On Minimum Elementary-Triplet Bases for Independence Relations

Janneke H. Bolt   J.H.Bolt@uu.nl
Linda C. van der Gaag   L.C.vanderGaag@uu.nl
Department of Information and Computing Sciences, Utrecht University, The Netherlands

Abstract

A semi-graphoid independence relation is a set of independence statements, called triplets, and is typically exponentially large in the number of variables involved. For concise representation of such a relation, a subset of its triplets is listed in a so-called basis; its other triplets are defined implicitly through a set of axioms. An elementary-triplet basis for this purpose consists of all elementary triplets of a relation. Such a basis however, may include redundant information. In this paper we provide two lower bounds on the size of an elementary-triplet basis in general and an upper bound on the size of a minimum elementary-triplet basis. We further specify the construction of an elementary-triplet basis of minimum size for restricted relations.

Keywords: Independence relations, Axioms of independence, Elementary triplets, Basis representation.

1. Introduction

The notion of conditional independence plays a key role in practical systems of uncertainty, since effective use of knowledge about independences allows these systems to deal with the computational complexity of their problem-solving tasks. The notion thus arises in various frameworks of uncertainty, and in probabilistic settings more specifically (see for example [3, 4, 8, 9, 24, 25]). Probabilistic conditional independence has been subject to multiple studies, from both a mathematics perspective and a computing-science perspective, with the latter focusing on problems of concise representation and efficient computation.

An *independence relation* over a set of random variables $\mathbf{V}$ is a set of *triplets* $\langle \mathbf{A}, \mathbf{B} \mid \mathbf{C} \rangle$ where $\mathbf{A}, \mathbf{B}, \mathbf{C} \subseteq \mathbf{V}$ are pairwise disjoint subsets with $\mathbf{A}, \mathbf{B} \neq \varnothing$. A triplet $\langle \mathbf{A}, \mathbf{B} \mid \mathbf{C} \rangle$ states that the sets of variables $\mathbf{A}$ and $\mathbf{B}$ are independent given the conditioning set $\mathbf{C}$; relative to any specific discrete joint probability distribution $\Pr$ over $\mathbf{V}$, the triplet thus states that $\Pr(\mathbf{A}, \mathbf{B} \mid \mathbf{C}) = \Pr(\mathbf{A} \mid \mathbf{C}) \cdot \Pr(\mathbf{B} \mid \mathbf{C})$ for all possible value combinations of $\mathbf{A}, \mathbf{B}, \mathbf{C}$. Well-known properties of classical probabilistic independence have been formulated as axiomatic systems to allow a study of independence without the numerical context involved (see for example [8, 10, 11, 12, 15, 17, 18, 20, 22]). The most often studied system includes four axioms, called the *semi-graphoid axioms*. Any (ternary) relation over $\mathbf{V}$ that is closed under these axioms, is then called a *semi-graphoid independence relation* [12, 18]. Although the semi-graphoid axioms have been formulated within the setting of classical probability theory, their validity and implications are also studied within other frameworks of uncertainty (see for example [5, 6, 7, 16, 26]).

Semi-graphoid independence relations in general are exponentially large in the number of random variables involved [22, 23], and representing them by mere enumeration of their triplets is not feasible for practical purposes. A more concise representation is arrived at by explicitly listing a small set of triplets, called a *basis*, and letting all other triplets be defined implicitly through the semi-graphoid axioms [23]. For such a representation, two types of basis have been proposed in the literature: an *elementary-triplet* basis is composed of triplets $\langle \mathbf{A}, \mathbf{B} \mid \mathbf{C} \rangle$ with $\mathbf{A}, \mathbf{B}$ singleton sets [19], and a *dominant-triplet* basis is composed of triplets such that any remaining triplet can be derived directly from *one* triplet from this set [23]. The latter type of basis has received considerably more attention from the research community, since it is commonly thought to be smaller in size than an elementary-triplet basis. While appropriate algorithms have been designed for constructing dominant-triplet bases from a given starting triplet set without the need to generate the full relation [1, 2, 13, 23], no such algorithms are available as yet for the construction of non-redundant elementary-triplet bases.

In this paper, we study elementary-triplet bases as representations of a semi-graphoid independence relation defined by a starting set of general, possibly non-elementary triplets. More specifically, we address the problem of constructing minimum elementary-triplet bases for independence relations defined by a starting set J of general triplets, without the need of actually generating all triplets of the relation. We will thereby first focus on the problem for starting sets including a single triplet and show that an elementary-triplet basis of minimum size can always be readily constructed for such sets. We then use this result to derive an upper bound on the size of a minimum elementary-triplet basis for arbitrarily-sized starting sets.

