

A Ternary Knowledge Relation on Secrets

Sara Miner More
Department of Mathematics
and Computer Science
McDaniel College
Westminster, Maryland 21157
smore@mcdaniel.edu

Pavel Naumov
Department of Mathematics
and Computer Science
McDaniel College
Westminster, Maryland 21157
pnaumov@mcdaniel.edu

Brittany Nicholls
Department of Mathematics
and Computer Science
McDaniel College
Westminster, Maryland 21157
brn002@mcdaniel.edu

Andrew Yang
Department of Mathematics
and Computer Science
McDaniel College
Westminster, Maryland 21157
asy001@mcdaniel.edu

ABSTRACT

The paper introduces and studies the ternary relation “secret a reveals at least as much information about secret c as secret b .” In spite of its seeming simplicity, this relation has many non-trivial properties. The main result is a complete infinite axiomatization of the propositional theory of this relation.

Categories and Subject Descriptors

I.2.4 [Artificial Intelligence]: Knowledge Representation Formalisms and Methods; F.4.1 [Mathematical Logic]: Mathematical Logic; I.2.3 [Artificial Intelligence]: Deduction and Theorem Proving

General Terms

Theory

Keywords

information flow, secret, knowledge, completeness

1. INTRODUCTION

In this paper, we study the properties of interdependencies between pieces of information. We call these pieces *secrets* to emphasize the fact that they might be unknown to some parties.

1.1 Functional Dependence and Independence

One of the simplest relations between two secrets is *functional dependence*. We denote it by $a \triangleright b$. It means that the value of secret a reveals the value of secret b . This relation

is reflexive and transitive. A more general and less trivial form of functional dependence is functional dependence between sets of secrets. If A and B are two sets of secrets, then $A \triangleright B$ means that, together, the values of all secrets in A reveal the values of all secrets in B . Armstrong [1] presented the following sound and complete axiomatization of this relation:

1. *Reflexivity*: $A \triangleright B$, if $A \supseteq B$,
2. *Augmentation*: $A \triangleright B \rightarrow A, C \triangleright B, C$,
3. *Transitivity*: $A \triangleright B \rightarrow (B \triangleright C \rightarrow A \triangleright C)$,

where here and everywhere below A, B denotes the union of sets A and B . The above axioms are known in database literature as Armstrong’s axioms [4, p. 81]. Beeri, Fagin, and Howard [2] suggested a variation of Armstrong’s axioms that describe properties of multi-valued dependency.

Not all dependencies between two secrets are functional. For example, if secret a is the area of a triangle and secret p is the perimeter of the same triangle, then there is an interdependence between these secrets in the sense that not every value of secret a is compatible with every value of secret p . However, neither $a \triangleright p$ nor $p \triangleright a$ is necessarily true. If there is no interdependence between two secrets, then we will say that the two secrets are *independent*. In other words, secrets a and b are independent if any possible value of secret a is compatible with any possible value of secret b . We denote this relation between two secrets by $a \parallel b$. This relation was introduced by Sutherland [14] and is known in the theory of information flow as *nondeducibility*. Halpern and O’Neill [6] proposed a closely related notion called *f*-secrecy. Kelvey, More, Naumov, and Sapp [9] gave a complete axiomatization of properties that connect relations $a \parallel b$ and $a \triangleright b$. More and Naumov also described properties of a multi-argument variation of the relation $a \parallel b$ under the assumption that the secrets are generated over an undirected graph [10], a directed acyclic graph [3], or a hypergraph [11] with a fixed topology as well as similar properties of relation $A \triangleright B$ over undirected graphs [12].

Like functional dependence, independence also can be generalized to relate two sets of secrets. If A and B are two such sets, then $A \parallel B$ means that any consistent combination

ACM COPYRIGHT NOTICE. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Publications Dept., ACM, Inc., fax +1 (212) 869-0481, or permissions@acm.org. TARK 2011, July 12-14, 2011, Groningen, The Netherlands. Copyright ©2011 ACM. ISBN 978-1-4503-0707-9, \$10.00.

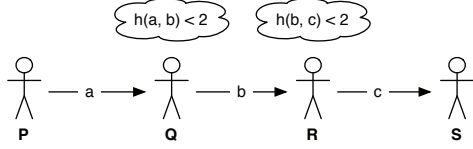


Figure 1: Telephone Game.

of values of secrets in A is compatible with any consistent combination of values of secrets in B . Note that “consistent combination” is an important condition here since some interdependence may exist between secrets in set A even while the entire set of secrets A is independent from the secrets in set B . A sound and complete axiomatization of this independence relation between sets was given by More and Naumov [12]:

1. *Empty Set*: $\emptyset \parallel A$,
2. *Monotonicity*: $A, B \parallel C \rightarrow A \parallel C$,
3. *Symmetry*: $A \parallel B \rightarrow B \parallel A$,
4. *Public Knowledge*: $A \parallel A \rightarrow (B \parallel C \rightarrow A, B \parallel C)$,
5. *Exchange*: $A, B \parallel C \rightarrow (A \parallel B \rightarrow A \parallel B, C)$.

Essentially the same axioms were shown by Geiger, Paz, and Pearl [5] to provide a complete axiomatization of the independence relation between random variables in probability theory.

Suppose now that a , b , and c are three secrets with integer values such that $a + b = c$. Note that $a \parallel b$ is true since every possible value of a is consistent with any possible value of b . Note, however, that if value of c is fixed, then not every possible value of secret a is compatible with every possible value of secret b . We will say that secrets a and b are not independent conditionally on c and denote this by $\neg(a \parallel_c b)$. The conditional independence relation is also known as *embedded multivalued dependency* in database theory. Herrmann [7, 8] proved the undecidability of the propositional theory of the conditional independence relation on *sets* of secrets. Studený [13] has shown that the related conditional independence in probability theory has no complete finite characterization.

