A note on the correspondences among entail relations, rough set dependencies, and logical consequence

Theoretical Note

Ivo Düntsch*
School of Information and Software Engineering
University of Ulster
Newtownabbey, BT 37 0QB, N.Ireland
I.Duentsch@ulst.ac.uk

Günther Gediga*
FB Psychologie / Methodenlehre
Universität Osnabrück
49069 Osnabrück, Germany
Guenther@Gediga.de

Abstract

In this note, we report that entail relations defined in the context of knowledge spaces are equivalent to the dependence relations of rough set data analysis and Tarski's consequence relation of monotone logic. We also discuss the connection between these and related structures, and give an example of their application from the area of computer assisted language learning.

1 Introduction

Several constructions based on finite nonempty sets such as *knowledge spaces* [3, 9], *dependency spaces* [15], *consequence relations* [23], *formal concept analysis* (FCA) [12, 24], and others are sometimes viewed to be "somehow the same", because they use similar systems of axioms and similar operators, yet, researchers in the respective areas are not always aware of developments in the related fields. In knowing the relationships among these systems, one will arrive at a more general and parsimonious theoretical framework, and one will obtain the chance of fruitfully applying ideas from one research field to the others.

In this note, we shall present some simple but quite general observations which we hope will help clarify the essence of the connections among these theoretical approaches.

2 Mathematical background

Throughout this note, we suppose that Ω is a nonempty set. We assume the reader is familiar with the basic definitions and properties of lattice– and order theory – suitable references being [2] or [13] –,

*The ordering of authors is alphabetical, and equal authorship is implied.

and we shall just recall a few concepts. A *sup–semilattice* is a partially ordered set $\langle L, \leq \rangle$, in which the supremum $a \vee b$ exists for all $a, b \in L$. In other words, for all $a, b \in S$ there is a smallest $c \in L$ such that $a \leq c,\ b \leq c$.

A *closure system* is a family of subsets of Ω which is closed under intersection and contains $\emptyset$ and Ω, and a *knowledge space* is a family of subsets of Ω which is closed under union, and contains $\emptyset$ and Ω. A *closure operator on* Ω is a mapping $c : \mathcal{P}(\Omega) \to \mathcal{P}(\Omega)$ for which

1. $A \subseteq c(A)$,
2. $c(A) = c(c(A))$,
3. $A \subseteq B$ implies $c(A) \subseteq c(B)$,
4. $c(\emptyset) = \emptyset$.

Suppose that $\mathcal{L} = \langle L, \vee \rangle$ is a sup – semilattice. A *congruence* on $\mathcal{L}$ is an equivalence relation θ on L which satisfies the substitution property

$$a\theta b \text{ and } c\theta d \text{ imply } (a \vee c)\theta(b \vee d).$$

It is well known that each class K of a congruence θ is a subsemilattice of $\mathcal{L}$, and thus it has a maximum, written as $\max K$, if $\mathcal{L}$ is finite [see e.g. 13, p. 20].

Consider the following conditions on a relation $R \subseteq \mathcal{P}(\Omega) \times \mathcal{P}(\Omega)$: Let $X, Y, Z \subseteq \Omega$. Then,

$$\text{If } X \supseteq Y, \text{ then } \langle X, Y \rangle \in R, \tag{2.1}$$

$$\text{If } \langle X, Y \rangle, \langle Y, Z \rangle \in R, \text{ then } \langle X, Z \rangle \in R, \tag{2.2}$$

$$\text{If } \langle X, Y_i \rangle \in R \text{ for all } i \in I, \text{ then } \langle X, \bigcup_{i \in I} Y_i \rangle \in R. \tag{2.3}$$

Such a relation R determines a congruence $K(R)$ on the semilattice $\langle \mathcal{P}(\Omega), \cup \rangle$ via the assignment

$$K(R) \stackrel{\text{def}}{=} R \cap R^{-1},$$

cf [17]. This assignment is one-one as the following shows:

Lemma 2.1. *Let* $R, S \subseteq \mathcal{P}(\Omega) \times \mathcal{P}(\Omega)$ *satisfy* (2.1) – (2.3) *and suppose that* $K(R) = K(S)$*; then,* $R = S$.

