2 B_3 \rangle \wedge \langle A, B_1 \mid B_3 B_0 \rangle \wedge \langle A, B_2 \mid B_0 B_1 \rangle \wedge \langle A, B_3 \mid B_1 B_2 \rangle$$

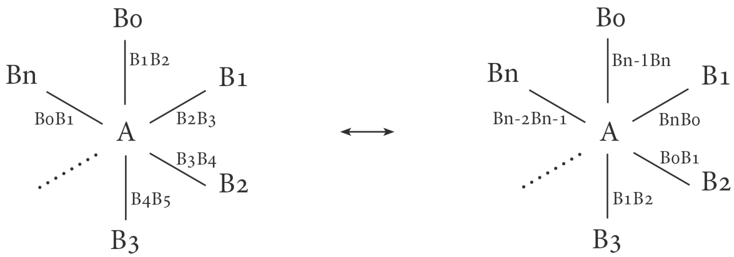

Fig. 3. The combinatorial structure of rule $G6$, for $n \geq 2$ and $k = 2$.

We observe that the rule's joint consequent cannot be derived from its joint premise using the original rules from the system $\mathcal{S}$. For $n+2$ sets of variables, rule $G6$ includes $n+1$ premise triplets and $n+1$ consequent triplets. When proceeding from $n+2$ to $n+3$ variable sets therefore, both the number of premise triplets and the number of consequent triplets increase by one; the parameter k does not affect the numbers of triplets involved. The combinatorial structure of rule $G6$ is shown in Fig. 3. The figure highlights the rule's star-shaped structure that originates from the fixed set A as the first argument of all triplets involved.

The idea underlying our generalization of rule $A4$ from $\mathcal{S}$, is somewhat less intuitive. Informally phrased, the generalized rule takes two fixed sets of variables A, B, and associates these sets with an ordered sequence $\mathbf{C}$ of sets C_i. Each C_i occurs as the second argument in two triplets, one with A for its first argument and one with B for its first argument. The two subsequences that remain after removing C_i from $\mathbf{C}$ are each positioned in the conditioning part of one of these triplets. In the case of an odd index i, moreover, the triplet involving A has the conditioning part extended with the set B, and vice versa. In the case of an even index, the conditioning parts of the triplets are not extended.

Proposition 3. *Let $A, B, C_1, \ldots, C_n$, with $n \geq 2$ even, be non-empty, mutually disjoint sets of variables. Then,*

$$\bigwedge_{i\in\{1,3,\ldots,n-1\}} [\,\langle A, C_i \mid \mathbf{C_{i^-}}\, B\,\rangle \;\wedge\; \langle B, C_i \mid \mathbf{C_{i^+}}\, A\,\rangle\,] \;\wedge\; \bigwedge_{i\in\{2,4,\ldots,n\}} [\,\langle A, C_i \mid \mathbf{C_{i^-}}\rangle \;\wedge\; \langle B, C_i \mid \mathbf{C_{i^+}}\rangle\,] \;\;\leftrightarrow$$

$$\bigwedge_{i\in\{1,3,\ldots,n-1\}} [\,\langle A, C_i \mid \mathbf{C_{i^+}}\, B\,\rangle \;\wedge\; \langle B, C_i \mid \mathbf{C_{i^-}}\, A\,\rangle\,] \;\wedge\; \bigwedge_{i\in\{2,4,\ldots,n\}} [\,\langle A, C_i \mid \mathbf{C_{i^+}}\rangle \;\wedge\; \langle B, C_i \mid \mathbf{C_{i^-}}\rangle\,]$$

where $\mathbf{C_{i^-}} = C_1 \cdots C_{i-1}$ and $\mathbf{C_{i^+}} = C_{i+1} \cdots C_n$, taking $C_j \cdots C_{j-1} := \varnothing$.