1.2 The Ternary Knowledge Relation

If secret b is functionally determined by secret a , or in our notation, $a \triangleright b$, then secret a reveals at least as much information as secret b . In this paper we study the ternary knowledge relation “secret a reveals at least as much information about secret c as secret b .” For instance, consider the variation of the Telephone game¹ depicted in Figure 1: person P picks a random binary string a and communicates it to Q . Person Q changes at most one bit of a , and communicates it to person R as b . Finally, R again changes at most one bit in b and communicates it to S as c . Note that in this situation secret c is not functionally determined by secret b , however, knowing string b reveals more about string a than knowing string c . Indeed, suppose that a_0 , b_0 , and c_0 are the values of a , b , and c , respectively, in a certain round of the game. This of course, means that $h(b_0, c_0) \leq 1$, where

¹This game is also known as Chinese Whispers, Grapevine, Broken Telephone, Whisper Down the Lane, and Gossip.

$h(x, y)$ is the Hamming distance between strings x and y . If somebody knows b_0 , then this person can conclude that $a_0 \in \text{Ball}(b_0, 1) = \{x \mid h(x, b_0) \leq 1\}$. At the same time, if one knows only c_0 , then all that can be concluded about string a_0 is that $a_0 \in \text{Ball}(c_0, 2) = \{x \mid h(x, c_0) \leq 2\}$. Note that $\text{Ball}(b_0, 1) \subset \text{Ball}(c_0, 2)$ due to $h(b_0, c_0) \leq 1$ and the triangle inequality. Therefore, in any round of the game, the value of secret b always reveals at least as much about the value of secret a the value of secret c . We will denote this by $b \triangleright_a^c$. One can similarly show that $b \triangleright_c^a$.

Of course, although statement $b \triangleright_a^c$ is true for the Telephone game semantics, it might be false for some other interpretation of secrets a , b , and c . In this paper we study the logical properties of relation $a \triangleright_c^b$ that are true for any secrets. A trivial example of such a property is transitivity:

$$a \triangleright_c^b \rightarrow (b \triangleright_a^d \rightarrow a \triangleright_c^d).$$

It turns out, however, that in spite of the seeming simplicity of this relation, it has many non-trivial properties. For example, the following statement is true for any secrets a , b , c , d , e , and f :

$$(a \triangleright_c^b) \wedge (b \triangleright_d^e) \wedge (c \triangleright_f^d) \wedge (d \triangleright_f^e) \rightarrow (a \triangleright_f^e).$$

To see the pattern in the assumptions of the above formula, we can arrange them into a “diamond” shape:

$$\begin{array}{ccc} & b \triangleright_d^e & \\ a \triangleright_c^b & & d \triangleright_f^e \\ & c \triangleright_f^d & \end{array} \rightarrow a \triangleright_f^e. \quad (1)$$

In some sense, this property is a ternary version of transitivity. An even more general version of transitivity is captured by the following formula, which, as we will show, is also true for any secrets:

$$\begin{array}{ccccc} & & e \triangleright_h^g & & \\ & b \triangleright_d^e & & h \triangleright_k^g & \\ a \triangleright_c^b & & d \triangleright_i^h & & k \triangleright_j^g \\ & c \triangleright_f^d & & i \triangleright_j^k & \\ & & f \triangleright_j^i & & \end{array} \rightarrow a \triangleright_j^g. \quad (2)$$

We will prove soundness of the principles (1) and (2) in Theorem 4.

The main result of this paper is a complete infinite axiomatization of relation $a \triangleright_c^b$ between three arbitrary secrets. The above principles (1) and (2) are two instances of the transitivity axiom schema in our logical system. In the conclusion of this paper, we discuss a connection between relation $a \triangleright_c^b$ and embedded multivalued dependency.

2. SEMANTICS

We assume a fixed alphabet of “secret” variables: a, b, \dots . By an atomic formula we mean either \perp or $a \triangleright_c^b$ for some secret variables a , b , and c . By formula we mean either an atomic formula or a combination of several atomic formulas using binary connective \rightarrow . All other boolean connectives are assumed to be defined through \perp and \rightarrow .

DEFINITION 1. A protocol is a pair $\mathcal{P} = \langle V, R \rangle$, where,

1. for any secret variable a , set $V(a)$ is an arbitrary set of “values” of secret a ,
2. R is a set of functions r on secret variables such that $r(a) \in V(a)$ for any secret variable a . Elements of R will be called “runs” of the protocol.

In a given protocol, if $b_0 \in V(b)$ is a value of secret b , than by $Ball_a(b_0)$ we will mean the set of all possible values of a that are consistent with value b_0 . We use the notation $Ball$ to emphasize connection with $Balls$ defined through the Hamming distance metric in the previous section. The formal definition of $Ball$, in the more general setting of an arbitrary protocol, is, of course, different:

DEFINITION 2. For any protocol $\langle V, R \rangle$, any two secret variables a and b , and any $b_0 \in V(b)$,

$$Ball_a(b_0) = \{r(a) \mid r(b) = b_0 \text{ and } r \in R\}.$$

DEFINITION 3. For any protocol $\mathcal{P} = \langle V, R \rangle$ and any formula ϕ , we define the binary relation $\mathcal{P} \models \phi$ as follows:

1. $\mathcal{P} \not\models \perp$,
2. $\mathcal{P} \models a \triangleright_c^b$ if and only if, for any $r \in R$,
 $Ball_c(r(a)) \subseteq Ball_c(r(b))$,
3. $\mathcal{P} \models \phi \rightarrow \psi$ if and only if $\mathcal{P} \not\models \phi$ or $\mathcal{P} \models \psi$.