Proof. By symmetry it suffices to show $R \subseteq S$; thus, let $\langle X, Y \rangle \in R$. Then, $\langle X, X \cup Y \rangle \in R$ follows from $\langle X, X \rangle \in R$ by (2.1), and (2.3). Again from (2.1) we have $\langle X \cup Y, X \rangle \in R$, and thus $\langle X, X \cup Y \rangle \in R \cap R^{-1} = S \cap S^{-1}$. We use (2.1) together with (2.2) once more to conclude $\langle X, Y \rangle \in S$. □

Conversely, it is proved in [17] that every congruence ψ on $\langle \mathcal{P}(\Omega), \cup \rangle$ determines a relation R_ψ on $\mathcal{P}(\Omega)$ which satisfies (2.1) – (2.3) via the assignment

$$\psi \mapsto \{\langle X, Y \rangle : \max(\psi Y) \subseteq \max(\psi X)\},$$

where for $Z \subseteq \Omega$, ψZ is the congruence class of Z with respect to ψ, and the maximum, which exists because Ω is finite, is taken with respect to $\subseteq$.

We remark in passing that Novotný [17] has shown that the congruences on $\langle \mathcal{P}(\Omega), \cup \rangle$ are in a bijective correspondence to the closure operators on $\langle \mathcal{P}(\Omega), \subseteq \rangle$, which in turn are in a bijective correspondence to the closure systems on Ω [cf 1].

In another direction, a relation with the properties (2.1) – (2.3) is determined by its restriction to $\mathcal{P}(\Omega) \times \{\{u\} : u \in \Omega\}$: If R is a binary relation on $\mathcal{P}(\Omega)$, we let

$$\mathrm{sng}(R) \stackrel{\mathrm{def}}{=} \{\langle X, y \rangle : \langle X, \{y\} \rangle \in R\} \subseteq \mathcal{P}(\Omega) \times \Omega.$$

Conversely, if $T \subseteq \mathcal{P}(\Omega) \times \Omega$, we set

$$\mathrm{mlt}(T) \stackrel{\mathrm{def}}{=} \{\langle X, Y \rangle : \langle X, y \rangle \in T \text{ for all } y \in Y\}.$$

The following was noted in [14], 4.2.(iii):

Proposition 2.2. *Let $T \subseteq \mathcal{P}(\Omega) \times \Omega$. Then, $\mathrm{mlt}(T) = R$ for some relation R on $\mathcal{P}(\Omega)$ which satisfies* (2.1) – (2.3)*, if and only if for all $X, Y \subseteq \Omega,\ x, y \in \Omega$,*

(2.4) $\langle X, x \rangle \in T,$ *if* $x \in X,$

(2.5) *If* $\langle X, y \rangle \in T$ *for all* $y \in Y,$ *and* $\langle Y, x \rangle \in T,$ *then* $\langle X, x \rangle \in T.$

Conversely, if R is a relation which satisfies (2.1) – (2.3)*, then $\mathrm{sng}(R)$ satisfies* (2.4) *and* (2.5)*, and* $\mathrm{mlt}(\mathrm{sng}(R)) = R$. □

It is not hard to see that conditions (2.4) and (2.5) are equivalent to

(2.6) If $x \in X,$ then $\langle X, x \rangle \in T$ (Reflexivity)

(2.7) If $\langle X, x \rangle \in T,$ then $\langle X \cup Y, x \rangle \in T$ (Monotony)

(2.8) If $\langle X, y \rangle \in T$ for all $y \in Y$ and $\langle X \cup Y, x \rangle \in T,$ then $\langle X, x \rangle \in T$ (Cut).

3 Applications

In this section we shall exhibit some contexts which use relations which satisfy (2.1) – (2.3), respectively the equivalent conditions (2.6) – (2.8).

Context 1. Entail relations [14]:

In the theory of knowledge spaces [3, 9], entail relations are used to compute an unknown knowledge space by querying an expert.