The paper is organised as follows. In Section 2, we provide some preliminaries on independence relations in general and thereby introduce our notations. In Section 3 we provide two lower bounds for elementary-triplet bases and in Section 4 we show that, for a singleton starting set, a

basis of lower-bound size can always be constructed. In Section 5 we give an upper bound for minimum elementary-triplet bases for arbitrarily-sized starting sets. Section 6 concludes the paper with our future plans.

2. Preliminaries

We consider a finite, non-empty set $\mathbf{V}$ of discrete random variables. A *triplet* over $\mathbf{V}$ is a statement of the form $\langle \mathbf{A}, \mathbf{B} \,|\, \mathbf{C} \rangle$, where $\mathbf{A}, \mathbf{B}, \mathbf{C} \subseteq \mathbf{V}$ are pairwise disjoint subsets with $\mathbf{A}, \mathbf{B} \neq \varnothing$. A triplet $\langle \mathbf{A}, \mathbf{B} \,|\, \mathbf{C} \rangle$ states that the sets of variables $\mathbf{A}$ and $\mathbf{B}$ are independent given the conditioning set $\mathbf{C}$. The set of all triplets over $\mathbf{V}$ will be denoted by $\mathbf{V}^{(3)}$.

A subset of $\mathbf{V}^{(3)}$ constitutes a *semi-graphoid independence relation* if it satisfies the four properties below.

Definition 1 *A* semi-graphoid independence relation *is a subset of triplets $\overline{J} \subseteq \mathbf{V}^{(3)}$ that satisfies the following properties for all sets $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D} \subseteq \mathbf{V}$:*

G1: *if $\langle \mathbf{A}, \mathbf{B} \,|\, \mathbf{C} \rangle \in \overline{J}$, then $\langle \mathbf{B}, \mathbf{A} \,|\, \mathbf{C} \rangle \in \overline{J}$* (Symmetry)

G2: *if $\langle \mathbf{A}, \mathbf{B} \,|\, \mathbf{C} \rangle \in \overline{J}$, then $\langle \mathbf{A}, \mathbf{B}' \,|\, \mathbf{C} \rangle \in \overline{J}$ for any non-empty subset $\mathbf{B}' \subseteq \mathbf{B}$* (Decomposition)

G3: *if $\langle \mathbf{A}, \mathbf{B}_1 \cup \mathbf{B}_2 \,|\, \mathbf{C} \rangle \in \overline{J}$ with $\mathbf{B}_1 \cap \mathbf{B}_2 = \varnothing$, then $\langle \mathbf{A}, \mathbf{B}_1 \,|\, \mathbf{C} \cup \mathbf{B}_2 \rangle \in \overline{J}$* (Weak Union)

G4: *if $\langle \mathbf{A}, \mathbf{B} \,|\, \mathbf{C} \cup \mathbf{D} \rangle \in \overline{J}$ and $\langle \mathbf{A}, \mathbf{C} \,|\, \mathbf{D} \rangle \in \overline{J}$, then $\langle \mathbf{A}, \mathbf{B} \cup \mathbf{C} \,|\, \mathbf{D} \rangle \in \overline{J}$* (Contraction)

The four properties stated above have been proven logically independent and taken to constitute an axiomatic system for the qualitative notion of independency [18]. The system is sound relative to the class of discrete probability distributions, yet not complete [21].

The semi-graphoid axioms of independence are viewed as *derivation rules* for generating, possibly new, triplets from a given triplet set. Given an arbitrary triplet set $J \subseteq \mathbf{V}^{(3)}$ and a designated triplet $\theta \in \mathbf{V}^{(3)}$, we write $J \vdash^* \theta$ if the triplet θ can be derived from J by finite application of the rules G1–G4. The *closure* $\overline{J}$ of a starting triplet set J then is the semi-graphoid independence relation composed of J and all triplets θ that can be derived from it. The starting set J can thus be viewed as a concise representation of $\overline{J}$, and as such is called a *basis* for $\overline{J}$.