3. DIAMOND NOTATION

Before stating the axioms of our logical system, we want to introduce a compact notation for the diamond-shaped patterns of formulas that has already appeared in formulas (1) and (2). In general, we will consider patterns depicted in Figure 2, where $\{a_j^i\}_{i,j}$ are secret variables. For such patterns, it will be assumed that $a_0^n = a_0^{n+1} = \dots = a_0^{2n-1}$ and $a_n^n = a_{n-1}^{n+1} = \dots = a_1^{2n-1}$. In other words, all variables along the upper-right edge of the diamond are the same and all variables along the lower-right edge of the diamond are also the same. No other assumptions about variables in the diamond pattern are made. In particular, the variables along the upper-right edge do not have to be the same as the variables along lower-right edge.

We will also use diamond patterns as propositional formulas. If a diamond pattern appears as a formula, then it should be viewed as notation for the conjunction

$$\bigwedge_{i,j} a_j^i \triangleright_{a_{j+1}^{i+1}}^{a_j^{i+1}}, \quad (3)$$

where the conjunction is taken for all pairs (i, j) except for those that correspond to variables a_j^i that are located along upper-right or lower-right edge of the diamond.

For example, the formula which appeared earlier as (1) can now be written more compactly as the following implication between two diamonds:

$$\begin{array}{ccccc} & & e & & \\ & b & & e & f \\ a & & d & & \\ & c & & f & e \\ & & f & & \end{array} \longrightarrow \begin{array}{ccccc} & & a & & \\ & & & & e \end{array}.$$

Similarly, formula (2) can now be written as the implication

$$\begin{array}{ccccc} & & g & & \\ & e & & g & \\ a & & b & & h & g \\ & d & & k & \\ & c & & i & j \\ & & f & & j \\ & & & & j \end{array} \longrightarrow \begin{array}{ccccc} & & a & & \\ & & & & j \end{array}.$$

Note a certain resemblance between condition (3) and the recurrence relation defining the Pascal triangle.

4. AXIOMS

In addition to the propositional tautologies and the Modus Ponens inference rule, our logical system includes the following axioms of Reflexivity, Symmetry, and Transitivity. Transitivity is technically a schema that generates infinitely many axioms for diamond patterns of different sizes.

Reflexivity

$$\begin{array}{c} a \\ a \\ b \end{array}$$

Symmetry

$$\begin{array}{ccc} & b & c \\ a & \longrightarrow & a \\ & c & b \end{array}$$

Transitivity

$$\begin{array}{ccccccc} & & & & b & & \\ & \dots & \dots & \dots & & & \\ & & & & & & \\ a & & \dots & \dots & & \longrightarrow & a \\ & \dots & \dots & \dots & c & & c \\ & & & & & & \\ & & & & c & & \end{array}$$

Of course, the Reflexivity and Symmetry axioms can be stated without diamond notation as: $a \triangleright_c^a$ and $a \triangleright_c^b \rightarrow a \triangleright_c^c$ respectively. Formulas (1) and (2) are instances of the Transitivity schema. While the soundness of the Reflexivity axiom is straightforward, the soundness of the Symmetry axiom and the Transitivity schema is not immediately obvious. We prove the soundness of all three axioms in the next section.

We will write $X \vdash \phi$ to state that that formula ϕ is provable in our logical system using additional (possibly empty) set of axioms X .

5. SOUNDNESS

THEOREM 1 (REFLEXIVITY). $\mathcal{P} \models a \triangleright_c^a$, for any protocol \mathcal{P} .

PROOF. For any run r of protocol \mathcal{P} ,

$$Ball_b(r(a)) \subseteq Ball_b(r(a))$$

due to the reflexivity of the subset relation. \square

Although relation $\mathcal{P} \models a \triangleright_c^b$ is defined in terms of sets $Ball_c(a)$ and $Ball_c(b)$, proving many properties of this relation is much easier using an alternative definition captured by the following definition and theorem:

DEFINITION 4. For any secret variable a , runs r_1 and r_2 are a -equivalent if $r_1(a) = r_2(a)$.

We denote this relation by $r_1 \equiv_a r_2$.

THEOREM 2. If \mathcal{P} is an arbitrary protocol, then $\mathcal{P} \models a \triangleright_c^b$ if and only if $\forall r_1 \forall r_2 (r_1 \equiv_a r_2 \rightarrow \exists r (r_1 \equiv_b r \equiv_c r_2))$, where the quantifiers are over the set of all runs of protocol \mathcal{P} .

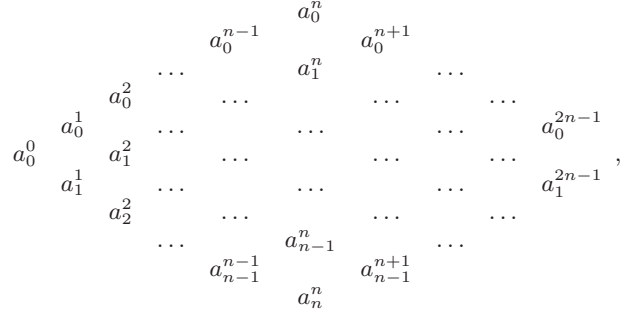


Figure 2: Diamond Pattern

PROOF. (\Rightarrow) Suppose r_1 and r_2 are runs of \mathcal{P} such that $r_1 \equiv_a r_2$. We will show that there is a run r such that $r_1 \equiv_b r \equiv_c r_2$. Indeed, by the assumption of the theorem, $\mathcal{P} \models a \triangleright_c^b$. Thus, $Ball_c(r_1(a)) \subseteq Ball_c(r_1(b))$. Taking into account the assumption $r_1 \equiv_a r_2$, we can conclude that $Ball_c(r_2(a)) \subseteq Ball_c(r_1(b))$. Note that this means

$$r_2(c) \in \{r(c) \mid r(a) = r_2(a)\} = Ball_c(r_2(a)) \subseteq Ball_c(r_1(b)) = \{r(c) \mid r(b) = r_1(b)\}.$$

Therefore, there must be a run r such that $r_1 \equiv_b r \equiv_c r_2$.

