INFORMATION AND COMPUTATION 122, 221–235 (1995)

On the Undecidability of Implications between Embedded Multivalued Database Dependencies

CHRISTIAN HERRMANN

Technische Hochschule Darmstadt, Fachbereich Mathematik, Schlossgartenstrasse 7, D61 Darmstadt, Germany

The implication and finite implication problems for embedded multivalued database dependencies are both shown to be algorithmically undecidable. The proof is by an interpretation of semigroup word problems via systems of permuting equivalence relations into database dependencies. In contrast, it is shown that for each fixed premise H one has a decision procedure for implications $H \Rightarrow F$.

1. INTRODUCTION

In the context of lossless decompositions, multivalued database dependencies (mvd's) have been introduced by Fagin (1977) and Zaniolo (1976) as a somehow weaker variant of functional dependencies (fd's). A sound and complete (also with respect to finite databases) axiomatization of mvd's and of mvd's and fd's together has been given by Beeri *et al.* (1977). In particular, the implication problem and the finite implication problem coincide and have an algorithmic solution. Here, the (finite) implication problem for a given class of database dependencies refers to a procedure deciding for any implication $F_1 \wedge \cdots \wedge F_n \Rightarrow F$, with dependencies F, F_i from the class, whether it is valid in all (finite) databases—where the attribute set of a database (which is commonly supposed to be finite) includes all the attributes occurring in the implication. The algorithm is required to be uniform for all finite attribute sets and antecedents $F_1 \wedge \cdots \wedge F_n$. In contrast, the restricted (finite) implication problem asks whether for each fixed finite set $F_1, \ldots, F_n$ of dependencies the set of all dependencies F with $F_1 \wedge \cdots \wedge F_n \Rightarrow F$ valid in all (finite) databases is a recursive set. Implication problems are closely related to first order logic decision problems. The problems to be considered here can indeed be viewed as instances of such.

Departing from a rather direct interpretation of groups or semigroups into relations, unsolvability has been shown for various classes of dependencies, in particular for the important class of embedded template dependencies (Beeri and Vardi, 1981; Chandra *et al.*, 1981) and even the subclass of projected join dependencies (Gurevich and Lewis, 1982; Vardi, 1984). For typed template dependencies, Vardi (1984) obtained the unsolvability of the restricted (finite) implication problem.

Embedded multivalued dependencies (emvd's) are mvd's on a projection and a special case of projected join dependencies. Emvd's being so closely related to the natural join, a positive solution of the implication problem would have been appreciated. The nonexistence of an axiomatization involving only a bounded number of attribute set variables has been shown by Parker and Parsaye-Ghomi (1980) and Sagiv and Walecka (1982). Sagiv *et al.* (1981) pointed out that no translation into propositional logic is possible. Day (1993) suggested an analogy with word problems for modular lattices.

The crucial reduction, providing further evidence towards unsolvability, was the result of Beeri and Vardi (1981) that a solution of the (finite) implication problem for emvd's would also solve the corresponding problem for implications of the form "a conjunction of emvd's and fd's implies an emvd." Based on this and the unsolvability of the word problem for (finite) semigroups (Markov, 1947; Post, 1947; resp. Gurevich, 1966), one gets the following.

THEOREM 1. *Implication and finite implication are algorithmically undecidable for embedded multivalued dependencies.*

COROLLARY 2. *Implication and finite implication do not coincide for embedded multivalued dependencies. There is no sound and complete recursive axiomatization of embedded multivalued dependencies with respect to finite databases.*

THEOREM 3. *The restricted (finite) implication problem is solvable for the class of functional and embedded multivalued dependencies.*

The undecidability proof relies on an interpretation of semigroups into certain systems of equivalence relations on a set, translating semigroup relations first into meet and sum relationships between subgroups of abelian groups and then into meet and product relationships between equivalence relations (Section 6). These in turn can be captured by fd's and emvd's, associating databases with systems of equivalences, canonically (Section 4); cf. Gurevich and Lewis (1982), Cosmadakis *et al.* (1986) and Rauszer (1987).

0890-5401/95 $12.00

The striking fact is that associativity does not come as a law encoded into dependencies of sufficient expressive power (which emvd's are not). Rather, it is a by-product of a coordinatization result (Section 7) for equivalence class geometries requiring only a finite number of meet and product relations for a finite system of "generating" equivalence relations which include a "coordinate system." It is this result which connects the rather general database methods of Section 4 with the special algebraic ones of Section 6 to provide the undecidability proof in Section 8.

As in Gurevich and Lewis (1982), the universal Horn theory of semigroups enters in a form where semigroups are conceived as ternary relational structures. In this version, the restricted decision problem has a "solution"—there is only one Horn formula to be produced from a given antecedent. So it comes as no surprise that the restricted problem is solvable for dependencies (Section 5). A primer on equivalence relations is given in Section 3; a general concept of interpretation of one theory in another is explained in Section 2. Finally, in Section 9, we discuss related results in the theory of lattices and relation algebras.

Quite a few techniques are borrowed from unsolvability results in modular lattices (Hutchinson, 1973; Lipshitz, 1974). In particular, von Neumann's (1960) concept of frame is adapted to systems of equivalence relations to get a coordinatization. Yet, the setting is definitely not a lattice theoretic one. For more comments on this, see Section 9. Most of the material is folklore in one context or other; proofs are included to keep the paper self-contained.

The author is grateful to Alan Day, Achim Jung, Bernhard Thalheim, and the referee for helpful comments.

2. THE METHOD OF INTERPRETATIONS

This method has been developed mainly by Tarski; cf. Shoenfield (1967). Here, a rather elementary version is needed. The framework is that of a class $\mathscr{C}$ of structures and a language L related by a notion $\models$ of validity: $S \models F$ to be read as "the sentence F is valid in the structure S" or "S satisfies F." If S does not satisfy F we also say that "S falsifies F." The language L is built from atomic sentences each involving finitely many of countably many individual constants resp. relation symbols. Validity has to be defined for the atomic sentences in such a manner that only the actually occurring constants matter. The language is supposed to consist of certain propositional combinations of the atomic sentences and the concept of validity is extended, accordingly. Thus, we may think of a structure as having only finitely many constants interpreted, including those occurring in the sentences under consideration. Saying that a structure satisfies or falsifies a sentence always includes that all the constants occurring in the sentence are interpreted in the structure. The sentences to be considered will be mostly of the form $H \Rightarrow F$ for which falsification means that H but not F is valid in the structure.

Now, a solution of the (finite) validity problem for L and $\mathscr{C}$ consists in an algorithm deciding for each sentence in L whether it is valid in all (finite) structures in $\mathscr{C}$.

An *interpretation* of $\mathscr{C}_1$ and L_1 into $\mathscr{C}_2$ and L_2 associates effectively with each F in L_1 a sentence F' in L_2 such that for each structure M in $\mathscr{C}_1$ falsifying F there is structure M' in $\mathscr{C}_2$ falsifying F' and, conversely, for each structure N in $\mathscr{C}_2$ falsifying F' a structure N' in $\mathscr{C}_1$ falsifying F. Here, finite structures are supposed to correspond to finite ones. Therefore, F is (finitely) valid in $\mathscr{C}_1$ if and only if F' is (finitely) valid in $\mathscr{C}_2$ and if the (finite) validity problem is unsolvable for L_1 and $\mathscr{C}_1$ then the same takes place for L_2 and $\mathscr{C}_2$.

If there is also an interpretation of $\mathscr{C}_2$ and L_2 into $\mathscr{C}_1$ and L_1, then we speak of an *equivalence*.

3. EQUIVALENCE RELATIONS

We outline some basics of the theory of equivalence relations; cf. Dubreil and Dubreil-Jacotin (1939), Ore (1942), and Cohn (1981).

Let $\Pi(E)$ denote the system of all equivalence relations on the set E—equivalently, one may think of the associated partitions. For α in $\Pi(E)$ and a, b in E we write $a\alpha b$ if and only if a is related to b under α (often written as $(a, b) \in \alpha$). Let $a/\alpha = \{b \in E \mid a\alpha b\}$ denote the α-class of a. Any such set will be called an (equivalence) *class* of α. The intersection $\gamma = \alpha \cap \beta$ is defined by: $a\gamma b$ if and only if $a\alpha b$ and $a\beta b$. For α, β in $\Pi(E)$ it is, of course, in $\Pi(E)$. α is finer than β, $\alpha \subseteq \beta$, if $a\alpha b$ implies $a\beta b$, in other words if $\alpha = \alpha \cap \beta$. We write $id = id_E$ for the identity relation and $\nabla = \nabla_E$ for the total relation.

Let Ξ_∞ be a countably infinite set of names (for equivalence relations). A *system of equivalences* is a pair (E, h) where E is a set and h a map from a finite set $\Xi \subseteq \Xi_\infty$ into $\Pi(E)$. In other terms, we consider relational structures with base set E and family h of binary relations under axioms requiring that these relations are equivalence relations. Let Π denote the class of all (E, h) and Π_0 the subclass consisting of those satisfying $\bigcap_{\xi \in \Xi} h(\xi) = id$.