Let Ω be a set of problems, and $\mathfrak{K}$ be a family of subsets of Ω, containing $\emptyset$ and Ω. The pair $\langle \Omega, \mathfrak{K} \rangle$ is called a *knowledge structure*, and the elements of $\mathfrak{K}$ are interpreted as the knowledge states of subjects. Define a relation $R_{\mathfrak{K}}$ on $\mathcal{P}(\Omega)$ by setting

(3.1) $$\langle A, B \rangle \in R_{\mathfrak{K}} \overset{\text{def}}{\iff} B \cap K \neq \emptyset \text{ implies } A \cap K \neq \emptyset \text{ for all } K \in \mathfrak{K}.$$

The interpretation of $\langle A, B \rangle \in R_{\mathfrak{K}}$ is that if subjects master some problem in B they also master at least one problem from A. The relation $R_{\mathfrak{K}}$ satisfies (2.1) – (2.3), and [14] call a binary relation on $\mathcal{P}(\Omega)$ which satisfies these conditions an *entail relation*.

It may be worthy to mention that the assignment $\mathfrak{K} \mapsto R_{\mathfrak{K}}$ is just one possibility to obtain an entail relation from a family of subsets of Ω. In [4] we describe all Galois connections between knowledge structures and binary relations on $\mathcal{P}(\Omega)$ in which the Galois closed relations are exactly the entail relations.

Context 2. Dependence relations [15]:

The method of *rough set data analysis* [19] is based on the fact that in many situations it is only possible to distinguish items of a domain up to a set of features or attribute values. Knowledge representation is done via *information systems* of the form $\mathcal{I} = \langle U, \Omega, (V_q, f_q)_{q \in \Omega} \rangle$, where

- U is a finite set of objects,
- Ω is a finite set of attributes,
- For each $q \in \Omega$, V_q is a set of attribute values for q,
- Each $f_q : U \to V_q$ is an *information function*.

If $V_q = \{0, 1\}$ for all $q \in \Omega$ we call $\mathcal{I}$ *binary*. It may be worthy to note that binary information systems can be expressed by contexts in the sense of [24].

For each $Q \subseteq \Omega$ we define an equivalence relation θ_Q on U by

$$x \equiv_{\theta_Q} y \overset{\text{def}}{\iff} f_q(x) = f_q(y) \text{ for all } q \in Q.$$

If $P, Q \subseteq \Omega$, we say that P *is dependent on* Q, and write $Q \to P$, if $\theta_Q \subseteq \theta_P$ [cf. 20].

The relation $\to \subseteq \mathcal{P}(\Omega) \times \mathcal{P}(\Omega)$ is called the *dependence relation* of $\mathcal{I}$, and it is shown in [15] that it satisfies (2.1) – (2.3). More generally, Novotný [17] calls a relation on $\mathcal{P}(\Omega)$ which satisfies these conditions a *dependence relation*. Algebraic properties of dependence relations can be found in [5, 17], and an up to date introduction to rough set data analysis is [8][1].

A particular kind of dependence are those of the form $Q \to \Omega$. If Q is minimal with respect to this property, then it is called a *reduct (of* Ω*)*. Reducts are a convenient way of reducing the amount of information and are helpful in discovering relationships among the attributes.

In the presence of an additional decision attribute d, one can consider dependencies of the form $Q \to d$, where $Q \subseteq \Omega$. If Q is a reduct with respect to d, i.e. if Q is minimal with respect to the property $Q \to d$, then the information given by the partition of the object set U by the attributes in Q completely determines the classification of the elements of U with respect to d.

In general, dependencies of the form $Q \to d$ predict less than dependencies of the form $Q \to \Omega$; however, for many applications it is sufficient to consider only one decision attribute, and further analysis is not necessary.

Knowledge structures may be viewed as binary information systems: More generally, if U is a group of subjects, Ω a set of problems, and $s : U \to \mathcal{P}(\Omega)$ the function which assigns to each subject x the set $s(x)$ of problems x is capable of solving, then we obtain a binary information system $\langle U, \Omega, \langle V_q, f_q \rangle_{q \in \Omega} \rangle$ by setting $V_q \stackrel{\text{def}}{=} \{0, 1\}$ for each $q \in \Omega$, and

$$f_q(x) = \begin{cases} 1, & \text{if } q \in s(x), \\ 0, & \text{otherwise,} \end{cases}$$

to form an information system, the rows of which correspond to the characteristic functions of the observed knowledge states via s.