(\Leftarrow) We will show that $Ball_c(r_1(a)) \subseteq Ball_c(r_1(b))$ for any run r_1 of protocol \mathcal{P} . Assume that $c_0 \in Ball_c(r_1(a))$. We will prove that $c_0 \in Ball_c(r_1(b))$. Note that the assumption $c_0 \in Ball_c(r_1(a))$, by Definition 2, implies that $c_0 = r_2(c)$ for some run r_2 such that $r_2 \equiv_a r_1$. Thus, by the assumption of the theorem, there must be a run r such that $r_1 \equiv_b r \equiv_c r_2$. Hence,

$$c_0 = r_2(c) \in \{r(c) \mid r(b) = r_1(b)\} = Ball_c(r_1(b)).$$

□

THEOREM 3 (SYMMETRY). For any protocol \mathcal{P} , if $\mathcal{P} \models a \triangleright_c^b$, then $\mathcal{P} \models a \triangleright_b^c$.

PROOF. Follows from Theorem 2 and symmetry of the relation $r_1 \equiv_a r_2$. □

THEOREM 4 (TRANSITIVITY). Suppose \mathcal{P} is a protocol such that $\mathcal{P} \models a_j^i \triangleright_{a_{j+1}^{i+1}}^{a_{j+1}^i}$ for every i and j , where a_j^i is not located on either the upper-right or lower-right edge of the diamond pattern (see Figure 2). For any runs r^- and r^+ of protocol \mathcal{P} such that $r^- \equiv_{a_0^n} r^+$, there is a run r of protocol \mathcal{P} such that $r^- \equiv_{a_0^{2n-1}} r \equiv_{a_1^{2n-1}} r^+$.

PROOF. Assume that $r^- \equiv_{a_0^n} r^+$.

LEMMA 1. For any $0 \leq i \leq n$, there are runs r_0, \dots, r_{i-1} such that

$$r^- \equiv_{a_0^i} r_0 \equiv_{a_1^i} r_1 \equiv_{a_2^i} \dots \equiv_{a_{i-1}^i} r_{i-1} \equiv_{a_i^i} r^+.$$

PROOF. We use induction on i . If $i = 0$, then $r^- \equiv_{a_0^0} r^+$ by our assumption. Suppose now that

$$r^- \equiv_{a_0^i} r_0 \equiv_{a_1^i} r_1 \equiv_{a_2^i} \dots \equiv_{a_{i-1}^i} r_{i-1} \equiv_{a_i^i} r^+. \quad (4)$$

By Theorem 2 and the equivalences from line (4), there must be runs r'_0, \dots, r'_i such that

$$\begin{aligned} r^- &\equiv_{a_0^{i+1}} r'_0 \equiv_{a_1^{i+1}} r_0 \\ r_0 &\equiv_{a_1^{i+1}} r'_1 \equiv_{a_2^{i+1}} r_1 \\ &\dots \\ r_{i-1} &\equiv_{a_{i-1}^{i+1}} r'_i \equiv_{a_i^{i+1}} r^+. \end{aligned}$$

Thus,

$$\begin{aligned} r^- &\equiv_{a_0^{i+1}} r'_0 \equiv_{a_1^{i+1}} r'_1 \equiv_{a_2^{i+1}} r'_2 \equiv_{a_3^{i+1}} \dots \\ &\dots \equiv_{a_{i-1}^{i+1}} r'_{i-1} \equiv_{a_i^{i+1}} r'_i \equiv_{a_{i+1}^{i+1}} r^+. \end{aligned}$$

□

LEMMA 2. For any integer $0 \leq i \leq n-1$, there are runs r_0, \dots, r_{n-i-1} such that

$$r^- \equiv_{a_0^{n+i}} r_0 \equiv_{a_1^{n+i}} r_1 \equiv_{a_2^{n+i}} \dots \equiv_{a_{n-i-1}^{n+i}} r_{n-i-1} \equiv_{a_{n-i}^{n+i}} r^+.$$

PROOF. Induction on i . If $i = 0$, then the statement is true by Lemma 1. Suppose now that

$$\begin{aligned} r^- &\equiv_{a_0^{n+i}} r_0 \equiv_{a_1^{n+i}} r_1 \equiv_{a_2^{n+i}} \dots \\ &\dots \equiv_{a_{n-i-1}^{n+i}} r_{n-i-1} \equiv_{a_{n-i}^{n+i}} r^+. \end{aligned} \quad (5)$$

By Theorem 2, there must be runs r'_0, \dots, r'_{n-i-2} such that

$$\begin{aligned} r_0 &\equiv_{a_0^{n+i+1}} r'_0 \equiv_{a_1^{n+i+1}} r_1 \\ r_1 &\equiv_{a_1^{n+i+1}} r'_1 \equiv_{a_2^{n+i+1}} r_2 \\ &\dots \\ r_{n-i} &\equiv_{a_{n-i-2}^{n+i+1}} r'_{n-i-2} \equiv_{a_{n-i-1}^{n+i+1}} r_{n-i-1}. \end{aligned}$$