An *isomorphism* between two systems (E, h) and (E', h') of equivalences (having the same domain Ξ) is given by a bijection $\phi: E \to E'$ such that for all $\xi \in \Xi$ the equivalence $h'(\xi)$ is the ϕ-image of $h(\xi)$.

For $E' \subseteq E$ we can form the *restriction* $(E', h \mid E')$ where $h \mid E'(\xi) = h(\xi) \cap E'^2$.

Given an equivalence relation μ on E the *quotient set* E/μ consists of all μ-classes a/μ, $a \in E$. For $\alpha \supseteq \mu$ in $\Pi(E)$ there is an equivalence relation α/μ on the quotient set E/μ such that two μ-classes are related if and only if they lie in a common α-class. In other words, for μ-classes $\tilde{a}$ and $\tilde{b}$ we have $\tilde{a}\,\alpha/\mu\,\tilde{b}$ if and only $a\alpha b$. So, for $\mu \subseteq \bigcap_{\xi \in \Xi} h(\xi)$ we can form the *quotient system* $(E, h)/\mu = (E/\mu, h/\mu)$ where $h/\mu(\xi) = h(\xi)/\mu$.

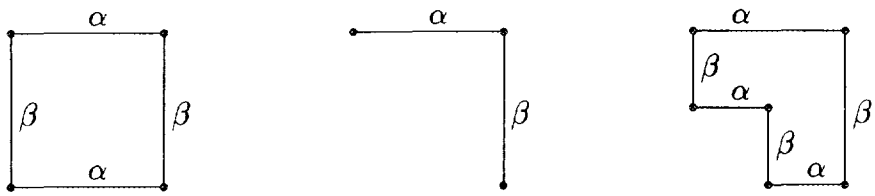

FIG. 1. Permutability.

In particular, with $\mu = \bigcap_{\xi \in \Xi} h(\xi)$ each (E, h) in Π has a *canonical quotient* in Π_0.

Considering a direct product $\prod_{i \in I} A_i$ let θ_i be the kernel of the projection π_i onto A_i and $\theta_i' = \bigcap_{j \neq i} \theta_j$, i.e. $s\theta_i t$ if and only if $s[i] = t[i]$ and $s\theta_i' t$ if and only if $s[j] = t[j]$ for all $j \neq i$.

If E is a subset of the direct product of sets A_ξ, $\xi \in \Xi$, and if $h(\xi) = \theta_\xi \mid E$, then we call (E, h) a *system of projection kernels* if the projections restricted to E are still onto: $\pi_\xi(E) = A_\xi$ for all ξ. One also speaks of a subdirect decomposition of E. Of course, such is in Π_0.

PROPOSITION 4. *Each system (E, h) of equivalences in Π_0 is isomorphic to a system of projection kernels.*

Proof. Let A_ξ the quotient set $E/h(\xi)$ of all $h(\xi)$-classes. Define ϕ from E into the direct product of the A_ξ by

$$\phi(e) = (e/h(\xi) \mid \xi \in \Xi).$$

Let E' the image of E under ϕ. Having (E, h) in Π_0 forces ϕ to be one-to-one. ∎

For two relations, $\alpha \circ \beta$ is the *relational product* (which may happen not to be an equivalence even if α, β are): $a(\alpha \circ \beta)\, b$ if and only if there is c with $a\alpha c$ and $c\beta b$. If α, β are reflexive then $\alpha, \beta \subseteq \alpha \circ \beta$; also, $\alpha \subseteq \beta$ implies $\beta = \alpha \circ \beta = \beta \circ \alpha$. Two equivalence relations α, β are said to *permute* if $\alpha \circ \beta = \beta \circ \alpha$; see Fig. 1.

PROPOSITION 5. (a) *α, β in $\Pi(E)$ permute if and only if $\alpha \circ \beta \in \Pi(E)$ if and only if $\beta \circ \alpha \in \Pi(E)$.*

(b) *The products formed from finitely many factors of a fixed set of pairwise permuting equivalences are again equivalences and give rise to a join semilattice; i.e., they satisfy the associative, commutative, and idempotency $(\alpha \circ \alpha = \alpha)$ laws, and $\alpha \subseteq \beta$ iff $\alpha \circ \beta = \beta$.*

(c) *Modular Law: For any α, β, γ in $\Pi(E)$*

$$\alpha \supseteq \gamma \text{ implies } \alpha \cap (\beta \circ \gamma) = (\alpha \cap \beta) \circ \gamma.$$

Proof. (a) The product operation, applied to any binary relations, is associative. Products of reflexive relations are reflexive. The transitivity of a reflexive relation γ amounts to $\gamma \circ \gamma = \gamma$. Now, for $\gamma = \alpha \circ \beta$, symmetry is equivalent to $\alpha \circ \beta = \beta \circ \alpha$ and implies $\gamma \circ \gamma = \alpha \circ \alpha \circ \beta \circ \beta = \alpha \circ \beta = \gamma$ if α and β were supposed to be equivalence relations.

(b) Follows, immediately.

(c) Let $a\alpha \cap (\beta \circ \gamma)\, b$, i.e., $a\alpha b$ and $a\beta c\gamma b$ for some c. Then $c\alpha b$ since $\alpha \supseteq \gamma$ and $a\alpha c$ since α is an equivalence relation. Thus $a\alpha \cap \beta c\gamma b$ which proves $\alpha \cap (\beta \circ \gamma) \subseteq (\alpha \cap \beta) \circ \gamma$. For the converse inclusion observe that $a\alpha \cap \beta c\gamma b$ implies $a\alpha b$ and $a\beta \circ \gamma b$, trivially. ∎

PROPOSITION 6. (a) *(E, α_1, α_2) is isomorphic to some $(A_1 \times A_2, \theta_1', \theta_2')$ if and only if $\alpha_1 \cap \alpha_2 = id$ and $\alpha_1 \circ \alpha_2 = \nabla$.*

(b) *$(E, \alpha_1, \alpha_2, \alpha_3)$ is isomorphic to some $(A_1 \times A_2 \times A_3, \theta_1', \theta_2', \theta_3')$ if and only if the α_i permute, $\alpha_1 \circ \alpha_2 \circ \alpha_3 = \nabla$ and $\alpha_i \cap (\alpha_j \circ \alpha_k) = id$ for any permutation i, j, k of $\{1, 2, 3\}$.*

Proof. In (b) let α_i' the product of the α_j, $j \neq i$. By the Modular Law, $\alpha_i' \cap \alpha_j' = \alpha_k$ for distinct i, j, k. So, the intersection of all α_i''s is id and Proposition 4 yields a subdirect decomposition into A_i's and one may assume that α_i is the restriction of θ_i' to E, whence $\alpha_i' = \theta_i \mid E$. Let $a_1 \in A_1$, $b_2 \in A_2$, and $c_3 \in A_3$ be given. Since the projection maps are onto, there are (a_1, a_2, a_3), (b_1, b_2, b_3), and (c_1, c_2, c_3) in E. Due to $\alpha_3' \circ \alpha_2' = \nabla$ there is (x, y, z) in E with $(c_1, c_2, c_3)\, \alpha_3' (x, y, z)\, \alpha_2' (b_1, b_2, b_3)$. But, necessarily, $z = c_3$ and $y = b_2$, so we have (x, b_2, c_3) in E. By hypothesis, $\alpha_1 \circ \alpha_1' = \nabla$, thus there is (u, v, w) in E such that $(x, b_2, c_3)\, \alpha_1 (u, v, w)\, \alpha_1' (a_1, a_2, a_3)$. Here, $v = b_2$, $w = c_3$, and $u = a_1$ whence (a_1, b_2, c_3) is in E. (a) is even simpler. Of course, this can be generalized to any finite number of factors. ∎

Combining Propositions 4 and 6 we arrive at a structural understanding of permutability: Define the *join* $\alpha \vee \beta$ of two equivalence relations as the smallest equivalence relation containing both α and β

$$\alpha \vee \beta = \bigcup \alpha \circ \beta \circ \alpha \circ \cdots \circ \beta.$$

Now, two equivalence relations α and β permute if and only if after factoring out $\alpha \cap \beta$ each class K of the join $\alpha \vee \beta$ is directly decomposed by the restrictions $\alpha \mid K$ and $\beta \mid K$.

The languages L_{Π} and $L_{\Pi s}$ associated with systems of equivalences are defined in terms of atomic sentences having one of the forms $\alpha \subseteq \beta$ (*inclusion*), $\gamma = \alpha \cap \beta$ (*meet sentence*), and $\gamma = \alpha \circ \beta$ (*product sentence*), respectively, with distinct α, β, γ in Ξ_{∞}. Let L_{Π} consist of all implications of the form

$$H_1 \wedge \cdots \wedge H_n \Rightarrow F_1 \wedge \cdots \wedge F_m,$$

where the H_i and F_j are atomic. Motivated by the special fd + emvd implications considered in Lemma 4 of Beeri and Vardi (1981), let $L_{\Pi s}$ consist of those sentences in L_{Π} where $m = 1$ and F_1 is a product sentence $\gamma = \alpha \circ \beta$ such that the inclusion sentences $\alpha \subseteq \gamma$ and $\beta \subseteq \gamma$ are among the H_i. The validity of an atomic sentence in a system (E, h) of equivalences is defined in the obvious fashion: e.g., $(E, h) \models \alpha \subseteq \beta$ if and only if $h(\alpha) \subseteq h(\beta)$—provided that α and β are in the domain of h. Of course, this means that the corresponding first-order sentences are valid in the relational structure $(E, h(\xi)_{\xi \in \Xi})$.