Definition 2 *A* basis *J for a semi-graphoid independence relation $\overline{J}$ is a subset $J \subseteq \overline{J}$ such that $\overline{J} = \{ \theta \mid J \vdash^* \theta \}$.*

Concise basis representations are of importance since independence relations in general are exponential in size. Two types of basis have been proposed. Studený [23] proposed as a basis for a relation $\overline{J}$, a subset of triplets $J \subseteq \overline{J}$ such that any remaining triplet can be derived directly from *one* triplet of J through the derivation rules G1–G3. Peña [19] proposed to represent $\overline{J}$ by its so-called elementary triplets.

An *elementary triplet* is a triplet of the form $\langle \{A\}, \{B\} \,|\, \mathbf{C} \rangle$, that is, a triplet with singleton sets for its first two arguments [19]; slightly abusing notation, we will omit in the sequel the set notation for singleton sets and write $\langle A, B \,|\, \mathbf{C} \rangle$. The set of all elementary triplets over $\mathbf{V}$ will be denoted by $\mathbf{E}^{(3)}$. For the subset $\overline{J}^{\mathbf{E}}$ of elementary triplets of an independence relation $\overline{J}$, the following properties hold.

Proposition 3 *Let $\overline{J}^{\mathbf{E}} = \overline{J} \cap \mathbf{E}^{(3)}$ be the set of elementary triplets of a semi-graphoid independence relation $\overline{J} \subseteq \mathbf{V}^{(3)}$. Then, $\overline{J}^{\mathbf{E}}$ satisfies the following properties for all $A, B, D \in \mathbf{V}$ and $\mathbf{C} \subseteq \mathbf{V}$:*

E1: *if $\langle A, B \,|\, \mathbf{C} \rangle \in \overline{J}^{\mathbf{E}}$, then $\langle B, A \,|\, \mathbf{C} \rangle \in \overline{J}^{\mathbf{E}}$* (Symmetry)

E2: *if $\langle A, B \,|\, \mathbf{C} \rangle \in \overline{J}^{\mathbf{E}}$ and $\langle A, D \,|\, B \cup \mathbf{C} \rangle \in \overline{J}^{\mathbf{E}}$, then $\langle A, D \,|\, \mathbf{C} \rangle \in \overline{J}^{\mathbf{E}}$ and $\langle A, B \,|\, \mathbf{C} \cup D \rangle \in \overline{J}^{\mathbf{E}}$* (Equivalence)

The two properties stated above in essence capture the semi-graphoid axioms G1–G4, tailored to elementary triplets. We use the term *equivalence* to refer to the property E2, to indicate that it essentially states an equivalence between two pairs of elementary triplets.

A well-known property associated with elementary triplets is that the set $\overline{J}^{\mathbf{E}}$ of all such triplets in an independence relation $\overline{J}$ constitutes a basis for $\overline{J}$ [14, 24]; $\overline{J}^{\mathbf{E}}$ was proposed in fact [19], as a representation of independence relations. The set of all elementary triplets in a relation may include redundant triplets due to the property E2 of equivalence however, and a smaller elementary-triplet basis may exist. In the sequel we will use K to denote an arbitrary elementary basis. Its closure under the semi-graphoid axioms will be indicated by $\overline{K}$ and its closure under E1–E2 by $\widetilde{K}$; $\widetilde{K}$ will be referred to as the *elementary closure* of K. We note that since E1–E2 in essence capture the semi-graphoid axioms G1–G4, we have that $\overline{K} = \overline{J}$ if and only if $\widetilde{K} = \overline{J}^{\mathbf{E}}$.

3. Lower Bounds on Elementary-Triplet Bases

We establish two lower bounds on the size of an elementary-triplet basis for a semi-graphoid independence relation. The first of these bounds is related to the number of different (unordered) variable pairs occurring in the first two arguments of any elementary triplet of the relation.

Proposition 4 *Let $\overline{J}$ be an independence relation and let $\overline{J}^{\mathbf{E}} = \overline{J} \cap \mathbf{E}^{(3)}$ be as before. Let U be the set of unordered variable pairs (A_i, B_j) with $\langle A_i, B_j \,|\, \mathbf{C} \rangle \in \overline{J}^{\mathbf{E}}$ for some $\mathbf{C} \subseteq \mathbf{V} \setminus \{A_i, B_j\}$. Then, any elementary-triplet basis for $\overline{J}$ includes at least $|U|$ elements.*

Proof The set U stated in the lemma includes all unordered variable pairs (A_i, B_j) occurring in the first two arguments of any elementary triplet of $\overline{J}^{\mathbf{E}}$. Now suppose that there exists an elementary-triplet basis K for $\overline{J}$ with $|K| < |U|$.