Thus, taking into account equivalencies (5),

$$\begin{aligned} r^- &\equiv_{a_0^{n+i}} r_0 \equiv_{a_0^{n+i+1}} r'_0 \equiv_{a_1^{n+i+1}} r_1 \equiv_{a_1^{n+i+1}} r'_1 \equiv_{a_2^{n+i+1}} \\ &\dots \equiv_{a_{n-i-2}^{n+i+1}} r'_{n-i-2} \equiv_{a_{n-i-1}^{n+i+1}} r_{n-i-1} \equiv_{a_{n-i}^{n+i+1}} r^+. \end{aligned}$$

Recall that a diamond pattern must contain the same variables along the upper-right and lower-right edges. In other words, a_0^{n+i} is the same variable as a_0^{n+i+1} and a_{n-i}^{n+i} is the same variable as a_{n-i-1}^{n+i+1} . Thus,

$$\begin{aligned} r^- &\equiv_{a_0^{n+i+1}} r_0 \equiv_{a_0^{n+i+1}} r'_0 \equiv_{a_1^{n+i+1}} r_1 \equiv_{a_1^{n+i+1}} r'_1 \equiv_{a_2^{n+i+1}} \\ &\dots \equiv_{a_{n-i-2}^{n+i+1}} r'_{n-i-2} \equiv_{a_{n-i-1}^{n+i+1}} r_{n-i-1} \equiv_{a_{n-i-1}^{n+i+1}} r^+. \end{aligned}$$

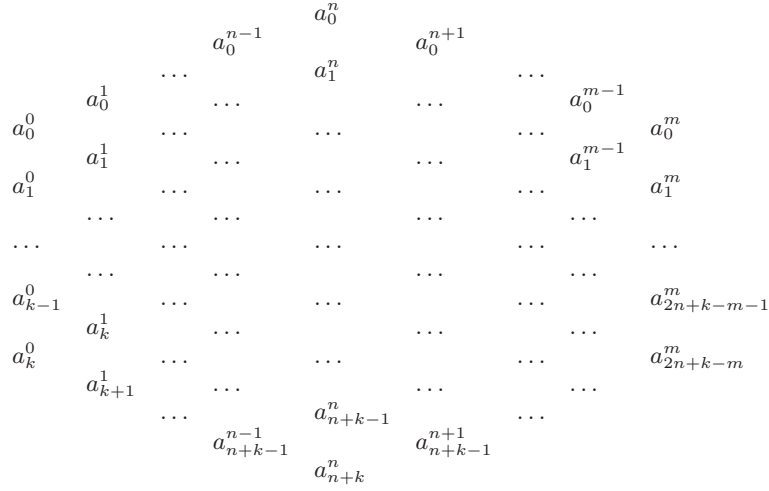
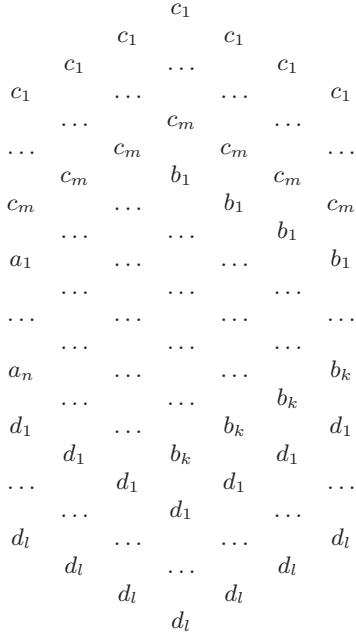


Figure 3: Hexagonal Pattern

pattern between layers of c_1, \dots, c_m and d_1, \dots, d_l :



To finish the proof, we need to show that condition 2 from Definition 5 is satisfied. Indeed, it follows from axioms of Reflexivity and Symmetry and the fact that the same condition is satisfied in the original pattern. \square

LEMMA 7. $A, C \circ_X A, b, C$, for any secret variable b and any two sequences A and C such that at least one of sequences A and C is not empty.

PROOF. Without loss of generality (due to Lemma 4), assume that sequence A is non-empty. Let $A = A', a$ for some secret variable a . Consider hexagonal pattern



Due to the Reflexivity axiom, $\vdash a \triangleright_b^a$. Thus, $a \circ_X a, b$.

By Lemma 6, we have $A', a, C \circ_X A', a, b, C$. Therefore, $A, C \circ_X A, b, C$. \square

LEMMA 8. $A, b, b, C \circ_X A, b, C$, for any secret variable b and any two sequences of secret variables A and C .

PROOF. Consider the hexagonal pattern



Thus, $b, b \circ_X b$. By Lemma 6, $A, b, b, C \circ_X A, b, C$. \square

DEFINITION 6. For any $n \geq 0$ and any secret variable a , by a^n we mean the sequence $\underbrace{a, \dots, a}_n$.

LEMMA 9. $a^n \circ_X a$, for any $n \geq 1$ and any variable a .

PROOF. We use induction on n . *Base Case:* If $n = 1$, then the required follows from Lemma 3. *Induction Step:* Let $n > 1$. Assume $a^{n-1} \circ_X a$. By Lemma 8, since $n > 1$, we have $a^n \circ_X a^{n-1}$. By Lemma 5, we can conclude that $a^n \circ_X a$. \square

LEMMA 10. $a^n, b^m \circ_X a, b$, for any secret variable a and any $n, m \geq 0$ such that $n + m \geq 1$.