PROPOSITION 7. *Let $\mu \subseteq h(\xi) \subseteq \tau$ for all $\xi \in \Xi$. Then for any atomic sentence F the following are equivalent:*

1. $(E, h) \models F$.
2. $(E, h)/\mu \models F$.
3. $(T, h \mid T) \models F$ *for every equivalence class T of τ.*

Proof. Consider a sentence $\alpha \circ \beta = \gamma$ and write α in place of $h(\alpha)$. For any a, b, c with μ-classes $\tilde{a}, \tilde{b}, \tilde{c}$ one has $a\alpha c\beta b$ if and only if $\tilde{a}\alpha/\mu\, \tilde{c}\, \beta/\mu\tilde{b}$ whence $(\alpha \circ \beta)/\mu = \alpha/\mu \circ \beta/\mu$. Similarly, $a\alpha b$ and $a\beta b$ if and only if $\tilde{a}\alpha/\mu\tilde{b}$ and $\tilde{a}\beta/\mu\tilde{b}$ whence $(\alpha \cap \beta)/\mu = \alpha/\mu \cap \beta/\mu$. In particular, $\alpha \subseteq \beta$ if and only if $\alpha/\mu \subseteq \beta/\mu$. This proves the equivalence of 1 and 2. The equivalence of 1 and 3 follows from $(\alpha \circ \beta) \mid T = \alpha \mid T \circ \beta \mid T$. ∎

One concludes:

PROPOSITION 8. *Π is equivalent to Π_0 with regard to any fragment of L_{Π} via passing to canonical quotients and leaving sentences unchanged.*

In Section 6 we will consider a certain structure on Ξ_{∞} and call certain finite subsets *closed*. For now, the only point of interest is that for each finite $\Xi \subseteq \Xi_{\infty}$ there is at least one closed $\Xi^c \supseteq \Xi$. Denote by Π^c the subclass of Π consisting of those (E, h) with domain of h being a closed set.

PROPOSITION 9. *Π^c is equivalent to Π with regard to any fragment of L_{Π} (leaving sentences unchanged).*

Proof. Consider (E, h) in Π falsifying $H \Rightarrow F$—in particular, all ξ occurring in $H \Rightarrow F$ are in the domain Ξ of h. Extend h to h' with domain Ξ^c defining $h'(\xi)$ arbitrarily for $\xi \notin \Xi$. Then (E, h') is in Π^c and falsifies $H \Rightarrow F$. ∎

4. DATABASES

We recall some definitions following Kanellakis (1990); cf. also Fagin and Vardi (1986) and Paredaens *et al.* (1989). Fix a countably infinite attribute set U_{∞}. Under the pure universal relation assumption (which is quite appropriate for implication problems), a database d is given by a finite subset U of U_{∞}, for each A in U a domain $\Delta[A]$ of values of the attribute A, and a subset (relation) r of the direct product $\prod_{A \in U} \Delta[A]$. For a tuple t in r and $X \subseteq U$ let $t[X]$ be the restriction of t to X.

The atomic sentences to be considered are the functional dependencies (fd's) $X \rightarrow Y$ and the embedded multivalued dependencies (emvd's) $[X, Y]$ with finite $X, Y \subseteq U_{\infty}$. Validity is defined as follows: $d \models X \rightarrow Y$ if and only if for all $s, t \in r$, if $s[X] = t[X]$ then $s[Y] = t[Y]$—provided X and Y are subsets of U, at all. $d \models [X, Y]$ if and only if for every $t_1, t_2 \in r$ with $t_1[X \cap Y] = t_2[X \cap Y]$ there exists $t \in r$ with $t[X] = t_1[X]$ and $t[Y] = t_2[Y]$; in other words, the restriction of r to $XY = X \cup Y$ is the natural join of the restrictions of X and Y. The original notion of an emvd $X \twoheadrightarrow Y \mid Z$ amounts to $[XY, X(Z - Y)]$ and, conversely, $[X, Y]$ to: $X \cap Y \twoheadrightarrow (X - Y) \mid Y$.

Let L_{emvd} (and L_s) consist of all implications $H \Rightarrow F$ where F is an emvd and H a conjunction of emvd's (and fd's). Let $L_{\text{fd + emvd}}$ consist of all implications

$$H_1 \wedge \cdots \wedge H_n \Rightarrow F_1 \wedge \cdots \wedge F_m$$

where the H_i and F_j are fd's or emvd's.

One could also consider a notion of fd or emvd and then of implication, which also mentions the attribute set U—to be interpreted only in databases with attribute set U (an mvd then is an emvd $[X, Y]$ with $U = XY$). A U-implication of fd's and emvd's can be considered a V-implication for every $V \supseteq U$ and the first holds for all databases with attribute set U if the latter does for V: a U-database can be turned into a V-database by adding singleton domains, the converse being achieved by projection; this change of point of view has no impact on the validity of dependencies with attributes in U. Consequently, an implication is valid if and only if so is the corresponding U-implication where U is the set of attributes actually occurring in the implication.

Also, one might introduce typed and untyped versions of the implication problem by considering only databases with $\Delta[A] \cap \Delta[B] = \varnothing$ for all $A \neq B$ resp. $\Delta[A] = \Delta[B]$ for all A, B. Again, for emvd's this does not change the implication problem. Each database can be turned into a untyped one by replacing each $\Delta[A]$ by $\Delta = \bigcup_{A \in U} \Delta[A]$ and into a typed one by replacing the values $v \in \Delta[A]$ by pairs (v, A). Thus, the question of decidability does not depend on which approach we take.

For U-implications there is a well-known translation into a language of first order logic based on one U-ary relation symbol (Beeri and Vardi, 1981). Since there are complete calculi working uniformly for all possible sets of relation symbols the above considerations yield the following.

THEOREM 10. *The set of valid fd + emvd-implications is recursively enumerable.* ∎

Now, consider a subset E of a direct product of sets Δ_A, $A \in U$, i.e., a database. Then one has projection maps $\pi_A: E \to \Delta_A$ yielding for each $t \in E$ its A-component $\pi_A(t) = t[A]$ in Δ_A. With each of these maps one has its kernel equivalence relation θ_A. For a set X of attributes write $\theta_X = \bigcap_{A \in X} \theta_A$. Thus $s \theta_X t$ if and only if $s[X] = t[X]$. The following is obvious but fundamental and can be found in Rauszer (1987), in principle, for the fd's in Cosmadakis *et al.* (1986).

LEMMA 11. *For any database d, and sets X, Y, Z of its attributes*

$$\theta_X \subseteq \theta_Y \quad \text{iff } d \models X \to Y$$

$$\theta_Z = \theta_X \cap \theta_Y \quad \text{iff } d \models XY \to Z \wedge Z \to XY$$

$$\theta_{X \cap Y} = \theta_X \circ \theta_Y \quad \text{iff } d \models [X, Y]$$

and, if X, Y, Z are pairwise disjoint, then

$$\theta_Z = \theta_X \circ \theta_Y \quad \text{iff } d \models [XZ, YZ] \wedge X \to Z \wedge Y \to Z.$$

Proof. $\theta_X \subseteq \theta_Y$ means that $s[X] = t[X]$ implies $s[Y] = t[Y]$ for all s, $t \in r$, i.e., that $X \to Y$ holds in d. In particular, $\theta_X = \theta_Y$ if and only if $X \to Y$ and $Y \to X$ hold in d. By definition $\theta_{XY} = \theta_X \cap \theta_Y$ yielding the second claim. The third is a special case of the fourth. $\theta_Z = \theta_X \circ \theta_Y$ implies $\theta_X \subseteq \theta_Z$ and $\theta_Y \subseteq \theta_Z$, whence the validity of $X \to Z$ and $Y \to Z$. Now, consider t_1, $t_2 \in r$ with $t_1[Z] = t_2[Z]$ and observe that $Z = XZ \cap YZ$. Then $t_1 \theta_Z t_2$, so by hypothesis there exists $t \in r$ with $t_1 \theta_X t \theta_Y t_2$ or, in other words, $t_1[X] = t[X]$ and $t[Y] = t_2[Y]$ from which we get $t_1[XZ] = t[XZ]$ and $t[YZ] = t_2[YZ]$ using the fd's already derived. But this shows that $d \models [XZ, YZ]$. In the converse direction, we have $\theta_X \subseteq \theta_Z$ and $\theta_Y \subseteq \theta_Z$ from the fd's. Finally, consider $t_1 \theta_Z t_2$ or $t_1[Z] = t_2[Z]$. Then the emvd yields a $t \in r$ with $t_1[X] = t[X]$ and $t[Y] = t_2[Y]$, which amounts to $t_1 \theta_X \circ \theta_Y t_2$. ∎

THEOREM 12. *Π and $L_\Pi(L_{\Pi_S})$ are equivalent to $\mathscr{L}$ and $L_{\mathrm{fd}+\mathrm{emvd}}(L_S)$.*

Proof. Choose a bijection ϕ of Ξ_∞ onto U_∞ writing the images of α, β, γ as A, B, C for convenience. Translate $H \Rightarrow F$ in L_Π into $H' \Rightarrow F'$ in $L_{\mathrm{fd}+\mathrm{emvd}}$, replacing each atomic sentence by a conjunction of atomic sentences according to Lemma 11. Given (E, h) in Π, by Propositions 8 and 4 we may assume it is a system of projection kernels. So, let $U = \phi(\Xi)$, $\Delta[\phi(\xi)] = \Delta_\xi$, and $r = E$ to obtain a database. Conversely, a database gives rise to a system of projection kernels. So we have an interpretation.