Then there must be at least one variable pair (A_r, B_s) such that $(A_r, B_s) \in U$ and $\langle A_r, B_s \mid \mathbf{C} \rangle, \langle B_s, A_r \mid \mathbf{C} \rangle \notin K$ for all $\mathbf{C} \subseteq \mathbf{V} \setminus \{A_r, B_s\}$. Since the derivation rules E1–E2 cannot construct a triplet with the variable pair (A_r, B_s) from the set K, it follows that the elementary closure $\widetilde{K}$ cannot be equal to $\overline{J}^{\mathbf{E}}$. Since $\overline{K} = \overline{J}$ if and only if $\widetilde{K} = \overline{J}^{\mathbf{E}}$, we conclude that K is not a basis for $\overline{J}$, which implies that any elementary-triplet basis for $\overline{J}$ should include at least $|U|$ elements. ∎

We note that, by the above proposition, an elementary-triplet basis may include $\mathcal{O}(n^2)$ triplets.

Another lower bound on the size of an elementary-triplet basis for an independence relation $\overline{J}$ is given by the number of different cardinalities of the conditioning sets $\mathbf{C}$ in $\overline{J}^{\mathbf{E}}$.

Proposition 5 *Let $\overline{J}, \overline{J}^{\mathbf{E}}$ be as before. Let U be the set of numbers t_k such that there is at least one triplet $\langle A_i, B_j \mid \mathbf{C} \rangle \in \overline{J}^{\mathbf{E}}$ with $|\mathbf{C}| = t_k$. Then, any elementary-triplet basis for $\overline{J}$ includes at least $|U|$ elements.*

Proof Suppose that there exists an elementary triplet basis K with $|K| < |U|$. Then there is at least one number t_r such that $t_r \in U$ and there is no triplet $\langle A_i, B_j \mid \mathbf{C} \rangle$ with $|\mathbf{C}| = t_r$ in K. Since the derivation rules E1–E2 cannot construct a triplet with a conditioning part of cardinality t_r from the triplet set K, it follows that $\widetilde{K}$ does not include such a triplet either and, hence, cannot be equal to $\overline{J}^{\mathbf{E}}$. Since $\overline{K} = \overline{J}$ if and only if $\widetilde{K} = \overline{J}^{\mathbf{E}}$, we conclude that K is not a basis for $\overline{J}$, which implies that any elementary-triplet basis for $\overline{J}$ should include at least $|U|$ elements. ∎

The two bounds stated in the Propositions 4 and 5 are tight in the sense that for each of these bounds independence relations exist for which it is attained. There are relations however, for which neither of the two bounds are met, as demonstrated by the following example.

Example 1 *We consider the following starting set*

$$J = \{ \langle 1, \{2,3\} \mid \varnothing \rangle, \langle 1, \{2,3\} \mid \{4,5,6\} \rangle, \langle 1, \{7,8\} \mid \varnothing \rangle \}$$

The set $\overline{J}^{\mathbf{E}}$ of the relation $\overline{J}$ includes:

$\langle 1,2 \mid \varnothing \rangle$	$\langle 1,2 \mid \{4,5,6\} \rangle$	$\langle 1,7 \mid \varnothing \rangle$
$\langle 1,2 \mid 3 \rangle$	$\langle 1,2 \mid \{3,4,5,6\} \rangle$	$\langle 1,7 \mid 8 \rangle$
$\langle 1,3 \mid \varnothing \rangle$	$\langle 1,3 \mid \{4,5,6\} \rangle$	$\langle 1,8 \mid \varnothing \rangle$
$\langle 1,3 \mid 2 \rangle$	$\langle 1,3 \mid \{2,4,5,6\} \rangle$	$\langle 1,8 \mid 7 \rangle$

The lower bounds found by Propositions 4 and 5 are both equal to 4. *It is easily verified through the derivation rules E1–E2 however, that the number of triplets in a minimum elementary-triplet basis for $\overline{J}$ equals* 6. □

4. Minimum Bases from Singleton Starting Sets

In this section we focus on elementary-triplet bases of minimum size for the closure of starting sets composed of a single, possibly non-elementary, triplet; Section 5 addresses elementary-triplet bases for arbitrarily-sized starting sets.

We begin by stating the number of elementary triplets implied by a general triplet θ through the semi-graphoid axioms G1–G4. In Lemma 7, we then show how to construct a generating triplet set from the elementary triplets implied by θ. Proposition 8 concludes by showing that the constructed triplet set is a basis for the closure of $\{\theta\}$.