PROOF. Due to Lemma 4, without loss of generality we may assume that $n > 0$. We will consider cases $m = 0$ and $m > 0$ separately. *Case I:* If $m = 0$, then, by Lemma 9, $a^n \circ_X a$. At the same time, by Lemma 7, we have $a \circ_X a, b$. Hence, by Lemma 5, $a^n \circ_X a, b$.

Case II: If $m > 0$, then, by Lemma 9, $a^n \circ_X a$ and $b^m \circ_X b$. By Lemma 6, $a^n, b^m \circ_X a, b^m$ and $a, b^m \circ_X a, b$. Finally, by Lemma 5, $a^n, b^m \circ_X a, b$. \square

6.2 Graph Semantics

In this section we will define a “graph semantics” for the relation $a \triangleright_c^b$ and prove the completeness of our formal system with respect to this new semantics. Later we will use this result to prove completeness with respect to the original semantics of secrets.

By graph we mean a (possibly **infinite**) undirected graph whose edges are labeled by secret variables. Each edge will be assumed to have a unique label. Multiple edges between the same vertices are allowed, but loop edges are not.

Let a be a secret variable. We say that two vertices are a -equivalent, if there is a path between these two vertices such that each edge along this path is labeled with a . Note that a -equivalence is an equivalence relation on vertices. If vertices u and v are a -equivalent, then we write $u \sim_a v$.

DEFINITION 7. For any graph G and any formula ϕ , we define the binary relation $G \models \phi$ as follows:

1. $G \not\models \perp$,
2. $G \models a \triangleright_c^b$ if and only if, for any vertices v and u such that $v \sim_a u$, there is a vertex w such that $v \sim_b w$ and $w \sim_c u$.
3. $G \models \phi \rightarrow \psi$ if and only if $G \not\models \phi$ or $G \models \psi$.

THEOREM 5. If $G \models \phi$, for each graph G , then $\vdash \phi$.

PROOF. Suppose that $\not\vdash \phi$. Let X be a (countable) maximal consistent set of formulas that contains $\neg\phi$. Let $\{a_i \triangleright_{c_i}^{b_i}\}_{i \in I}$ be the (at most countable) set of all atomic formulas in X and $\{d_j \triangleright_{f_j}^{e_j}\}_{j \in J}$ be the (at most countable) set of all atomic formulas that do not belong to X .

For each $j \in J$, we define an infinite chain of finite graphs $G_0^j \subset G_1^j \subset G_2^j \subset \dots$ such that G_k^j is a subgraph of G_{k+1}^j for each k . Let G_0^j be a graph with just two vertices, denoted by v^- and v^+ , and a single edge between these two vertices labeled by d_j .

Assume that G_k^j is already defined and that vertices u and v are a_i -equivalent in graph G_k^j for some $i \in I$. We define graph G_{k+1}^j by adding a new vertex w and edges (u, w) and (w, v) to graph G_k^j . Edge (u, w) is labeled with b_i and edge (w, v) is labeled with c_i . Note that the construction of graph G_{k+1}^j depends on the particular choice of u , v , and i . We will specify this choice later. Let $G^j = \bigcup_k G_k^j$.

LEMMA 11. If there is a simple² path π in graph G^j from v^- to v^+ labeled by sequence $L = l_1, \dots, l_n$, then $d_j \circ_X L$.

PROOF. Consider the chain $G_0^j \subset G_1^j \subset \dots$, and let G_k^j be the first graph in the chain that contains the entire path π . We will prove the lemma by induction on k .

Base Case: If π existed in G_0^j , then $L = l_1 = d_j$. Hence, by Lemma 3, $d_j \circ_X L$.

Induction Step: Suppose now that path π first appeared in graph G_{k+1}^j , which was obtained by adding new vertex w and edges (u, w) and (w, v) labeled with b_i and c_i respectively, where $a_i \triangleright_{c_i}^{b_i} \in X$ and $u \sim_{a_i} v$. Thus, path π must contain edges (u, w) and (w, v) . There are two possible orders in which path π can go through these two edges (see Figure 4). We consider these two cases separately.

Case 1: Path π , in the direction from v^- to v^+ , first passes through edge (u, w) and then edge (w, v) . Thus, we have $L = L_1, b_i, c_i, L_2$, where labels L_1 are on the edges along path π between vertices v^- and u and L_2 are on the edges along path π between vertices v and v^+ . Since $u \sim_{a_i} v$, there must be a path between u and v in graph G_k^j whose edges are all labeled by a_i . Thus, in graph G_k^j , there was a

²without self-intersections

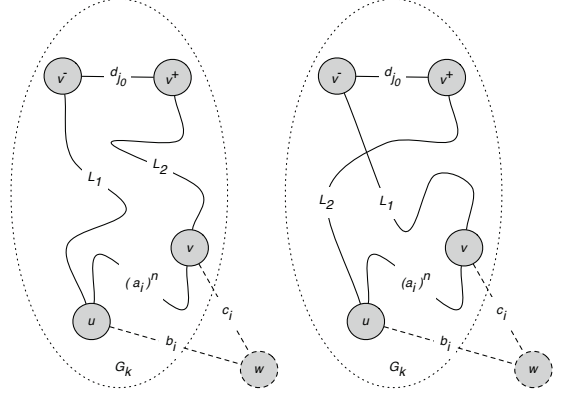


Figure 4: Graph G_{k+1}^j . Case 1 (left) and Case 2 (right).

path between v^- and v^+ labeled by $L_1, (a_i)^n, L_2$ for some $n \geq 0$. Hence, by the Induction Hypothesis,

$$d_{j_0} \circ_X L_1, (a_i)^n, L_2. \quad (6)$$

First, assume that $n = 0$. Thus,

$$d_{j_0} \circ_X L_1, L_2. \quad (7)$$

Note that since v^- and v^+ are two distinct vertices, the sum of the lengths of sequences L_1 and L_2 is not zero. Thus, by Lemma 7,

$$L_1, L_2 \circ_X L_1, b_i, L_2$$

and

$$L_1, b_i, L_2 \circ_X L_1, b_i, c_i, L_2.$$

Hence, Lemma 5, $L_1, L_2 \circ_X L_1, b_i, c_i, L_2$. By statement (7) and Lemma 5, $d_{j_0} \circ_X L_1, b_i, c_i, L_2$.