The property stated in the proposition is taken as the independence rule $G4$. Note that $G4$ embeds rule $A4$ as a special case with $n = 2$, as is seen by setting

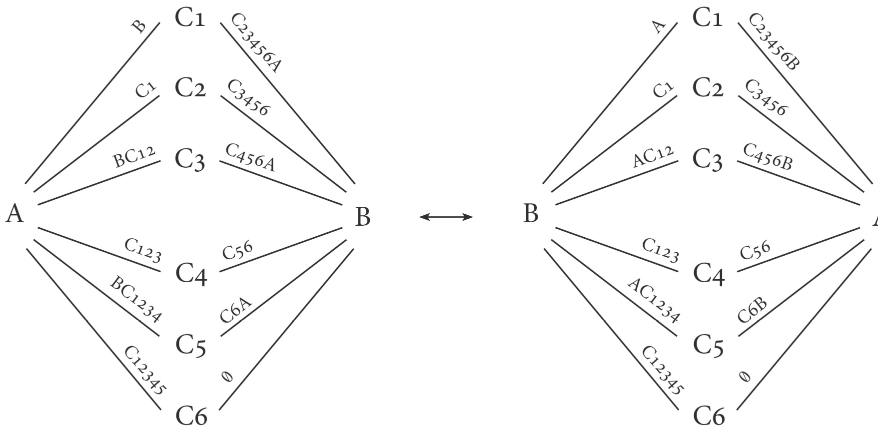

Fig. 4. The combinatorial structure of rule $G4$, for $n = 6$.

$A \leftarrow D$, $B \leftarrow B$, $C_1 \leftarrow A$ and $C_2 \leftarrow C$, where each set from $G4$ is assigned a set from rule $A4$. As an example of the generalization, rule $G4$ is now detailed for $n = 4$:

$$\begin{aligned}
&\langle A, C_1 \mid B\rangle \ \wedge\ \langle A, C_3 \mid C_1C_2B\rangle \ \wedge\ \langle B, C_1 \mid C_2C_3C_4A\rangle \ \wedge\ \langle B, C_3 \mid C_4A\rangle \ \wedge \\
&\langle A, C_2 \mid C_1\rangle \ \wedge\ \langle A, C_4 \mid C_1C_2C_3\rangle \ \wedge\ \langle B, C_2 \mid C_3C_4\rangle \ \wedge\ \langle B, C_4 \mid \varnothing\rangle \quad \leftrightarrow \\
&\langle A, C_1 \mid C_2C_3C_4B\rangle \ \wedge\ \langle A, C_3 \mid C_4B\rangle \ \wedge\ \langle B, C_1 \mid A\rangle \ \wedge\ \langle B, C_3 \mid C_1C_2A\rangle \ \wedge \\
&\langle A, C_2 \mid C_3C_4\rangle \ \wedge\ \langle A, C_4 \mid \varnothing\rangle \ \wedge\ \langle B, C_2 \mid C_1\rangle \ \wedge\ \langle B, C_4 \mid C_1C_2C_3\rangle
\end{aligned}$$

Note that the rule's consequent is not derivable from its premise using the original rules in $\mathcal{S}$. For $n+2$ sets of variables, rule $G4$ includes $2 \cdot n$ premise triplets and the same number of consequent triplets. When proceeding from $n + 2$ to $(n + 2) + 2$ variable sets therefore, both the number of premise triplets and the number of consequent triplets increase by four. The rule's combinatorial structure is shown in Fig. 4, for $n = 6$. The rule has a bipartitely linked structure, originating from the fixed sets A and B in the first argument positions per triplet linking to the exact same sets from the sequence $\mathbf{C}$. We further note that the lines from A and from B to the same set C_i swap labels (replacing A by B and vice versa) between the premise and consequent parts of the rule.