When L_S is concerned and F is $\gamma = \alpha \circ \beta$, then form H' as before and let F' be the emvd $[AC, BC]$. Observe that by the definition of L_{Π_S} the conjunction H contains $\alpha \subseteq \gamma$ and $\beta \subseteq \gamma$, so H' has the fd's $A \to C$ and $B \to C$ as conjuncts. Again, by Lemma 11 H holds in d if and only H' holds. But, assuming H or H' we now have F equivalent to F', too.

For the converse interpretation let $X \mapsto \alpha_X$ be a bijection from the set of all finite subsets of U_∞ onto Ξ_∞. Given $H \Rightarrow F$ in $L_{\mathrm{fd}+\mathrm{emvd}}$ add in H' the conjunction of all $\alpha_X \cap \alpha_Y = \alpha_{X \cup Y}$ where X, Y are subsets of the attribute set of $H \Rightarrow F$ while translating the fd's and emvd's according to Lemma 11: $X \to Y$ by $\alpha_X \subseteq \alpha_Y$ and $[X, Y]$ by $\alpha_{X \cap Y} = \alpha_X \circ \alpha_Y$. The structures are again associated via Propositions 8 and 4. The additional hypotheses make sure that $h(\alpha_X)$ is associated with θ_X. ∎

TABLE 1

Example

A	B	C	A'	C'
a	b	c	a'	c
a	b'	c	a'	c
a_1	c	c'	a''	a_1
a	b'	c'	a''	c'
a	b	c'	a''	c'
a_2	c'	c'	a''	a_2

EXAMPLE. We slightly modify example 2.3.12 in Kanellakis (1990). Consider attributes A, B, C, A', C' and the relation r given by Table 1.

The fd $AA' \to C$ is witnessed by $\alpha \cap \alpha' \subseteq \gamma$ and the emvd $[AB, AC]$ by $\alpha = \beta \circ \gamma'$. On the other hand, $\gamma' = \alpha \cap \gamma$ is witnessed by the conjunction $C' \to AC \wedge AC \to C'$ of fd's and $\alpha = \beta \circ \gamma'$ by the conjunction $[AB, AC'] \wedge B \to A \wedge C' \to A$ of an emvd and fd's.

Let $\alpha, \beta, \gamma, \alpha', \gamma'$ denote the associated partitions. Listing the tuples in r as 1, ..., 6 we indicate the partitions in the figure by encircling and labeling their equivalence

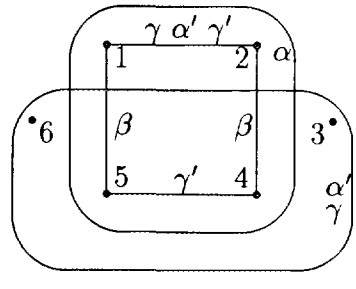

FIG. 2. Example in partition form.

TABLE 2

Subdirect Decomposition

A	B	C	A'	C'
1	1	1	1	1
1	2	1	1	1
3	3	3	3	3
1	2	3	3	4
1	1	3	3	4
6	6	3	3	6

classes—not caring about singletons and representing 2-element classes by an edge; see Fig. 2.

Constructing a database from a system of equivalences one has to list the classes of a δ, somehow, e.g., in the form $k_1^\delta, \ldots, k_m^\delta$ where each k_i is the first element in its δ-class. Of course, the superscripts δ may be omitted; see Table 2.

5. SOLVABILITY

In this section, we prove some positive results about dependencies. In accordance with Lemma 11 we conceive a database d as a set E together with a map $A \mapsto f(A)$ of U_∞ into $\Pi(E)$. We write $f(X) = \bigcap_{A \in X} f(A)$. For the validity of a dependency or implication, what matters is only the values $f(A)$ of attributes A actually occurring in it. A database $d = (E, \theta)$ is called *monolithic* if there is an equivalence relation μ on E such that $\mu \subseteq \theta_A$ for all A and all μ-classes are at least 2-element.

LEMMA 13. *For each database there is a monolithic one satisfying precisely the same fd's and emvd's (and preserving finiteness).*

Proof. Trivial set theory yields a database $\hat{d} = (\hat{E}, \hat{f})$ and an equivalence relation $\mu \subseteq \bigcap_{A \in U} \hat{f}_A$ on $\hat{E}$ having each class 2-element and $\hat{d}/\mu$ isomorphic to d. Now apply Proposition 7 and Lemma 11. ∎

Let H a finite conjunction of fd's and emvd's and V the set of attributes occurring in H. We use $H \models F$ to mean that $H \Rightarrow F$ holds in all databases—restricting to finite databases claims and proofs are verbally the same. We write $X - Y$ for the difference set.

LEMMA 14. $H \models X \to Y$ if and only if $Y - V \subseteq X - V$ and $H \models X \cap V \to Y \cap V$.

Proof. Assume $H \models X \to Y$ and consider a database d satisfying H. Construct d' putting $f'(A) = \nabla$ for $A \notin V$ and $f'(A) = f(A)$ for $A \in V$. Then in d' we have $f'(Z) = f'(Z \cap V)$ for all Z. Since $d \models X \to Y$ it follows that $d' \models X \cap V \to Y \cap V$. Since d and d' coincide on V, this fd holds in d, too. Now, assume there is $A \in Y - V$, $A \notin X$. By Lemma 13, we may suppose that d is monolithic. But, redefining $f'(A) = id$, we get d' satisfying H and $f'(Y) = id \subset \mu \subseteq f'(X)$, contradicting the validity of $X \to Y$. The converse implication is trivial. ∎

LEMMA 15. $H \models [X, Y]$ *if and only if one of the following takes place:*

1. $XY \subseteq V$ *and* $H \models [X, Y]$.
2. $X - V = Y - V$ *and* $H \models X \to Y \vee Y \to X$.
3. $Y - V \subseteq X$ *and* $H \models X \cap Y \cap V \to Y \cap V$.
4. $X - V \subseteq Y$ *and* $H \models X \cap Y \cap V \to X \cap V$.

Proof. Assume $H \models [X, Y]$ and consider monolithic d satisfying H. If there were $A \in X - V$, $A \notin Y$ and $B \in Y - V$, $B \notin X$ we could redefine $f'(A) = f'(B) = id$ and $f'(C) = f(C)$, elsewhere, to obtain $d' \models H$ but $f'(X) \circ f'(Y) = id \subset \mu \subseteq f'(X \cap Y)$. Now, assume $Y - V \subset X - V$. Choosing $A \in X - V$, $A \notin Y$, and redefining $f'(A) = id$ we get d' satisfying H and $f'(X) = id$ whence $f'(Y) = f'(X) \circ f'(Y) = f'(X \cap Y)$. Since $f'(B) = f(B)$ for all $B \in Y$, it follows that $f(Y) = f(X \cap Y)$ and so $H \models X \cap Y \to Y$. By the preceding lemma, we get $H \models X \cap Y \cap V \to Y \cap V$, i.e., case 3. Similarly, $X - V \subset Y - V$ leads to case 4. So $X - V = Y - V \neq \varnothing$ remains for us to consider. Assume that $f(X \cap V) \neq f(X \cap Y \cap V) \neq f(Y \cap V)$. Then there is a pair $(a, b) \in f(X \cap Y \cap V)$ with $(a, b) \notin f(X \cap V)$ and $(a, b) \notin f(Y \cap V)$. For all $A \in X - V$ redefine $f'(A)$ as the equivalence α consisting of the diagonal and (a, b) and (b, a) whereas $f'(B) = f(B)$, else. Then $d' \models H$ and $f'(X) \circ f'(Y) = id \subset \alpha = f'(X \cap Y)$, a contradiction. So we arrive at 2 in this case. Again, the converse implication is trivial. ∎

Proof of Theorem 3. Let a conjunction H of fd's and emvd's be given. Consider fd-consequences, first. By Lemma 14, the set $\{X \to Y \mid H \models X \to Y\}$ of fd-consequences is the union of the sets $\{SZ \to TW \mid W \subseteq Z \subseteq U_\infty\}$ where S, T are contained in the attribute set of H and $H \models S \to T$. Each of these sets is recursive, obviously. And there are only finitely many such $S \to T$, so the union is recursive, too. For emvd's the reasoning is analogous using Lemma 15. ∎

The following is, implicitly, in Lemma 4 in Beeri and Vardi (1981). It allows, one to remove the fd's from the antecedent of an implication. From the authors point of view, this is the crucial step in the undecidability result, so it deserves that a proof be included here. The proof is immediate by Lemmas 17 and 18, below.