Second, suppose that $n > 0$ and consider the pattern

$$\begin{array}{c} b_i \\ a_i \\ c_i \end{array} \quad (8)$$

Recall that $a_i \triangleright_{c_i}^{b_i} \in X$. Thus, $a_i \circ_X b_i, c_i$. By Lemma 9, $(a_i)^n \circ_X a_i$. Hence, by Lemma 5, $(a_i)^n \circ_X b_i, c_i$. By Lemma 6, $L_1, (a_i)^n, L_2 \circ_X L_1, b_i, c_i, L_2$. Taking into account statement (6) and Lemma 5, $d_{j_0} \circ_X L_1, b_i, c_i, L_2$.

Case 2: Path π , in the direction from v^- to v^+ , first passes through edge (v, w) and then edge (w, u) . See Figure 4. In this case, instead of pattern (8), consider pattern

$$\begin{array}{c} c_i \\ a_i \\ b_i \end{array}$$

To show that $a_i \circ_X c_i, b_i$, notice that, by our assumption, $a_i \triangleright_{c_i}^{b_i} \in X$. Thus, by the Symmetry axiom, $X \vdash a_i \triangleright_{b_i}^{c_i}$. The rest of the proof is identical to Case 1. \square

LEMMA 12. $G^j \not\models d_j \triangleright_{f_j}^{e_j}$, for each $j \in J$.

PROOF. Assume that $G^j \models d_j \triangleright_{f_j}^{e_j}$. Note that $v^- \sim_{d_j} v^+$ by the definition of G_0^j . By Definition 7, there must be a vertex w such that $v^- \sim_{e_j} w$ and $w \sim_{f_j} v^+$. Thus, graph G^j contains a path π from v^- to v^+ labeled by the sequence

$(e_j)^n, (f_j)^m$ for some integers n and m . Since v^- and v^+ are different vertices, $n + m > 0$. By Lemma 11,

$$d_j \circ_X (e_j)^n, (f_j)^m.$$

By Lemma 9 and Lemma 5,

$$d_j \circ_X e_j, f_j.$$

By the Transitivity Axiom, $X \vdash d_j \triangleright_{f_j}^{e_j}$. By the maximality of X , $d_j \triangleright_{f_j}^{e_j} \in X$, which is a contradiction with $\{d_j \triangleright_{f_j}^{e_j}\}_{j \in J}$ being the set of all atomic formulas that do not belong to X . \square

Recall now that we left some flexibility in the choice of u , v , and i , when we defined extension G_{k+1}^j of graph G_k^j . We can use this flexibility as well as the countability of set I and the set of vertices in graph G^j to guarantee that, at some point, the expansion is applied to each possible triple u , v , and i such that $u \sim_{a_i} v$ in graph G^j . This will imply that the following statement is true:

PROPOSITION 1. *For any $i \in I$ and any vertices u and v in G^j such that $u \sim_{a_i} v$, there is a vertex w in G^j such that $u \sim_{b_i} w$ and $w \sim_{c_i} v$.*

Let graph G be the disjoint union of graphs $\{G_j\}_{j \in J}$.

LEMMA 13. *For any formula ψ ,*

$$G \models \psi \quad \text{iff} \quad \psi \in X.$$

PROOF. We use induction on the structural complexity of formula ψ . If ψ is \perp , then the statement is true due to the consistency of set X . Suppose now that ψ is formula $p \triangleright_r^q$. (\Rightarrow) Assume that $p \triangleright_r^q \notin X$. Thus $p \triangleright_r^q$ is $d_{j_0} \triangleright_{f_{j_0}}^{e_{j_0}}$ for some $j_0 \in J$. By Lemma 12, $G^{j_0} \not\models p \triangleright_r^q$. It means that there are vertices v and u in graph G^{j_0} such that $v \sim_p u$, but for any vertex w of G^{j_0} either $v \not\sim_q w$ or $w \not\sim_r u$. Since G is the disjoint union of graphs $\{G_j\}_{j \in J}$, the same is true for graph G . Therefore, $G \not\models p \triangleright_r^q$.

(\Leftarrow) Let $p \triangleright_r^q \in X$. Thus $p \triangleright_r^q$ is $a_{i_0} \triangleright_{c_{i_0}}^{b_{i_0}}$ for some $i_0 \in I$. Consider any vertices v and u in graph G such that $v \sim_p u$. Since G is the disjoint union of graphs $\{G_j\}_{j \in J}$, vertices v and u must belong to the same component G^{j_0} of the graph G . By Proposition 1, there is a vertex w in component G^{j_0} such that $v \sim_q w$ and $w \sim_r u$.

When formula ψ is an implication, the induction step of the proof follows trivially from the maximality and consistency of set X . \square

Finally, $\phi \notin X$ due to the consistency of set X . Thus, by Lemma 13, $G \not\models \phi$. This concludes the proof of Theorem 5.

6.3 Semantics of Secrets

In this section, we will use the graph completeness result from the previous section to prove the completeness of our logical system with respect to the original semantics of secrets from Definition 1.