Context 3. Logical consequence [23]:

Let Fml be the set of formulas of a logic $\mathcal{L}$. A *consequence relation* $\vdash$ is a subset of $\mathcal{P}(\mathsf{Fml}) \times \mathsf{Fml}$ which satisfies (2.6), (2.7), and (2.8), and, more generally, we call a relation $T \subseteq \mathcal{P}(\Omega) \times \Omega$ which satisfies these conditions a *consequence relation* on Ω.

With each consequence relation T – and thus with each entail relation by Proposition 2.2 – on Ω one can associate an operator f on $\mathcal{P}(\Omega)$ by

$$f(X) \stackrel{\text{def}}{=} \{y \in \Omega : \langle X, y \rangle \in T\}.$$

The conditions on T ensure that f is a closure operator (which is called a *consequence operator* in [23]). It is mentioned in [11] that Tarski and Scott have shown that each closure operator on a set Ω is induced by the consequence relation of a logic.

For completeness, we mention the following related construction: Davey & Priestley [2] call a structure $\mathbf{A} = \langle A, Con, \vdash \rangle$ an *information system*,[2] if

[1] Technical report version available from `htpp://www.infj.ulst.ac.uk/~cccz23/papers/rsda.html`.

[2] Hilary Priestley has informed us that this notion is due to Dana Scott.

1. A is a set of *tokens*.

2. Con is a nonempty set of finite subsets of A which satisfy

   (IS1) $Y \in Con,\ Z \subseteq Y \Rightarrow Z \in Con$,

   (IS2) Con contains all singleton subsets of A.

3. $\vdash$ is a subset of $Con \times A$ for which

   (a) If $Y \in Con,\ a \in A$ and $Y \vdash a$, then $Y \cup \{a\} \in Con$,

   (b) If $Y \in Con$ and $a \in Y$, then $Y \vdash a$,

   (c) If $Y, Z \in Con,\ a \in A$ such that $Y \vdash b$ for all $b \in Z$ and $Z \vdash a$, then $Y \vdash a$.

An area which is related to all of these contexts is *formal concept analysis* (FCA) which has been developed by Rudolf Wille and his group since the early 1980s [24]. Its basic structure are triples $\langle G, M, I\rangle$, where G is a set of objects, M a set of attributes, and $I \subseteq G \times M$ a relation in which gIm is interpreted as "g has attribute m". A *concept* (of the context I) is a pair $\langle A, B\rangle \in \mathcal{P}(\mathcal{G}) \times \mathcal{P}(\mathcal{M})$ such that

$$(\forall g \in G)[g \in A \iff (\forall m \in B)\langle g, m\rangle \in I],$$
$$(\forall m \in M)[m \in B \iff (\forall g \in A)\langle g, m\rangle \in I].$$

These notions of "context" and "concept" are so general – in fact, concepts correspond to Galois correspondences – that everything mentioned thus far can be expressed in terms of contexts, resp. concepts.

The relationship between FCA and dependence spaces and knowledge structures is well known, and we invite the reader to consult [2, 12] for an introduction to FCA, [16] for the relationship of rough set dependencies and contexts, and [21] for the connections of FCA to knowledge structures.

It may also be interesting to consult [22] and [10] for an earlier approach to determine concepts (using the term "monothetic groups") in the additional context of taxonomy.

4 Discussion

It is possible to discover connections among groups of problems by studying the dependence relation, and entailment of problem solving can be directly applied to relate groups of subjects according to their problem solving behaviour: If Ω is a set of problems, and $Q \subseteq \Omega,\ p \in \Omega$, the dependence rule $Q \to p$ tells us that whenever two subjects agree in their (non-)ability to solve problems in Q, then they will also agree in their (non-) ability to solve p.

Knowledge spaces are often used to describe results of a test, avoiding psychometric theories to describe the results in "underlying dimension(s)". The aim of testing is to discriminate between groups of subjects in an interesting entity $\mathcal{E}$ (e.g. intelligence tests should discriminate subjects' intelligence), and assume that $\mathcal{E}$ is measurable by a scaling function E. In this situation, rough set dependency analysis may be applied to find those $Q \subseteq \Omega$ which are sufficient to predict the outcome E.

As an example, we should like to present the following situation adapted from [7]: The test data were obtained in a *German as a foreign language* course at beginner-level, which introduced learners to idioms and vocabulary of various situations, using basic grammar and sentence structure. A range of exercises was given which aimed at developing and enforcing the language acquisition skills. 17 subjects were taught by a tutor using a computer assisted language learning package (Group CT), seven were taught by a tutor only (Group T).