Lemma 6 *Let $\theta = \langle \mathbf{A}, \mathbf{B} \mid \mathbf{C} \rangle$ be an arbitrary triplet in $\mathbf{V}^{(3)}$, with $\mathbf{A} = \{A_1, \ldots, A_n\}$, $n \geq 1$, and $\mathbf{B} = \{B_1, \ldots, B_m\}$, $m \geq 1$. Then, θ implies $n \cdot m \cdot 2^{n+m-2}$ different elementary triplets (ignoring symmetry).*

Proof From the triplet θ, by the *decomposition* and *weak union* rules, all elementary triplets of the form $\langle A_i, B_j \mid \mathbf{D} \rangle$ with $\mathbf{C} \subseteq \mathbf{D} \subseteq \mathbf{C} \cup (\mathbf{A} \cup \mathbf{B}) \setminus \{A_i, B_j\}$ are derived, for $i = 1, \ldots, n$, $j = 1, \ldots, m$. For a single pair A_i, B_j, the number of elementary triplets thus constructed equals the size of the power set of $(\mathbf{A} \cup \mathbf{B}) \setminus \{A_i, B_j\}$, that is, 2^{n+m-2}. With $n \cdot m$ pairs A_i, B_j, therefore, a total of $n \cdot m \cdot 2^{n+m-2}$ different elementary triplets are derived from θ. ∎

The set of elementary triplets derived from a single triplet θ as described in Lemma 6, is known to constitute a basis for the closure of $\{\theta\}$. As this set adheres to the axioms E1–E2 from Proposition 3, however, it may include derivable elements and not be minimal with respect to set inclusion.

The following lemma and associated proposition now detail the construction of an elementary-triplet basis for the closure of $\{\theta\}$ of a size that exactly matches the lower bound from Proposition 4 for any starting set composed of a single independence statement.

Lemma 7 *Let $\theta = \langle \mathbf{A}, \mathbf{B} \mid \mathbf{C} \rangle$ be a triplet in $\mathbf{V}^{(3)}$ with $\mathbf{A} = \{A_1, \ldots, A_n\}$, $n \geq 1$, and $\mathbf{B} = \{B_1, \ldots, B_m\}$, $m \geq 1$. Furthermore, let*

$$K_\theta = \big\{ \langle A_i, B_j \mid \mathbf{A} \setminus \{A_i, \ldots A_n\} \cup \mathbf{B} \setminus \{B_j, \ldots B_m\} \cup \mathbf{C} \rangle \mid \\ i = 1, \ldots, n, j = 1, \ldots, m \big\}$$

Then, $K_\theta \vdash^* \theta$.

Proof We consider the elementary triplet set K_θ and show by induction that $K_\theta \vdash^* \theta$.

Induction base: Let $\theta = \theta_{1,1} = \langle A_1, B_1 \mid \mathbf{C} \rangle$. Then, $K_{\theta_{1,1}} = \{\langle A_1, B_1 \mid \mathbf{C} \rangle\}$ clearly implies the triplet $\theta_{1,1}$.

Induction hypothesis: Let $\theta_{i,j} = \langle \{A_1, \ldots, A_i\}, \{B_1, \ldots B_j\} \mid \mathbf{C} \rangle$, for some $i \in \{1, \ldots, n\}, j \in \{1, \ldots, m\}$. Our hypothesis has that $K_{\theta_{k,\ell}} \vdash^* \theta_{k,\ell}$ for all $k = 1, \ldots, i, \ell = 1, \ldots, j$.

Induction step: Without loss of generality we consider

$$\theta_{i,j+1} = \langle \{A_1,\ldots,A_i\},\{B_1,\ldots,B_{j+1}\} \mid \mathbf{C} \rangle$$

that is, the triplet $\theta_{i,j+1}$ that extends $\theta_{i,j}$ by a single variable in the second argument. The triplet set $K_{\theta_{i,j+1}}$ equals

$$K_{\theta_{i,j}} \cup \{\langle A_k,B_{j+1} \mid \mathbf{A}_k^- \cup \{B_1,\ldots,B_j\} \cup \mathbf{C} \rangle \mid k=1,\ldots,i\}$$