THEOREM 6. *If $\mathcal{P} \models \phi$, for each protocol \mathcal{P} , then $\vdash \phi$.*

PROOF. Suppose that $\not\vdash \phi$. By Theorem 5, there is a graph G such that $G \not\models \phi$. We will define a protocol $\mathcal{P} = \langle V, R \rangle$ and prove that $\mathcal{P} \models \phi$. In the previous section, we defined relation \sim_a on the vertices of graph G for any label

a . Let $V(a)$ be the set of all equivalence classes of vertices of graph G with respect to equivalence relation \sim_a .

For any vertex v of graph G , define function r_v on labels of graph G in such way that $r_v(a)$ is the equivalence class of vertex v with respect to relation \sim_a . Let R be the set of such functions for all possible vertices v . This concludes the definition of the protocol \mathcal{P} .

LEMMA 14. *For any vertices u and v ,*

$$u \sim_a v \quad \text{iff} \quad r_u \equiv_a r_v$$

PROOF. Follows from the above definition of run $r_v(a)$. \square

LEMMA 15. *For any secret variables p, q, r ,*

$$\mathcal{P} \models p \triangleright_s^q \quad \text{iff} \quad G \models p \triangleright_s^q.$$

PROOF. Immediately follows from Theorem 2, Definition 7, and Lemma 14. \square

LEMMA 16. *For any formula ψ ,*

$$\mathcal{P} \models \psi \quad \text{iff} \quad G \models \psi.$$

PROOF. We use induction on the structural complexity of formula ψ . If ψ is \perp , then both statements are false. If ψ is $p \triangleright_s^q$, then the claim follows from Lemma 15. The case where ψ is an implication is trivial. \square

Note that $\mathcal{P} \not\models \phi$ by Lemma 16. This concludes the proof of Theorem 6.

7. CONCLUSION

In this paper, we studied the ternary relation $a \triangleright_c^b$ between secrets. Note that due to Lemma 2, this relation can be defined alternatively as

$$\forall r_1 \forall r_2 (r_1 \equiv_a r_2 \rightarrow \exists r (r_1 \equiv_b r \equiv_c r_2)).$$

In this alternate form, the definition of $a \triangleright_c^b$ is very similar to the definition of the embedded multivalued dependency $b \parallel_a c$:

$$\forall r_1 \forall r_2 (r_1 \equiv_a r_2 \rightarrow \exists r (r_1 \equiv_{a,b} r \equiv_{a,c} r_2)),$$

where $r' \equiv_{x,y} r''$ means that runs r' and r'' agree on secret variable x and secret variable y . It would be interesting to see if the techniques developed in this paper could be generalized to produce a complete axiomatization of the embedded multivalued dependency.

8. REFERENCES

- [1] W. W. Armstrong. Dependency structures of data base relationships. In *Information processing 74 (Proc. IFIP Congress, Stockholm, 1974)*, pages 580–583. North-Holland, Amsterdam, 1974.
- [2] Catriel Beeri, Ronald Fagin, and John H. Howard. A complete axiomatization for functional and multivalued dependencies in database relations. In *SIGMOD '77: Proceedings of the 1977 ACM SIGMOD international conference on Management of data*, pages 47–61, New York, NY, USA, 1977. ACM.
- [3] Michael Donders, Sara Miner More, and Pavel Naumov. Information flow on directed acyclic graphs. In L. Beklemishev and R. de Queiroz, editors, *Proceedings of 18th Workshop on Logic, Language, Information and Computation (Philadelphia, United States)*, pages 95–109. Springer, 2011.

- [4] Hector Garcia-Molina, Jeffrey Ullman, and Jennifer Widom. *Database Systems: The Complete Book*. Prentice-Hall, second edition, 2009.
- [5] Dan Geiger, Azaria Paz, and Judea Pearl. Axioms and algorithms for inferences involving probabilistic independence. *Inform. and Comput.*, 91(1):128–141, 1991.
- [6] Joseph Y. Halpern and Kevin R. O’Neill. Secrecy in multiagent systems. *ACM Trans. Inf. Syst. Secur.*, 12(1):1–47, 2008.
- [7] Christian Herrmann. On the undecidability of implications between embedded multivalued database dependencies. *Inf. Comput.*, 122(2):221–235, 1995.
- [8] Christian Herrmann. Corrigendum to “on the undecidability of implications between embedded multivalued database dependencies” [inform. and comput. 122(1995) 221-235]. *Inf. Comput.*, 204(12):1847–1851, 2006.
- [9] Robert Kelvey, Sara Miner More, Pavel Naumov, and Benjamin Sapp. Independence and functional dependence relations on secrets. In *Proceedings of 12th International Conference on the Principles of Knowledge Representation and Reasoning (Toronto, 2010)*, pages 528–533. AAAI, 2010.
- [10] Sara Miner More and Pavel Naumov. On interdependence of secrets in collaboration networks. In *Proceedings of 12th Conference on Theoretical Aspects of Rationality and Knowledge (Stanford University, 2009)*, pages 208–217, 2009.
- [11] Sara Miner More and Pavel Naumov. Hypergraphs of multiparty secrets. In *11th International Workshop on Computational Logic in Multi-Agent Systems CLIMA XI (Lisbon, Portugal), LNAI 6245*, pages 15–32. Springer, 2010.
- [12] Sara Miner More and Pavel Naumov. The functional dependence relation on hypergraphs of secrets. In *12th International Workshop on Computational Logic in Multi-Agent Systems CLIMA XI (Barcelona, Spain)*. Springer, 2011. (to appear).
- [13] Milan Studený. 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*, pages 377–396. Kluwer, 1990.
- [14] David Sutherland. A model of information. In *Proceedings of Ninth National Computer Security Conference*, pages 175–183, 1986.