As the above generalization, our generalization of rule $A3$ takes two fixed sets of variables A, B. Instead of being associated with a single sequence of sets, A and B are now associated with two separate sequences $\mathbf{C}$ and $\mathbf{D}$, respectively. Each set C_i occurs as the second argument in two triplets, both with A as the first argument. The two remaining subsequences of $\mathbf{C}$ after removing C_i, are each positioned in the conditioning part of one of these triplets, supplemented with a subsequence of $\mathbf{D}$. A similar pattern is seen with the set B and its associated sequence $\mathbf{D}$. The triplet pairs with A and those with B as their first arguments,

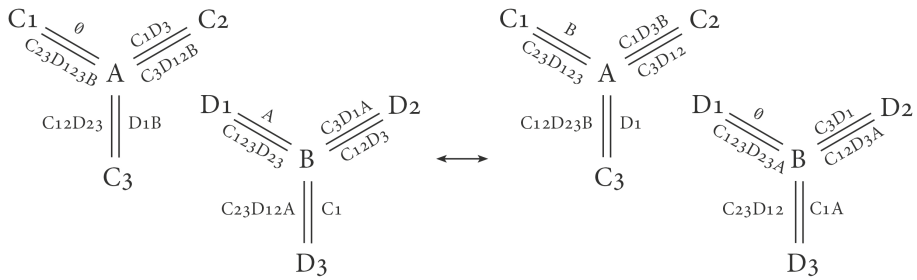

Fig. 5. The combinatorial structure of rule $G3$, for $n = 3$.

are again related through the inclusion of B in the conditioning part of one of the triplet pairs with A, and vice versa.

Proposition 4. *Let A, B, C_i, D_i, $i = 1, \ldots, n$, $n \geq 1$, be non-empty, mutually disjoint sets of variables, and let $\mathbf{C} = C_1 \cdots C_n$ and $\mathbf{D} = D_1 \cdots D_n$. Then,*

$$\bigwedge_{i\in\{1,\ldots,n\}} \begin{array}{l} [\, \langle A, C_i \mid \mathbf{C_{i^-}}\, \mathbf{D_{i^{++}}} \rangle \;\wedge\; \langle A, C_i \mid \mathbf{C_{i^+}}\, \mathbf{D_{i^{--}}}\, B \rangle \;\wedge \\ \langle B, D_i \mid \mathbf{C_{i^{++}}}\, \mathbf{D_{i^-}}\, A \rangle \;\wedge\; \langle B, D_i \mid \mathbf{C_{i^{--}}}\, \mathbf{D_{i^+}} \rangle \,] \end{array} \quad \leftrightarrow$$

$$\bigwedge_{i\in\{1,\ldots,n\}} \begin{array}{l} [\, \langle A, C_i \mid \mathbf{C_{i^-}}\, \mathbf{D_{i^{++}}}\, B \rangle \;\wedge\; \langle A, C_i \mid \mathbf{C_{i^+}}\, \mathbf{D_{i^{--}}} \rangle \;\wedge \\ \langle B, D_i \mid \mathbf{C_{i^{++}}}\, \mathbf{D_{i^-}} \rangle \;\wedge\; \langle B, D_i \mid \mathbf{C_{i^{--}}}\, \mathbf{D_{i^+}}\, A \rangle \,] \end{array}$$

where $\mathbf{Z_{i^-}} = Z_1 \cdots Z_{i-1}$, $\mathbf{Z_{i^+}} = Z_{i+1} \cdots Z_n$, $\mathbf{Z_{i^{--}}} = Z_1 \cdots Z_{n-i+1}$ and $\mathbf{Z_{i^{++}}} = Z_{n-i+2} \cdots Z_n$, taking $Z_j \cdots Z_{j-1} := \varnothing$, for $\mathbf{Z} = \mathbf{C}, \mathbf{D}$.