where $\mathbf{A}_k^- = \varnothing$ for $k=1$ and $\mathbf{A}_k^- = \{A_1,\ldots,A_{k-1}\}$ otherwise. Since $K_{\theta_{i,j}} \vdash^* \theta_{i,j}$ by the induction hypothesis and $K_{\theta_{i,j}} \subseteq K_{\theta_{i,j+1}}$, we have that $K_{\theta_{i,j+1}} \vdash^* \theta_{i,j}$. From the triplets in $K_{\theta_{i,j+1}} \backslash K_{\theta_{i,j}}$, we can now construct the triplet

$$\theta' = \langle \{A_1,\ldots,A_i\},B_{j+1} \mid \{B_1,\ldots,B_j\} \cup \mathbf{C} \rangle$$

by repeatedly applying the contraction rule G4. From θ' and $\theta_{i,j}$, we then construct the triplet $\theta_{i,j+1}$ by once more applying G4. ∎

From Lemma 7 we have that the constructed elementary-triplet set allows the derivation of the starting triplet θ. The following proposition now states that the constructed set actually is a minimum elementary-triplet basis.

Proposition 8 *Let $\theta = \langle \mathbf{A},\mathbf{B} \mid \mathbf{C} \rangle$ be a triplet with $\mathbf{A} = \{A_1,\ldots,A_n\}$, $n \geq 1$, and $\mathbf{B} = \{B_1,\ldots,B_m\}$, $m \geq 1$. Let $J = \{\theta\}$ and let $\bar{J}$ be its closure. Furthermore, let K_θ be the elementary-triplet set constructed from J in Lemma 7. Then, K_θ is a minimum elementary-triplet basis for $\bar{J}$.*

Proof From Lemma 7, we have that the set K_θ implies the triplet θ. As K_θ is a subset of the closure $\bar{J}$ of $\{\theta\}$, we further have that any triplet τ with $K_\theta \vdash^* \tau$ is implied by θ. The set K_θ thus is an elementary-triplet basis for $\bar{J}$. As it includes $n \cdot m$ elements by its construction moreover, we have by Proposition 4 that it is a minimum basis for $\bar{J}$. ∎

We note that from a different perspective, Peña proved the related proposition stating that it suffices to check $n \cdot m$ elementary triplets from a basis $\bar{J}^{\mathbf{E}}$ to establish whether some triplet θ is included in $\bar{J}$ [19].

The elementary-triplet basis K_θ from Lemma 7 is not unique. It is readily seen that by choosing different orderings of the variables in the sets $\mathbf{A}$ and $\mathbf{B}$, a different basis will result. Yet, there exist also minimum elementary-triplet bases that cannot be obtained by different orders of processing $\mathbf{A}$ and $\mathbf{B}$, as shown in the following example.

Example 2 *Let $J = \{\langle\{1,2\},\{3,4\} \mid \varnothing\rangle\}$. By Lemma 7, the set*

$$\{\langle 1,3 \mid \varnothing\rangle,\langle 1,4 \mid 3\rangle,\langle 2,3 \mid 1\rangle,\langle 2,4 \mid \{1,3\}\rangle\}$$

is a minimum elementary-triplet basis for $\bar{J}$. The first and third triplet of this basis can be substituted by their equivalent triplet pair through the rule E2, resulting in

$$\{\langle 1,3 \mid 2\rangle,\langle 1,4 \mid 3\rangle,\langle 2,3 \mid \varnothing\rangle,\langle 2,4 \mid \{1,3\}\rangle\}$$

which is not of the form K_θ for any order of the variables of the sets $\{1,2\}$ and $\{3,4\}$. □

To conclude we note that for a singleton starting set $\{\langle\mathbf{A},\mathbf{B} \mid \mathbf{C}\rangle\}$ with $\mathbf{A} = \{A_1,\ldots,A_n\}$, $n \geq 1$, and $\mathbf{B} = \{B_1,\ldots,B_m\}$, $m \geq 1$, the lower bound of Proposition 4 equals $n \cdot m$ while the lower bound of Proposition 5 is $n + m - 1$. For a singleton starting set therefore, the lower bound of Proposition 4 is equal to or larger than the bound of Proposition 5.

5. Minimum Elementary Bases in General

Proposition 8 established the number of triplets in a minimum elementary-triplet basis for a starting set composed of a single triplet. We now address starting sets with arbitrary numbers of triplets and provide an upper bound on the number of triplets in a minimum elementary-triplet basis for such starting sets.