A test was given after 6 hours of lectures, and contained 18 questions listed in Table 1.

The usual procedure is to find the knowledge structure belonging to this table, group the subjects, and then compare the groups. An alternative way, based on the rough set philosophy is to ask right away, which problems suffice to determine membership in one of the two groups. Observe that the emphasis is in the problem solving abilities of groups of subjects, and not on the relationship among the problems. Using dependency analysis, we have found that there are 212 reducts which determine in which group a subject was taught. This shows that there is a large range of substitution possibilities among the problems. Further analysis showed that the rules given by the reduct containing problems A_2, A_5, A_{10} and A_{11} are significant [6], and that the expected prediction quality of these attributes using the jackknife method is 82%.

The aim of this note was to give an example how, for a given set of data, equivalent approaches to its analysis such as knowledge structure theory and rough set dependency analysis can lead to different insights. Hopefully, this will aid better communication among different areas of application, such as psychology, artificial intelligence, and applied logic. Moreover, we have shown that, by borrowing from the field of dependence relations of rough set data analysis, it is possible to investigate a problem of another area, namely, knowledge structures in psychology, in a way which gives additional insight into the problem under consideration.

References

[1] Birkhoff, G. (1948). Lattice Theory, vol. 25 of *Am. Math. Soc. Colloquium Publications*. Providence: AMS, 2nd Edn.

[2] Davey, B. A. & Priestley, H. A. (1990). Introduction to Lattices and Order. Cambridge University Press.

[3] Doignon, J. P. & Falmagne, J. C. (1999). Knowledge spaces. Berlin: Springer–Verlag.

[4] Düntsch, I. & Gediga, G. (1996). On query procedures to build knowledge structures. *J. Math. Psych.*, **40**, 160–168.

Table 1: Test questions

A_1	Wie . . . es Ihnen? (a) trinkt (b) ist (c) kommt (d) geht
A_2	Wo . . . Sie? (a) kommen (b) nehmen (c) wohnen (d) heissen
A_3	. . . kommen Sie? (a) wo (b) wer (c) wie (d) woher
A_4	Verzeihung, wie . . . Ihr Name? (a) bist (b) ist (c) sein (d) sind
A_5	Wie . . . das auf deutsch? (a) heisse (b) heissen (c) heisst (d) heiss
A_6	. . . du auch Englisch? (a) sprichst (b) spricht (c) spreche (d) sprechen
A_7	Siebzehn (a) 27 (b) 6 (c) 17 (d) 16
A_8	Dreiundzwanzig (a) 30 (b) 32 (c) 20 (d) 23
A_9	Guten M . . .
A_{10}	Verzeihung, w . . . ist Ihr Name?
A_{11}	Wie s . . . man das?
A_{12}	Wo w . . . Sie?
A_{13}	Konjugieren Sie bitte “sein”.
A_{14}	Konjugieren Sie bitte “wohnen”.
A_{15}	Antworten Sie oder fragen Sie: Das ist Frau Lim. . . . ?
A_{16}	Antworten Sie oder fragen Sie: Sie kommt aus Korea. . . . ?
A_{17}	Antworten Sie oder fragen Sie: Wie heissen Sie? . . .
A_{18}	Antworten Sie oder fragen Sie: Wo arbeiten Sie? . . .

[5] Düntsch, I. & Gediga, G. (1997a). Algebraic aspects of attribute dependencies in information systems. *Fundamenta Informaticae*, **29**, 119–133.

[6] Düntsch, I. & Gediga, G. (1997b). Statistical evaluation of rough set dependency analysis. *International Journal of Human–Computer Studies*, **46**, 589–604.

[7] Düntsch, I. & Gediga, G. (1998). Knowledge structures and their applications in CALL systems. In S. Jager, J. Nerbonne & A. van Essen (Eds.), *Language teaching and language technology*, 177–186. Swets & Zeitlinger.