The property stated in the proposition is taken as the independence rule $G3$. Note that $G3$ embeds rule $A3$ as a special case with $n = 1$, as is seen by setting $A \leftarrow A$, $B \leftarrow C$, $C_1 \leftarrow B$ and $D_1 \leftarrow D$, where each set from $G3$ is assigned a set from rule $A3$. As an example of the generalization, rule $G3$ is now detailed for $n = 2$:

$$\begin{array}{l} \langle A, C_1 \mid \varnothing \rangle \;\wedge\; \langle A, C_1 \mid C_2 D_1 D_2 B \rangle \;\wedge\; \langle B, D_1 \mid A \rangle \;\wedge\; \langle B, D_1 \mid C_1 C_2 D_2 \rangle \;\wedge \\ \langle A, C_2 \mid C_1 D_2 \rangle \;\wedge\; \langle A, C_2 \mid D_1 B \rangle \;\wedge\; \langle B, D_2 \mid C_2 D_1 A \rangle \;\wedge\; \langle B, D_2 \mid C_1 \rangle \qquad \leftrightarrow \\ \langle A, C_1 \mid B \rangle \;\wedge\; \langle A, C_1 \mid C_2 D_1 D_2 \rangle \;\wedge\; \langle B, D_1 \mid \varnothing \rangle \;\wedge\; \langle B, D_1 \mid C_1 C_2 D_2 A \rangle \;\wedge \\ \langle A, C_2 \mid C_1 D_2 B \rangle \;\wedge\; \langle A, C_2 \mid D_1 \rangle \;\wedge\; \langle B, D_2 \mid C_2 D_1 \rangle \;\wedge\; \langle B, D_2 \mid C_1 A \rangle \end{array}$$

We note that it is not possible to derive the rule's consequent from its premise by means of the rule system $\mathcal{S}$. For $2 \cdot n + 2$ sets of variables, rule $G3$ includes $4 \cdot n$ premise triplets and the same number of consequent triplets. When proceeding from $2 \cdot n + 2$ to $2 \cdot (n + 1) + 2$ variable sets therefore, both the number of premise triplets and the number of consequent triplets increase by four. The combinatorial structure of the rule is illustrated in Fig. 5, for $n = 3$. This structure consists

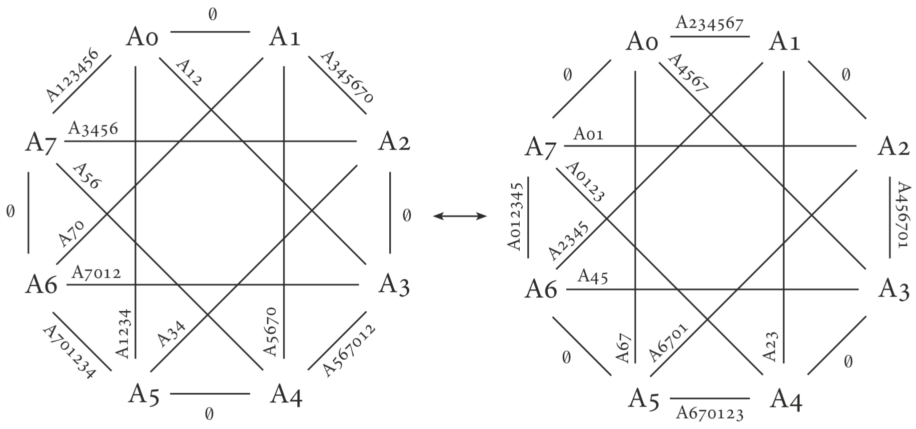

Fig. 6. The combinatorial structure of rule $G7$, for $n = 7$.

of double-edged stars with the cardinalities of the conditioning sets of the paired triplets adding up to $2 \cdot n$.

To conclude this section, we present our generalization of rule $A7$ from $\mathcal{S}$. The generalized rule involves a fixed sequence with an even number of sets of variables A_i. Each set with an even index i is paired with each set with an odd index i in the first two argument positions of a triplet. The conditioning part of such a triplet includes either the intervening variables sets of the sequence or all remaining variable sets.