Corollary 9 *Let $J = \{\theta_1,\ldots,\theta_k\}$, with $\theta_i = \langle \mathbf{A}_i,\mathbf{B}_i \mid \mathbf{C}_i\rangle$, $i = 1,\ldots,k$, be a starting triplet set and let $\bar{J}$ be its closure. Let n_i,m_i be the number of variables in the sets $\mathbf{A}_i,\mathbf{B}_i$, respectively. A minimum elementary-triplet basis for $\bar{J}$ has at most $\sum_{i=1,\ldots,k} n_i \cdot m_i$ elements.*

Proof The result follows directly from Proposition 8. ∎

The upper bound stated above is tight, in the sense that there exist starting sets for which this bound is attained.

For any starting set J, an elementary triplet basis of a size equal to the upper bound provided in the corollary above is constructed straightforwardly by adding for every triplet $\theta \in J$ the subset $K_\theta \in J$ detailed in Lemma 7. The thus constructed basis may include redundant triplets, as elementary triplets found from one triplet in J may imply elementary triplets from another such triplet. A non-redundant elementary-triplet basis can now be readily found by iteratively removing triplets and checking whether the resulting basis still has $\bar{J}$ for its closure. Such a procedure unfortunately does not guarantee that a minimum basis is found, as illustrated by the following example.

Example 3 *Let $\bar{J}^{\mathbf{E}} = \{\theta_i\}, i = 1,\ldots,7$, be the elementary triplets of a semi-graphoid independence relation $\bar{J}$ with:*

$$\theta_1 = \langle 1,2 \mid 4\rangle \qquad \theta_5 = \langle 1,3 \mid \{2,4\}\rangle$$
$$\theta_2 = \langle 1,2 \mid \{3,4\}\rangle \qquad \theta_6 = \langle 1,4 \mid 2\rangle$$
$$\theta_3 = \langle 1,3 \mid 2\rangle \qquad \theta_7 = \langle 1,4 \mid \{2,3\}\rangle$$
$$\theta_4 = \langle 1,3 \mid 4\rangle$$

In this basis $\bar{J}^{\mathbf{E}}$, the triplet pairs (θ_1,θ_5) and (θ_2,θ_4) are equivalent by the derivation rule E2, as are the pairs (θ_3,θ_7) and (θ_5,θ_6). Now suppose that the basis is reduced in size by first removing the triplets θ_1,θ_5, and then deleting the triplet θ_6. The resulting basis $\{\theta_2,\theta_3,\theta_4,\theta_7\}$ cannot be further reduced in size, and includes four triplets. An example elementary-triplet basis for $\bar{J}$ of minimum size is $\{\theta_1,\theta_5,\theta_6\}$, including three triplets. □

The example above shows that a procedure of iteratively removing redundant triplets from an elementary-triplet basis does not necessarily result in a basis of minimum size. Such a procedure would therefore constitute just a heuristic resulting in a minimal elementary-triplet basis.

6. Conclusions and Future Work

For concisely representing a semi-graphoid independence relation, a subset of its triplets is listed explicitly in a basis, leaving its other triplets implicit by the associated axioms. In this paper, we addressed the use of an elementary-triplet basis for this purpose, composed of all elementary triplets of an independence relation, as proposed by Peña [19]. Since elementary triplets may occur in equivalent pairs, a relation's set of all such triplets may include redundant information. We provided two lower bounds on the size of an elementary-triplet basis in general and showed that an elementary-triplet basis of minimum size can be readily constructed for semi-graphoid relations defined by a single arbitrary triplet. We further derived an upper bound on the size of a minimum elementary-triplet basis for semi-graphoid relations in general, and detailed a procedure for constructing minimal elementary-triplet bases.

In our future research we intend to further investigate algorithms for constructing elementary-triplet bases of minimum size for semi-graphoid independence relations in general. Having focused so far on semi-graphoid relations, we further intend to study alternative concepts of probabilistic independence and irrelevance, as found for example in frameworks of imprecise probability, for which the semi-graphoid axioms of independence may not all hold (see for example [4, 7, 16]), and address their concise representation by triplet bases.

Acknowledgments

References

[1] M. Baioletti, G. Busanello, and B. Vantaggi. Conditional independence structure and its closure: Inferential rules and algorithms. *International Journal of Approximate Reasoning*, 50:1097–1114, 2009.

[2] M. Baioletti, G. Busanello, and B. Vantaggi. Closure of independencies under graphoid properties: Some experimental results. In Th. Augustin, F.P.A. Coolen, S. Moral, and M.C.M. Troffaes, editors, *Proceedings 6th International Symposium on Imprecise Probability: Theories and Applications, ISIPTA'09*, pages 11–19. SIPTA, 2009.