[8] Düntsch, I. & Gediga, G. (2000). Rough set data analysis. In *Encyclopedia of Computer Science*

Table 2: Rules of the reduct A_2, A_5, A_{10}, A_{11}

$$A_2 = 1, A_5 = 0, A_{11} = 0 \Rightarrow \text{ Class } = CT$$
$$A_2 = 1, A_{10} = 1, A_{11} = 0 \Rightarrow \text{ Class } = CT$$
$$A_2 = 1, A_5 = 1, A_{10} = 1 \Rightarrow \text{ Class } = CT$$
$$A_5 = 1, A_{10} = 1 A_{11} = 0 \Rightarrow \text{ Class } = CT$$
$$A_5 = 0, A_{10} = 0 \Rightarrow \text{ Class } = CT$$
$$A_2 = 0, A_5 = 1, A_{11} = 0 \Rightarrow \text{ Class } = CT$$
$$A_2 = 0, A_{10} = 0 \Rightarrow \text{ Class } = CT$$
$$A_2 = 0, A_5 = 0 \Rightarrow \text{ Class } = T$$
$$A_2 = 0, A_{10} = 1 A_1 1 = 1 \Rightarrow \text{ Class } = T$$
$$A_2 = 1, A_5 = 1, A_{10} = 0 \Rightarrow \text{ Class } = T$$
$$A_2 = 1, A_{10} = 0, A_{11} = 1 \Rightarrow \text{ Class } = T$$
$$A_5 = 0, A_{11} = 1 \Rightarrow \text{ Class } = T$$
$$A_5 = 1, A_{10} = 0 A_{11} = 0 \Rightarrow \text{ Class } = T$$

and Technology. Marcel Dekker. To appear.

[9] Falmagne, J.-C., Koppen, M., Villano, M., Doignon, J.-P. & Johannesen, J. (1990). Introduction to knowledge spaces: How to build, test and search them. *Psychological Review*, **97**, 202–234.

[10] Farris, J. S. (1978). An efficient method for finding monothetic groups. *Systematic Zoology*, **27**, 468–472.

[11] Gabbay, D. M. (1985). Theoretical foundations for non–monotonic reasoning in expert systems. In K. R. Apt (Ed.), *Logics and Models of Concurrent Systems*, vol. F13 of *NATO Advanced Studies Institute*, 439–457. Berlin: Springer.

[12] Ganter, B. & Wille, R. (1999). Formal concept analysis: Mathematical foundations. Berlin: Springer–Verlag.

[13] Grätzer, G. (1978). General Lattice Theory. Basel: Birkhäuser.

[14] Koppen, M. & Doignon, J.-P. (1990). How to build a knowledge space by querying an expert. *J. Math. Psych.*, **34**, 311–331.

[15] Novotný, J. & Novotný, M. (1992). Notes on the algebraic approach to dependence in information systems. *Fundamenta Informaticae*, **16**, 263–273.

[16] Novotný, J. & Novotný, M. (1993). On dependence in Wille's contexts. *Fundamenta Informaticae*, **19**, 343–353.

[17] Novotný, M. (1997). Dependence spaces of information systems. In [18], 193–246.

[18] Orłowska, E. (Ed.) (1997). Incomplete Information – Rough Set Analysis. Heidelberg: Physica – Verlag.

[19] Pawlak, Z. (1982). Rough sets. *Internat. J. Comput. Inform. Sci.*, **11**, 341–356.

[20] Pawlak, Z. & Rauszer, C. (1985). Dependency of attributes in information systems. *Bull. Polish Acad. Sci. Math.*, **9–10**, 551–559.

[21] Rusch, A. & Wille, R. (1996). Knowledge spaces and formal concept analysis. In H.-H. Bock & W. Polasek (Eds.), *Data Analysis and Information Systems*, Studies in Classification, Data Analysis, and Knowledge Organization, 427–436, Berlin, Germany. Springer–Verlag.

[22] Sharrock, G. & Felsenstein, J. (1975). Finding all monothetic subsets of a taxonomic group. *Systematic Zoology*, **24**, 373–377.

[23] Tarski, A. (1930). Fundamentale Begriffe der Methodologie der deduktiven Wissenschaften. *Monatsh. Math. Phys.*, **37**, 361–404.

[24] Wille, R. (1982). Restructuring lattice theory: an approach based on hierarchies of concepts. In I. Rival (Ed.), *Ordered sets*, vol. 83 of *NATO Advanced Studies Institute*, 445–470. Dordrecht: Reidel.