Proposition 5. *Let $A_0, \ldots, A_n$, with $n \geq 1$ an odd number, be non-empty, mutually disjoint sets of variables. Then,*

$$\bigwedge_{\substack{i \in \{0,2,\ldots,n-1\},\\ k \in \{1,3,\ldots,n\}}} \langle A_i, A_{\mu(i+k)} \mid \mathbf{A_{i^+}} \rangle \quad \leftrightarrow \quad \bigwedge_{\substack{i \in \{0,2,\ldots,n-1\},\\ k \in \{1,3,\ldots,n\}}} \langle A_i, A_{\mu(i-k)} \mid \mathbf{A_{i^-}} \rangle$$

where $\mathbf{A_{i^+}} = A_{\mu(i+1)} \cdots A_{\mu(i+k-1)}$ and $\mathbf{A_{i^-}} = A_{\mu(i-k+1)} \cdots A_{\mu(i-1)}$, taking $\mathbf{A_{i^+}}, \mathbf{A_{i^-}} := \varnothing$ for $k = 1$, and where $\mu(x) := x \bmod (n+1)$, as before.

The property stated in the proposition is taken as the independence rule $G7$. Note that for $n = 1$ the rule reduces to a tautology. $G7$ further embeds rule $A7$ as a special case with $n = 3$, as is seen by setting $A_0 \leftarrow A$, $A_1 \leftarrow C$, $A_2 \leftarrow D$ and $A_3 \leftarrow B$. As an example of the generalization, rule $G7$ is now detailed for $n = 5$:

$$\begin{aligned}
&\langle A_0, A_1 \mid \varnothing\rangle \;\wedge\; \langle A_0, A_3 \mid A_1A_2\rangle \;\wedge\; \langle A_0, A_5 \mid A_1A_2A_3A_4\rangle \;\wedge\\
&\langle A_2, A_3 \mid \varnothing\rangle \;\wedge\; \langle A_2, A_5 \mid A_3A_4\rangle \;\wedge\; \langle A_2, A_1 \mid A_3A_4A_5A_0\rangle \;\wedge\\
&\langle A_4, A_5 \mid \varnothing\rangle \;\wedge\; \langle A_4, A_1 \mid A_5A_0\rangle \;\wedge\; \langle A_4, A_3 \mid A_5A_0A_1A_2\rangle \quad\leftrightarrow\\
&\langle A_0, A_5 \mid \varnothing\rangle \;\wedge\; \langle A_0, A_3 \mid A_4A_5\rangle \;\wedge\; \langle A_0, A_1 \mid A_2A_3A_4A_5\rangle \;\wedge\\
&\langle A_2, A_1 \mid \varnothing\rangle \;\wedge\; \langle A_2, A_5 \mid A_0A_1\rangle \;\wedge\; \langle A_2, A_3 \mid A_4A_5A_0A_1\rangle \;\wedge\\
&\langle A_4, A_3 \mid \varnothing\rangle \;\wedge\; \langle A_4, A_1 \mid A_2A_3\rangle \;\wedge\; \langle A_4, A_5 \mid A_0A_1A_2A_3\rangle
\end{aligned}$$

We observe that it is not possible to derive the rule's consequent from its premise using the rule system $\mathcal{S}$ considered thus far. For $n+1$ sets of variables, rule $G7$ includes $((n+1)/2)^2$ premise triplets and the same number of consequent triplets. When proceeding from $n+1$ to $n+3$ sets of variables therefore, both the number of premise triplets and the number of consequent triplets are increased by $n+2$. The combinatorial structure of the rule is illustrated in Fig. 6 for $n = 7$. The conditioning sets of a premise triplet are found on the outermost circle, in between the sets of its first two arguments, going clockwise starting from its first argument. For a consequent triplet a similar observation holds, now however the conditioning sets are found going counterclockwise.