THEOREM 16. *For the class $\mathscr{D}$ of all databases, there is an equivalence between L_{emvd} and L_s.*

Let $\hat{U}_\infty$ a disjoint copy of U_∞ and $A \mapsto \hat{A}$ a bijection. Let $U_\infty^+ = U_\infty \cup \hat{U}_\infty$, L_s^+ and L_{emvd}^+ be the corresponding languages, and $\mathscr{D}^+$ be the corresponding class of databases. For finite U let I_U be the conjunction of all $A \to \hat{A}$ and $\hat{A} \to A$, A in U, and let $U^+ = U \cup \hat{U}$. Let L_{ss}^+ the fragment of L_s^+ consisting of $H \Rightarrow F$ where F is an emvd over U_∞ and H is a conjunction of emvd's over U^+ and of I_U.

LEMMA 17. *$\mathscr{S}$ and L_s can be interpreted into $\mathscr{S}^+$ and L_{ss}^+.*

Proof. Given $H \Rightarrow F$ let $F' = F$. One may assume that the fd's in H are of the form $X \to A$ with A not in X. Under the hypothesis I_U each such fd is equivalent to the U^+-mvd $[U^+ - A, XA]$. Namely, under this hypothesis, $f^+(U^+ - A) \subseteq f(XA)$ whence $fX = f(XA) \cdot f(U^+ - A)$ implies $fX \subseteq f(XA) \cdot f(XA) = f(XA)$. The converse (that the fd implies the mvd) is trivial: from $fX = f(XA)$ one gets $fX = f(XA) \cdot f(U^+ - A)$.

So let H' arise from H by replacing the fd's by the corresponding mvd's and forming the conjunction with I_U. For a model (E, f) of H in $\mathscr{S}$ define (E, f^+) in $\mathscr{S}^+$ such that $f^+A = f^+\hat{A} = fA$. Conversely, from a model (E, f^+) of H' pass to (E, f) just by restricting f^+. ∎

LEMMA 18. *$\mathscr{S}^+$ and L_{ss}^+ can be interpreted into $\mathscr{S}^+$ and L_{emvd}^+.*

Proof. For $X \subseteq U^+$ let $X^+ \supseteq X$ be the smallest set containing with A also $\hat{A}$ and conversely. Given $H \Rightarrow F$ in L_{ss}^+ let $H' \Rightarrow F'$ arise by omitting all fd's and replacing $[X, Y]$ by $[X^+, Y^+]$. Under the hypothesis I_U, H is equivalent to H', F to F' (since $fX = fX^+$). Thus any model of H not satisfying F is also a model of H' not satisfying F'. Conversely, given a model (E, f) of H' not satisfying F', define $f'A = f'\hat{A} = fA \cap f\hat{A}$ to obtain a model of H not satisfying F. ∎

6. ABELIAN CONGRUENCE SYSTEMS

As a first step towards unsolvability, we interpret the universal Horn theory of semigroups into subgroup systems of abelian groups. The basic ideas are well known; cf. Lipshitz (1974). However, having the language L_{IIs} in mind, we have to a use a language for semigroups which corresponds to viewing them is ternary relational structures; cf. Gurevich and Lewis (1982). As a primer in (general) algebra, we recommend Kurosh (1963).

To define the language L_{SG}, let X_∞ be a countably infinite set of generator symbols and consider atomic sentences of the form $z = xy$ and $x = y$. Let L_{SG} consist of all sentences $H \Rightarrow F$ where H is a conjunction of atomic sentences and F of the form $x = y$. Let $\mathscr{S}$ consist of all pairs (S, σ) where S is a semigroup and σ a map of a finite $X \subseteq X_\infty$ into S. A *semigroup* is an algebraic structure with an associative binary operation, usually written as multiplication. It is a *monoid* if it has a *neutral* element e satisfying $ae = ea$ for all a. Any semigroup can be embedded into a monoid by just adding a neutral element.

An *abelian group* M is a commutative monoid (written additively with neutral element 0) having for each a a (uniquely determined) *inverse* $-a$ such that $a + (-a) = 0$. A *subgroup* is a subset containing 0 and closed under addition and inversion. With two subgroups N, K of M the smallest subgroup containing both N and K is

$$N + K = \{n + k \mid n \in N, k \in K\} = K + N.$$

For subgroups one has the Modular Law

$$A \supseteq C \Rightarrow A \cap (B + C) = A \cap B + C.$$

Namely, if a is in the left-hand side, i.e., $a = b + c$ with $a \in A$, $b \in B$, $c \in C$, then $b = a - c \in A$, whence $b \in A \cap B$ and a also in the right-hand side. The converse is trivial. *Products* of groups are direct products of the underlying sets with operations component wise. An *homomorphism* r of an abelian group M into another, N, is a map $a \mapsto ar$ (exceptionally, we write application of maps on the right, in this context) of M to N such that

$$(a + b)r = ar + br \qquad \text{for all} \quad a, b \in M.$$

It easily follows that

$$0r = 0 \qquad \text{and} \qquad (-a)r = -(ar).$$

A homomorphism is an *embedding* if it is one-to-one and an *endomorphism*, if $M = N$ as groups. Under composition, the endomorphisms of M form a monoid $End(M)$, the *endomorphism monoid*, with neutral element $e = id$,

$$a(rs) = (ar)s \quad \text{and} \quad ae = a \quad \text{for all} \quad a \in M, r, s \in End(M).$$

These identities mean that $End(M)$ *acts* on M on the right via $a \mapsto ar$. The following is well known (cf. Kurosh, 1963, 10.5) and basic for the intended interpretation.

PROPOSITION 19. *Every (finite) monoid can be embedded into the endomorphism monoid of some (finite) abelian group.*

Proof. Choose a finite field k, and let A consist of all formal sums

$$\sum_{s \in S} \lambda_s s \text{ with } \lambda_s \in k \text{ and all but finitely many } \lambda_s = 0$$

with addition defined by

$$\sum_{s \in S} \lambda_s s + \sum_{s \in S} \mu_s s = \sum_{s \in S} (\lambda_s + \mu_s) s.$$

In particular, two sums are equal if and only if $\lambda_s = \mu_s$ for all their coefficients. In other words, A is a k-vector space with basis S. With $r \in S$ associate the map ϕ_r given by

$$\sum_{s \in S} \lambda_s s \mapsto \left(\sum_{s \in S} \lambda_s s \right) r = \sum_{t \in S} \left(\sum_{s \in S,\, sr = t} \lambda_s \right) t.$$

Straightforward calculation shows that $(a+b)r = ar + br$ and $(ars) = (ar)s$ for all $a, b \in A$ and $r, s \in S$. If $\phi_r = \phi_s$, then $1r = (1e)r = (1e)s = 1s$, whence $r = s$. Thus, $r \mapsto \phi_r$ is an embedding of S into $End(A)$. ∎

Let 0 denote the one-element subgroup of the abelian group A. With each endomorphism r of A we associate its (negative) *graph*

$$\Gamma_r = \{(a, -ar) \mid a \in A\}.$$

LEMMA 20. *The graph Γ_r of an endomorphism of A is a subgroup of the abelian group A^2 and $\Gamma_r \subseteq \Gamma_s$ if and only if $r = s$. A subgroup B of A^2 is of the form Γ_r if and only if*

$$B \cap (0 \times A) = 0^2 \quad \text{and} \quad B + (0 \times A) = A^2.$$

Proof. Γ_r is a subgroup since $(a, -ar) + (b, -br) = (a+b, -(a+b)r)$ and $-(a, -ar) = (-a, -(-a)r$. Also, $\Gamma_r \cap (0 \times A) = \{(0, 0r)\} = \{(0, 0)\}$ and $\Gamma_r + (0 \times A) = \{(a, -ar) + (0, b) \mid a, b \in A\} = A^2$. Conversely, given B define $\phi(a) = b$ if and only if $(a, -b) \in B$. ϕ is well defined, since from $(a, -c) \in B$ it follows $(0, b-c) = (a, -c) - (a, -b) \in B$, whence $(0, b-c) \in B \cap (0 \times A)$ and so $b - c = 0$ and $b = c$. ϕ is defined on all of A since for any $a \in A$ by hypothesis there are $(b, -br) \in B$ and $(0, c) \in 0 \times A$ with $(a, 0) = (b, -br) + (0, c)$, whence $a = b$ and $(a, -ar) \in B$. Finally, if $\Gamma_r \subseteq \Gamma_s$ then for all a there is b with $(a, -ar) = (b, -bs)$, i.e., $a = b$ and $ar = bs$, which means that r and s are the same (being maps). ∎

In order to recapture the composition of endomorphisms, too, we have to use the abelian group A^3 and its *canonical frame of subgroups* consisting of the following subgroups:

$$A_1 = A \times 0^2, \quad A_2 = 0 \times A \times 0, \quad A_3 = 0^2 \times A$$

$$E_{12} = \{(a, -a, 0) \mid a \in A\}, \quad E_{13} = \{(a, 0, -a) \mid a \in A\}$$

$$E_{23} = \{(0, a, -a) \mid a \in A\}.$$

Observe that $A_i + A_j$ is isomorphic to $A \times A$; e.g. $(a, b, 0) \mapsto (a, b)$ for $i = 1$, $j = 2$. Now, we associate with each endomorphism r of A and ordered pair $i \neq j$ in $\underline{3} = \{1, 2, 3\}$ the graph subgroup Γ_{rij} of $A_i + A_j$, e.g.,

$$\Gamma_{r12} = \{(a, -ar, 0) \mid a \in A\}.$$

In particular, for the identity endomorphism e we have $\Gamma_{eij} = E_{ij}$.