[3] G. Coletti and R. Scozzafava. Zero probabilities in stochastic independence. In B. Bouchon-Meunier, R. Yager, and L. Zadeh, editors, *Information, Uncertainty, Fusion*, pages 185–196. Kluwer Academic Publishers, Dordrecht, 2000.

[4] I. Couso, S. Moral, and P. Walley. A survey of concepts of independence for imprecise probabilities. *Risk, Decision and Policy*, 5:165–181, 2000.

[5] F.G. Cozman. Sets of probability distributions, independence, and convexity. *Synthese*, 186(2):577–600, 2012.

[6] F.G. Cozman and C. Polpo de Campos. Kuznetsov independence for interval-valued expectations and sets of probability distributions: Properties and algorithms. *International Journal of Approximate Reasoning*, 55: 666–682, 2014.

[7] F.G. Cozman and P. Walley. Graphoid properties of epistemic irrelevance and independence. *Annals of Mathematics and Artificial Intelligence*, 45(1):173–195, 2005.

[8] A.P. Dawid. Conditional independence in statistical theory. *Journal of the Royal Statistical Society B*, 41: 1–31, 1997.

[9] A.P. Dawid. Separoids: A mathematical framework for conditional independence and irrelevance. *Annals of Mathematics and Artificial Intelligence*, 32(1):335–372, 2001.

[10] P. de Waal and L.C. van der Gaag. Stable independence and complexity of representation. In M. Chickering and J. Halpern, editors, *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence*, pages 112–119. AUAI Press, Arlington, 2004.

[11] D. Geiger and J. Pearl. Logical and algorithmic properties of conditional independence and graphical models. *Annals of Statistics*, 21:2001–2021, 1993.

[12] D. Geiger, A. Paz, and J. Pearl. Axioms and algorithms for inferences involving probabilistic independence. *Information and Computation*, 91:128–141, 1991.

[13] S. Lopatatzidis and L.C. van der Gaag. Concise representations and construction algorithms for semi-graphoid independency models. *International Journal of Approximate Reasoning*, 80:377–392, 2017.

[14] F. Matúš. Ascending and descending conditional independence relations. In S. Kubik and J.A. Visek, editors, *Proceedings of the Joint Session of the 11th Prague Conference on Asymptotic Statistics and the 13th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes*, pages 189–200. Kluwer, 1992.

[15] F. Matúš. Towards classification of semigraphoids. *Discrete Mathematics*, 277:115–145, 2004.

[16] S. Moral and A. Cano. Strong conditional independence for credal sets. *Annals of Mathematics and Artificial Intelligence*, 35(1):295–321, 2002.

[17] M. Niepert, B. Sayrafi, M. Gyssens, and D. Van Gucht. The conditional independence implication problem: A lattice-theoretic approach. *Artificial Intelligence*, 202:29–51, 2013.

[18] J. Pearl. *Probabilistic Reasoning in Intelligent Systems. Networks of Plausible Inference*. Morgan Kaufmann, 1988.

[19] J.M. Peña. Representing independence models with elementary triplets. *International Journal of Approximate Reasoning*, 88:587–601, 2017.

[20] W. Spohn. Stochastic independence, causal independence and shieldability. *Journal of Philosophical Logic*, 9:73–99, 1980.

[21] M. Studený. Conditional independence relations have no finite complete characterization. In S. Kubik and J.A. Visek, editors, *Proceedings of the Joint Session of the 11th Prague Conference on Asymptotic Statistics and the 13th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes*, pages 377–396. Kluwer, 1992.

[22] M. Studený. Semigraphoids and structures of probabilistic conditional independence. *Annals of Mathematics and Artificial Intelligence*, 21:71–98, 1997.

[23] M. Studený. Complexity of structural models. In *Proceedings of the Joint Session of the 6th Prague Conference on Asymptotic Statistics and the 13th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes*, pages 521–528, 1998.

[24] M. Studený. *Probabilistic Conditional Independence Structures*. Springer, 2005.

[25] B. Vantaggi. Conditional independence in a coherent setting. *Annals of Mathematics and Artificial Intelligence*, 32:287–313, 2001.

[26] B. Vantaggi. Conditional independence structures and graphical models. *International Journal of Uncertainty Fuzziness and Knowledge-Based Systems*, 11(5):545–571, 2003.