We now define $\mathcal{G} = \mathcal{A} \cup \{G3, \ldots, G7\}$ for the new system of independence rules. Compared to the system $\mathcal{S}$, system $\mathcal{G}$ further restricts the number of relations closed under all known independence rules and thereby constitutes an enhanced characterisation of the class of structural semi-graphoid relations. Figure 1(b) positions the new system $\mathcal{G}$ into the hierarchy of independence relations.

4 Conclusions and Future Research

With existing rule systems for probabilistic independence being sound yet incomplete, we presented generalizations of the five rules for independence formulated by Studený [6], to larger numbers of (sets of) variables. With these generalized rules we uncovered combinatorial structures of probabilistic independence, and thereby enable their further study. We also closed in on the class of probabilistic independence relations, as the class of relations closed under the generalized rules is a proper subclass of the class closed under the previously existing rule system. By further exploiting multiinformation proof constructs, we hope to uncover in future research further combinatorial structures of independence and to formulate new generalized rules for probabilistic independence.

Acknowledgment. We would like to thank Milan Studený for his quick answers to our questions related to his work.

Appendix: Sketches of Proofs

All propositions stated in Sect. 3 are proven through multiinformation proof constructs. We find that the multiinformation equations of the four rules $G3$–$G6$ share the property that each term occurs exactly twice. In these equations, moreover, any multiinformation term originating from a premise triplet has a matching term, that occurs either in the rule's consequent with the same sign or in another premise triplet with the opposite sign. All terms therefore cancel out of the equation. For the fifth rule $G7$ the same observation holds, except for the terms $M(\varnothing)$ and $M(A_0 \cdots A_n)$. These terms may occur more than twice, however both terms occur equally many times in the multiinformation expressions of the rule's premise and of the rule's consequent and thus cancel out as well. In the proof sketches per rule, we restrict ourselves to indicating how each multiinformation term from an arbitrarily chosen premise triplet is canceled out from the multiinformation equation. Upon doing so, we abbreviate the multiinformation terms $M(XYZ)$, $M(Z)$, $M(XZ)$ and $M(YZ)$ of a triplet $\theta = \langle X, Y \mid Z \rangle$ as $M_{\mathrm{I}}\theta$, $M_{\mathrm{II}}\theta$, $M_{\mathrm{III}}\theta$ and $M_{\mathrm{IV}}\theta$, respectively.

Proposition 1. Let $\theta = \langle A_j, A_{\mu(j+1)} \mid A_{\mu(j+2)} \rangle$ be a premise triplet of rule $G5$, for some index $i = j$. Then, the triplet $\theta' = \langle A_{\mu(j+1)}, A_{\mu(j+2)} \mid A_j \rangle$, with $i = \mu(j+1)$, in the rule's consequent, has $M_{\mathrm{I}}\theta' = M_{\mathrm{I}}\theta$ and $M_{\mathrm{IV}}\theta' = M_{\mathrm{III}}\theta$. The consequent triplet $\theta'' = \langle A_{\mu(j+3)}, A_{\mu(j+4)} \mid A_{\mu(j+2)} \rangle$, with $i = \mu(j+3)$, further has $M_{\mathrm{II}}\theta'' = M_{\mathrm{II}}\theta$, and the consequent triplet $\theta''' = \langle A_{\mu(j+2)}, A_{\mu(j+3)} \mid A_{\mu(j+1)} \rangle$, with $i = \mu(j+2)$, to conclude, has $M_{\mathrm{III}}\theta''' = M_{\mathrm{IV}}\theta$. □