LEMMA 21. *For any abelian group A, permutation i, j, k of $\underline{3}$, and endomorphisms r, s, t of A,*

$$\Gamma_{tik} = (\Gamma_{rij} + \Gamma_{sjk}) \cap (A_i + A_k) \quad \text{iff} \quad t = rs$$

$$(\Gamma_{rij} + E_{kj}) \cap (A_i + A_k) = \Gamma_{rik},$$

$$(\Gamma_{rij} + E_{ik}) \cap (A_k + A_j) = \Gamma_{rkj}.$$

Proof. Consider i, j, k being 1, 2, 3. Then $\Gamma_{r12} + \Gamma_{s23} = \{(a, -ar, 0) + ((0, b, -bs) \mid a, b \in A\} = \{(a, -ar + b, -bs) \mid a, b \in A\} =: B$ and $B \cap (A_1 + A_3) = \{(a, -ar + b, -bs) \mid a, b \in A, -ar + b = 0\} = \{(a, 0, -ars) \mid a \in A\}$. The second claim follows using $s = e$ resp. $r = e$ (and renaming the indices). ∎

The concept of a group with system of subgroups will be based on the language $L_{\Pi S}$ considered in Section 3. We have some structure on the set Ξ_∞ of constants. Namely, we assume that it includes particular names δ, α_I, ε_{ij}, ε, ε_{ijk}, where $I \subseteq \underline{3}$ and i, j, k is a permutation of $\underline{3}$. Also, we assume that for every $x, y \in X_\infty$ there are names ρ_{xij}, ρ_{xijk}, ρ_{xijlm}, and ρ_{xyijk}, where $l \neq m$ in $\underline{3}$. Of course, all these names are supposed to be distinct; we also assume that every name in Ξ_∞ is one of these.

The basic names are the α_i and ε_{ij} referring to the frame and the ρ_{xij} referring to endomorphisms $r_x = r(x)$, as well as, δ for a special purpose. The remaining names are needed to speak about the sums resp. relational products involved in the definition of a frame, the characterization of graph congruences (Lemma 31) corresponding to graph subgroups, and the "geometric multiplication" of the latter (Lemma 21). Thus, for a given finite $X \subseteq X_\infty$, we want all these to be available and define

$$\Xi(X) = \{\delta, \varepsilon\} \cup \{\varepsilon_{ij}, \rho_{xij} \mid i \neq j \in \underline{3}, x \in X\}$$

$$\cup \{e_{ijk}, \rho_{xijk} \mid i, j, k \text{ a permutation of } \underline{3}, x \in X\}$$

$$\cup \{\alpha_I \mid I \subseteq \underline{3}\}$$

$$\cup \{\rho_{xijlm} \mid i \neq j, l \neq m \in \underline{3}, x \in X\}$$

$$\cup \{\rho_{xyijk} \mid i, j, k \text{ a permutation of } \underline{3}, x \neq y \in X\}.$$

Call a finite subset Ξ *closed* if and only $\Xi = \Xi(X)$ for some finite $X \subseteq X_\infty$.

Let $\mathcal{A}$ consist of all quadruples (A^3, f, r, X), called *systems of subgroups*, where A is an abelian group, f a map from the closed subset $\Xi(X)$ of Ξ_∞ into the system of all subgroups of A^3, and r a map of X into $End(A)$ such that for all $x, y \in X$

$$f(\alpha_I) = \sum_{i \in I} A_i, \quad f(\alpha_\varnothing) = 0, \quad f(\varepsilon_{ij}) = E_{ij}$$

$$f(\varepsilon) = E_{12} + E_{23}, \quad f(\varepsilon_{ijk}) = E_{ij} + A_k$$

$$f(\rho_{xij}) = \Gamma_{r(x)\,ij}, \quad f(\rho_{xijk}) = \Gamma_{r(x)\,ij} + A_k$$

$$f(\rho_{xijlm}) = \Gamma_{r(x)\,ij} + E_{lm}, \quad f(\rho_{xyijk}) = \Gamma_{r(x)\,ij} + \Gamma_{r(y)\,jk}.$$

In other words, given A, we have no choice implementing a "frame-name." For "ρ-names" implementation is determined by the map r. Only δ can be interpreted, arbitrarily.

For α, β, γ in Ξ the validity of an atomic formula is defined as follows:

$$(A^3, f, r, X) \models \alpha \subseteq \beta \quad \text{iff} \quad f(\alpha) \subseteq f(\beta)$$
$$(A^3, f, r, X) \models \alpha \cap \beta = \gamma \quad \text{iff} \quad f(\alpha) \cap f(\beta) = f(\gamma)$$
$$(A^3, f, r, X) \models \alpha \cdot \beta = \gamma \quad \text{iff} \quad f(\alpha) + f(\beta) = f(\gamma).$$

THEOREM 22. *$\mathscr{S}$ and L_{SG} can be interpreted into $\mathscr{A}$ and L_{IIs}.*

Proof. Given a formula $H \Rightarrow F$ in L_{SG} having generator symbols in X, let H' be the conjunction of the following sentences (the first two types of which encode the given multiplicative relationships in H):

$$\rho_{zik} = \alpha_{ik} \cap \rho_{xyijk} \quad \text{where} \quad z = xy \quad \text{occurs in} \quad H$$
$$\rho_{xij} = \rho_{yij} \quad \text{where} \quad x = y \quad \text{occurs in} \quad H$$
$$\delta \subseteq \alpha_{12}, \text{ and } \delta = \rho_{x_0 12} \cap \rho_{y_0 12} \quad \text{where} \quad F \text{ is } x_0 = y_0.$$

Then choose F' as $\alpha_2 \cdot \delta = \alpha_{12}$. Observe that for systems in $\mathscr{A}$ satisfying H' the claim F' is equivalent to $r(x_0) = r(y_0)$. Indeed, if F' holds then by the Modular Law $\Gamma_{r(x_0)12} = \Gamma_{r(x_0)12} \cap (A_1 + A_2) = (\Gamma_{r(x_0)12} \cap A_2) + f(\delta) = 0 + f(\delta) = f(\delta) \subseteq \Gamma_{r(y_0)12}$, whence $r(x_0) = r(y_0)$ by Lemma 20. Conversely, if $r(x_0) = r(y_0)$, then $f(\delta) = \Gamma_{r(x_0)12}$ and F' follows from $\Gamma_{r(x_0)12} + A_2 = A_1 + A_2$.

Given (S, σ) embed S into $End(A)$ according to Proposition 19. For $x \in X$ let $r(x)$ the image of $\sigma(x)$ under this embedding. Let (A^3, f, r, X) be the associated structure in $\mathscr{A}$ such that $f(\delta) = f(\rho_{x_0 12} \cap \rho_{y_0 12})$. By Lemma 21, if (S, σ) satisfies H then (A^3, f, r, X) satisfies H' and $\sigma(x_0) \neq \sigma(y_0)$ amounts to F' not being satisfied.

Conversely, let (A^3, f, r, X) in $\mathscr{A}$ be a model of H' not satisfying F'. Choose $\sigma(x) = r(x)$, $x \in X$, and S as the subsemigroup these generate in $End(A)$. If $z = xy$ is in H then $\rho_{zik} = \alpha_{ik} \cap \rho_{xyjk}$ is in H' and we get

$$\Gamma_{r(z)\,ik} = (A_i + A_k) \cap (\Gamma_{r(x)\,ij} + \Gamma_{r(y)\,jk}) = \Gamma_{r(x)\,r(y)\,ik}$$

by Lemma 21 whence $r(z) = r(x)\, r(y)$ by Lemma 20 and so $\sigma(z) = \sigma(x)\, \sigma(y)$ by definition of σ. Similarly, $\sigma(x) = \sigma(y)$ if $x = y$ occurs in H. Hence, H is satisfied in (S, σ). Moreover, by the above remark the failure of F' in (A^3, f, r, X) entails that of F in (S, σ). ∎

From systems of subgroups we can pass to systems of equivalence relations naturally. A *congruence* of the abelian group M is an equivalence relation θ on M such that

$$a \theta b \quad \text{implies} \quad a + c \theta b + c.$$

Full compatibility with $+$ and $-$ follows, easily; cf. Kurosh (1963, 18.5).