Proposition 2. Let $\theta = \langle A, B_j \mid \mathbf{B}^k_{\mathbf{j}+} \rangle$ be a premise triplet of rule $G6$, for some index $i = j$ and some k. Then, the consequent triplet $\theta' = \langle A, B_{\mu(j+k)} \mid \mathbf{B}^k_{\mu(\mathbf{j}+\mathbf{k})-} \rangle$, with $i = \mu(j+k)$, has $M_{\mathrm{I}}\theta' = M_{\mathrm{I}}\theta$ and $M_{\mathrm{IV}}\theta' = M_{\mathrm{IV}}\theta$. The consequent triplet $\theta'' = \langle A, B_{\mu(j+k+1)} \mid \mathbf{B}^k_{\mu(\mathbf{j}+\mathbf{k}+\mathbf{1})-} \rangle$, with $i = \mu(j+k+1)$, further has $M_{\mathrm{II}}\theta'' = M_{\mathrm{II}}\theta$ and $M_{\mathrm{III}}\theta'' = M_{\mathrm{III}}\theta$. □

Proposition 3. Let $\theta = \langle A, C_j \mid C_1 \cdots C_{j-1}\, B \rangle$ be a premise triplet of rule $G4$, for some index $i = j$. Then, the consequent triplet $\theta' = \langle B, C_j \mid C_1 \cdots C_{j-1}\, A \rangle$, with $i = j$, has $M_{\mathrm{I}}\theta' = M_{\mathrm{I}}\theta$ and $M_{\mathrm{III}}\theta' = M_{\mathrm{III}}\theta$. The consequent triplet $\theta'' = \langle B, C_{j+1} \mid C_1 \ldots C_j \rangle$, with $i = j+1$, has $M_{\mathrm{III}}\theta'' = M_{\mathrm{IV}}\theta$. For $j > 1$, the consequent triplet $\theta''' = \langle B, C_{j-1} \mid C_1 \ldots C_{j-2} \rangle$, with $i = j-1$, has $M_{\mathrm{I}}\theta''' = M_{\mathrm{II}}\theta$. For $j = 1$, the term $M_{\mathrm{II}}\theta$ is canceled out of the multiinformation equation by the term M_{III}, with the opposite sign, of the premise triplet $\langle B, C_n \mid \varnothing \rangle$. Similar arguments apply to the premise triplets of another form than θ. □

Proposition 4. Let $\theta = \langle A, C_j \mid C_1 \cdots C_{j-1} D_{n-j+2} \cdots D_n \rangle$ be a premise triplet of $G3$, for some index $i = j$. Then, the consequent triplet $\theta' = \langle B, D_{n-j+1} \mid C_1 \cdots C_j D_{n-j+2} \cdots D_n A \rangle$, with $i = n-j+1$, has $M_{\mathrm{II}}\theta' = M_{\mathrm{I}}\theta$. The premise triplet $\theta'' = \langle B, D_{n-j+1} \mid C_1 \cdots C_j D_{n-j+2} \cdots D_n \rangle$, with $i = n-j+1$, further has $M_{\mathrm{II}}\theta'' = -M_{\mathrm{IV}}\theta$. For all $j > 1$, the premise triplet $\theta''' = \langle B, D_{n-j+2} \mid C_1 \cdots C_{j-1} D_{n-j+3} \cdots D_n \rangle$, with $i = n-j+2$, has $M_{\mathrm{IV}}\theta''' = -M_{\mathrm{II}}\theta$ and the consequent triplet $\theta'''' = \langle B, D_{n-j+2} \mid C_1 \cdots C_{j-1} D_{n-j+3} \cdots D_n A \rangle$, also with $i = n-j+2$, has $M_{\mathrm{IV}}\theta'''' = M_{\mathrm{III}}\theta$. For $j = 1$ to conclude, the term $M_{\mathrm{II}}\theta$ is matched by the M_{II} term of the consequent triplet $\langle B, D_1 \mid \varnothing \rangle$ and the term $M_{\mathrm{III}}\theta$ is canceled out by the term M_{II} of opposite sign of the premise triplet $\langle B, D_1 \mid A \rangle$, with $i = 1$. Similar arguments apply to premise triplets of other form than θ. □