PROPOSITION 23. *For an abelian group M there is a bijective correspondence*

$$N_\theta = \{a \in M \mid a \theta 0\}, \qquad a \theta_N b \quad \textit{iff} \quad a - b \in N$$

between subgroups N of M and congruences θ of M. Also

$$\theta_{N \cap K} = \theta_N \cap \theta_K$$
$$\theta_{N+K} = \theta_N \circ \theta_K = \theta_K \circ \theta_N.$$

In particular, any two congruences of an abelian group permute.

Proof. N_θ is a subgroup since $a\theta 0$, $b\theta 0$ leads to $a + b\ \theta\ a + 0 = a\theta 0$ and $-a = -a + 0\ \theta\ -a + a = 0$. θ_N is a congruence since $a - b$, $b - c \in N$ implies $b - a = -(a - b) \in N$, $a - c = (a - b) + (b - c) \in N$, and $(a + d) - (b + d) = a - b \in N$. Moreover, $a \theta_{N_\theta} b$ iff $a - b \in N_\theta$ iff $a - b\ \theta 0$ iff $a = (a - b) + b\ \theta\ 0 + b = b$ and $a \in N_{\theta_N}$ iff $a \theta_N 0$ iff $a = a - 0 \in N$. Finally, $a \theta_{N+K} b$ implies $a - b = n + k$ for some $n \in N$ and $k \in K$, whence $a \theta_N a - n = b + k\ \theta_K b$, and, conversely, $a \theta_N c \theta_K b$ implies $a - c \in N$, $c - b \in K$, and $a - b = (a - c) + (c - b) \in N + K$. ∎

A system (A^3, h, r, X) is in $\mathscr{A}G$ (an *abelian congruence system*) if and only if there is (A^3, f, r, X) in $\mathscr{A}$ such that f and h have the same finite domain $\Xi \subset \Xi_\mathscr{A}$ and $h(\xi)$ is the congruence associated with the subgroup $f(\xi)$ for each $\xi \in \Xi$. Then, from Theorem 22 we have, immediately

COROLLARY 24. *There is an interpretation of $\mathscr{S}$ and L_{SG} into $\mathscr{A}G$ and L_{IIs}.* ∎

The congruences associated with the canonical frame of subgroups of A^3 form the *canonical frame of congruences* of A^3. The congruence associated with a Γ_{rij} is called an *ij-graph congruence.*

7. COORDINATIZATION

Coordinatization can be understood as showing that certain systems of equivalences are isomorphic to abelian congruence systems. Since we have to work with the very restricted language L_{IIs} we cannot introduce any "axioms" —besides those expressing the fact that we work with equivalence relations. All we are allowed to do is to form conjunctions of atomic sentences and make them part of an antecedent, i.e., to require some of the equivalence relations to be meets or products of certain others, or some of them to be comparable. This really is not much to work with. So it may be a surprise that one arrives at an abelian group, after all. As an introduction, we study the 2-dimensional case; cf. Reidemeister (1968) or Denes and Keedwell (1974).

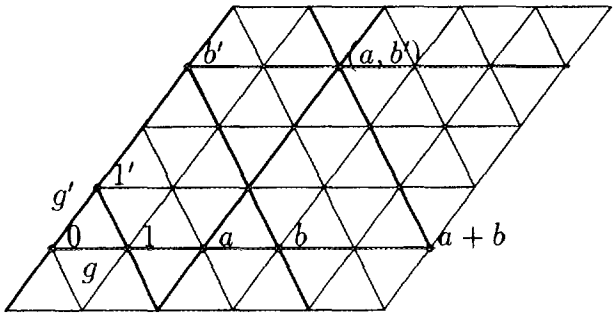

FIG. 3. Coordinatization.

A 3-*net* is a set E together with 3 equivalence relations α_1, α_2, ε_{12} such that

$$\alpha_1 \cap \alpha_2 = \alpha_1 \cap \varepsilon_{12} = \alpha_2 \cap \varepsilon_{12} = id$$

$$\alpha_1 \circ \alpha_2 = \alpha_1 \circ \varepsilon_{12} = \alpha_2 \circ \varepsilon_{12} = \nabla.$$

A *loop* is an algebraic structure A with a binary operation $+$ and a constant 0 such that $a + 0 = 0 + a = a$ for all $a \in A$ and such that for all a, b in A there are unique c, d in A such that $a + c = b$ and $d + a = b$.

PROPOSITION 25. *For a loop A we get a 3-net on $E = A^2$ such that $\alpha_1 = \theta_2$ and $\alpha_2 = \theta_1$ are the kernels of the projections and*

$$(a, 0)\ \varepsilon_{12}(0, b) \qquad \text{if and only if} \quad a = b$$

$$a + b = c \qquad \text{if and only if} \quad (a, b)\ \varepsilon_{12}(c, 0).$$

Each 3-net is isomorphic to the 3-net associated with a suitable loop.

Geometrically, think of a "plane," say the real plane. Choose an origin 0 and two "coordinate axes" g, g' passing through it. Using the parallels to g and g', each point of the plane is uniquely determined by its pair of coordinates on g and g'. Thus, the plane can be understood as a direct product $A_1 \times A_2$, where A_1 corresponds to g and A_2 to g'. Choosing units 1 and 1' on g and g' one can establish a one-to-one correspondence $r \leftrightarrow r'$ between g and g' such that the lines through r, r' and 1, 1' are parallel. So, one can identify the plane with a set A^2 and g with $A \times 0$, g' with $0 \times A$. Now, addition on A can be defined as in Fig. 3.

The system of equivalences considered here consists of the plane E and h listing the three partitions into parallels given by the two axes and the line through the two unit points. For the proof we use the following:

LEMMA 26. *Let γ be an equivalence relation on $A_1 \times A_2$ such that $\gamma \circ \theta_2' = id$ and $\gamma \circ \theta_2' = \nabla$ and let $0_1 \in A_1$, $0_2 \in A_2$ be fixed. Then there is a (unique) map $f: A_1 \to A_2$ with $f(a) = b$ iff $(a, 0_2)\ \gamma(0_1, b)$ and, in particular, $f(0_1) = 0_2$. f is a bijection if, in addition, $\gamma \cap \theta_1' = id$ and $\gamma \circ \theta_1' = \nabla$.*

Proof. f is well defined since $(a, 0_2)\ \gamma(0_1, b)$ and $(a, 0_2)\ \gamma(0_1, b')$ jointly imply $(0_1, b)\ \gamma(0_1, b')$ whence $(0_1, b)\ \gamma \cap \theta_2'(0_1, b')$ and $b = b'$ by hypothesis. On the other hand, given $a \in A_1$, in view of $\gamma \circ \theta_2' = \nabla$ there is (x, b) with $(a, 0_2)\ \gamma(x, b)\ \theta_2'(0_1, 0_2)$ where, necessarily, $x = 0_1$, whence $f(a) = b$. The analogous reasoning establishes the inverse map. ∎

Proof of Proposition 25. Given a loop A define ε_{12} on A^2 such that $(a, b)\ \varepsilon_{12}(c, d)$ if and only if $a + b = c + d$. It is easily seen that this yields a 3-net together with the projection kernels α_i. Conversely, if we consider systems of equivalences up to isomorphism only, by Lemma 6(a) we may assume that $E = A_1 \times A_2$ and $\alpha_i = \theta_i'$. Moreover, choosing $0_i \in A_i$, Lemma 26 provides us with a bijection $f: A_1 \to A_2$ with $f(a) = b$ if and only if $(a, 0)\ \varepsilon_{12}(0, b)$. Identifying A_2 with $A = A_1$ we have $(a, 0)\ \varepsilon_{12}(0, b)$ if and only if $a = b$ while retaining all other properties. Defining $a + b = c$ if and only if $(a, b)\ \varepsilon_{12}(c, 0)$ we have to show that this results into a loop. First, from $\varepsilon_{12} \circ \alpha_2 = \nabla$ for any a, b we have c, d such that $(a, b)\ \varepsilon_{12}(c, d)\ \alpha_2(0, 0)$, whence $d = 0$, which means that $a + b$ is always defined. Having $(a, b)\ \varepsilon_{12}(c', 0)$, too, we get from transitivity $(c, 0)\varepsilon_{12}(c', 0)$, whence $(c, 0)\ \varepsilon_{12} \cap \alpha_2(c', 0)$ and so $c = c'$. Therefore, addition is well defined. The other properties (which are not needed later on) are proved in a similar fashion. ∎

In the planar case, one needs some kind of Desargues' Law in order to obtain an abelian group. Geometers know that this comes free in 3 dimensions but did not bother to write that down for such rudimentary geometries as the ones we have to consider. So we have to follow Herrmann (1987). From the notions of the canonical frame of subgroups resp. congruences of A^3 we abstract the following properties which a 3-dimensional analogue of a 3-net should have. A *permuting frame of equivalences* on E is a family α_i, ε_{ij}, $(i \neq j \in \underline{3})$ of equivalence relations on E such that every two of them permute and

$$\alpha_i \cap (\alpha_j \circ \alpha_k) = id \tag{1}$$

$$\varepsilon_{ij} \cap \alpha_i = id \tag{2}$$

$$\varepsilon_{ij} \circ \alpha_i = \alpha_i \circ \alpha_j \tag{3}$$

$$\varepsilon_{ik} = \varepsilon_{ki} = (\varepsilon_{ij} \circ \varepsilon_{jk}) \cap (\alpha_i \circ \alpha_k) \tag{4}$$

for every permutation i, j, k of $\underline{3}$ and, in addition, $\alpha_1 \circ \alpha_2 \circ \alpha_3$ is the total equivalence relation on E.