Proposition 5. Let $\theta = \langle A_j, A_{\mu(j+h)} \mid A_{\mu(j+1)} \cdots A_{\mu(j+h-1)} \rangle$ be a premise triplet of $G7$, for some indices $i = j$, $k = h$. The term $M_{\mathrm{III}}\theta$ is canceled out from the multiinformation equation by the term $M_{\mathrm{III}}\theta'$ of the consequent triplet $\theta' = \langle A_{\mu(j+h-1)}, A_{\mu(j-1)} \mid A_j \cdots A_{\mu(j+h-2)} \rangle$, with $k = h$, $i = \mu(j+h-1)$. The term $M_{\mathrm{IV}}\theta$ is canceled out by $M_{\mathrm{IV}}\theta''$ of the consequent triplet $\theta'' = \langle A_{\mu(j+h+1)}, A_{\mu(j+1)} \mid A_{\mu(j+2)} \cdots A_{\mu(j+h)} \rangle$, with $k = h$, $i = \mu(j+h+1)$. For all $k \leq n-2$, the term $M_{\mathrm{I}}\theta$ is canceled out by the term $M_{\mathrm{II}}\theta'''$ of the consequent triplet $\theta''' = \langle A_{\mu(j+h+1)}, A_{\mu(j-1)} \mid A_{\mu(j)} \cdots A_{\mu(j+h)} \rangle$, with $k = h+2$, $i = \mu(j+h+1)$; for $k = n$, the M_{I} terms of all premise and consequent triplets have $M_{\mathrm{I}} \cdot = M(A_0 \cdots A_n)$ and hence match. For all $k \geq 3$, the term $M_{\mathrm{II}}\theta$ is canceled out by the term $M_{\mathrm{I}}\theta''''$ of the consequent triplet $\theta'''' = \langle A_{\mu(j+h-1)}, A_{\mu(j+1)} \mid A_{\mu(j+2)} \cdots A_{\mu(j+h-2)} \rangle$, with $k = h-2$, $i = \mu(j+h-1)$; for $k = 1$, the M_{II} terms of all premise and consequent triplets have $M_{\mathrm{II}} = M(\varnothing)$ and thus match. □

References

1. Dawid, A.: Conditional independence in statistical theory. J. R. Stat. Soc. B **41**, 1–31 (1979)
2. Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann (1988)
3. Studený, M.: Attempts at axiomatic description of conditional independence. Kybernetika **25**, 72–79 (1989)
4. Studený, M.: Multiinformation and the problem of characterization of conditional independence relations. Probl. Control Inf. Theor. **18**, 01 (1989)
5. Studený, M.: Conditional independence relations have no finite complete characterization. In: Information Theory, Statistical Decision Functions and Random Processes. Transactions of the 11th Prague Conference, vol. B, pp. 377–396. Kluwer, Dordrecht (1992)
6. Studený, M.: Structural semigraphoids. Int. J. Gen. Syst. **22**, 207–217 (1994)
7. Studený, M.: Description of structures of stochastic conditional independence by means of faces and imsets. 1st part: introduction and basic concepts. Int. J. Gen. Syst. **23**, 123–137 (1995)
8. Studený, M.: Description of structures of stochastic conditional independence by means of faces and imsets. 2rd part: basic theory. Int. J. Gen. Syst. **23**, 201–219 (1995)
9. Studený, M.: Description of structures of stochastic conditional independence by means of faces and imsets. 3rd part: examples of use and appendices. Int. J. Gen Syst **23**, 323–341 (1995)
10. Studený, M.: Probabilistic Conditional Independence Structures. Springer, London (2005). https://doi.org/10.1007/b138557
11. Studený, M., Vejnarová, J.: The multiinformation function as a tool for measuring stochastic dependence. In: Learning in Graphical Models, pp. 261–298. Kluwer Academic Publishers (1998)
12. Yeung, R.: A First Course in Information Theory. Kluwer Academic Publishers, Dordrecht (2002)