Geometrically, we have 3 coordinate axes and a unit point on each of them, resp. the 3 lines connecting the 3 unit points—and all parallel classes determined by these lines. In particular, restricting any equivalence class K of an $\alpha_i \circ \alpha_j$ (which one might view as a coordinate plane) one obtains a 3-net $(K, \alpha_i \mid K, \alpha_j \mid K, \varepsilon_{ij} \mid K)$. Whereas the loop associated with a 3-net is not a really big deal, the heavenly

gift of a third dimension turns frames of permuting equivalences into easily understood algebraic objects. This has been stated in slightly more general form as Thm.1 of Herrmann (1987).

THEOREM 27. *The sets E with frames of permuting equivalences are up to isomorphism exactly the abelian groups A^3 with canonical frames of congruences.*

The basic steps are the isomorphism from arbitrary E to an $E = A^3$ with certain conditions on the frame including that the α_i are intersections of projection kernels (Lemma 29), the definition of an abelian group structure on A in that situation (Lemma 30), and the characterization of those equivalence relations which are graph congruences (Lemma 31). With (1_{12}) we quote the instance $i = 1, j = 2$ of (1), etc.

LEMMA 28. *For a frame of permuting equivalences let ε be the product of all ε_{ij} in arbitrary order and $\alpha'_i = \alpha_j \circ \alpha_k$ for i, j, k, a permutation of $\underline{3}$. Then*

$$\varepsilon_{12} \circ \varepsilon_{23} = \varepsilon_{12} \circ \varepsilon_{13} = \varepsilon_{23} \circ \varepsilon_{13} = \varepsilon \tag{5}$$

$$\alpha_i \cap \varepsilon = id \tag{6}$$

$$\varepsilon_{ij} = \alpha'_k \cap (\varepsilon_{ij} \circ \alpha_k) \tag{7}$$

Proof. (5) is immediate by Proposition 5 (b) and the fact that $\varepsilon_{ik} \subseteq \varepsilon_{ij} \circ \varepsilon_{jk}$ by (4). With (4) and (2) it follows that $\alpha_i \cap \varepsilon \subseteq \alpha_i \cap \varepsilon_{ik} = id$. Finally, since $\varepsilon_{ij} \subseteq \alpha'_k$ by (3), the Modular Law applies and with (1) one gets $\alpha'_k \cap (\varepsilon_{ij} \circ \alpha_k) = \varepsilon_{ij} \circ (\alpha'_k \cap \alpha_k) = \varepsilon_{ij} \circ id = \varepsilon_{ij}$. ∎

AD HOC DEFINITION. Let A be a set and $0 \in A$. A *0-nice frame* of A^3 is a frame of permuting equivalences on A^3 such that for every permutation i, j, k of $\underline{3}$

$$(a_1, a_2, a_3)\, \alpha_i (b_1, b_2, b_3) \quad \text{iff} \quad a_j = b_j, \quad a_k = b_k \tag{8}$$

$$(a_1, a_2, a_3)\, \alpha'_i (b_1, b_2, b_3) \quad \text{iff} \quad a_i = b_i \tag{9}$$

$$(a_1, a_2, a_3)\, \varepsilon_{ij} (b_1, b_2, b_3) \quad \text{iff} \quad (a'_1, a'_2, a'_3)\, \varepsilon_{ij} (b'_1, b'_2, b'_3)$$

where $a'_i = a_i, b'_i = b_i, a'_j = a_j, b'_j = b_j, a'_k = b'_k = 0, a_k = b_k$

or $a'_i = a_i, b'_i = 0, a'_j = 0, b'_j = b_j, a'_k = a_k = b'_k = b_k, a_i = b_j$. (10)

Every canonical frame of congruences is 0-*nice*.

LEMMA 29. *Every set E with frame of permuting equivalences is isomorphic to some set A^3 with 0-nice frame.*

Proof. By Proposition 6(c) we may assume that $E = A_1 \times A_2 \times A_3$ with (8) and (9). For each i choose an element 0_i in A_i and let $0 = (0, 0, 0)$—not mentioning the subscript where there is no need for doing so. Now, the α'_3-class of 0 can be identified with $A_1 \times A_2$ (forgetting the third component 0) and with the restrictions of α_1, α_2, and ε_{12} Lemma 26 applies to provide a bijection f_{12} of A_1 onto A_2. More precisely,

$$f_{12}(x) = y \quad \text{iff} \quad (x, 0, 0)\, \varepsilon_{12} (0, y, 0).$$

Similarly, we have f_{13}. Let $A = A_1$ and identify A_2 with A via f_{12}^{-1} and A_3 with A via f_{13}^{-1}. More formally, map $A_1 \times A_2 \times A_3$ bijectively onto A^3

$$(a, b, c) \mapsto (a, f_{12}^{-1}(b), f_{13}^{-1}(c)),$$

and continue with the image of the old system under this map using the old notations. Observe that all frame relations and (8), (9) remain in power and that

$$(a, 0, 0)\, \varepsilon_{12} (0, b, 0) \quad \text{iff} \quad a = b, \quad (a, 0, 0)\, \varepsilon_{13} (0, 0, c) \quad \text{iff} \quad a = c. \tag{11}$$

Now, $(0, b, 0)\, \varepsilon_{23} (0, 0, c)$ implies $(0, b, 0)\, \varepsilon_{12} (x, y, z)\, \varepsilon_{13} (0, 0, c)$ for some some (x, y, z) by (4_{23}). Having $\varepsilon_{12} \subseteq \alpha'_3$ and $\varepsilon_{13} \subseteq \alpha'_2$ (cf. (4)) it follows that $z = 0$ and $y = 0$, whence, by (11), $b = x = c$. Conversely, $(0, b, 0)\, \varepsilon_{12} (b, 0, 0)\, \varepsilon_{13} (0, 0, b)$ whence $(0, b, 0)\, \varepsilon_{12} \circ \varepsilon_{13} (0, 0, b)$ and $(0, b, 0)\, \varepsilon_{23} (0, 0, b)$ by (4_{23}). Summarizing,

$$(0, b, 0)\, \varepsilon_{23} (0, 0, c) \quad \text{iff} \quad b = c$$

so no symmetry has been lost. Coming to the proof of (10), first, $a_3 = b_3$ is necessary since $\varepsilon_{12} \subseteq \alpha'_3$. So we assume $a_3 = b_3$. Since $\alpha_3 \circ \varepsilon_{12} \circ \alpha_3 \circ \alpha_3 = \varepsilon_{12} \circ \alpha_3$ by permutability and transitivity, the pair $(a_1, a_2, 0), (b_1, b_2, 0)$ is in $\varepsilon_{12} \circ \alpha_3$ if and only so is the pair $(a_1, a_2, a_3), (b_1, b_2, a_3)$ is so. Both being in α'_3, one of them is in ε_{12} if and only the other is (in view of (7_{12}). But, in the case $a_2 = 0$ and $b_1 = 0$, for the first pair we already know by (11) that this happens if and only if $a_1 = b_2$. ∎

LEMMA 30. *If $\alpha_i, \varepsilon_{ij}$ is a 0-nice frame on A^3, then A is an abelian group such that for all permutations i, j, k of $\underline{3}$,*

$$a + b = c \quad \textit{iff} \quad (a_1, a_2, a_3)\, \varepsilon_{ij} (b_1, b_2, b_3)$$
$$\textit{where} \quad a_i = a, a_j = b, b_i = c, a_k = b_j = b_k = 0. \tag{12}$$

Proof. Define

$$a +_{12} b = c \quad \text{iff} \quad (a, b, 0)\, \varepsilon_{12} (c, 0, 0) \quad \text{and}$$
$$a +_{21} b = c \quad \text{iff} \quad (a, b, 0)\, \varepsilon_{21} (0, c, 0)$$

and, similarly, $a +_{ij} b$ according to $(12)_{ij}$. Recall that $\varepsilon_{ij} = \varepsilon_{ji}$ due to (4). That addition is well defined is seen as in the proof of Proposition 25. By this definition and (10_{12})

$$(a +_{12} b, 0, 0)\, \varepsilon_{12} (a, b, 0)\, \varepsilon_{21} (0, b +_{21} a, 0)\, \varepsilon_{12} (b +_{21} a, 0, 0),$$

whence $(a +_{12} b, 0, 0)\ \varepsilon_{12} \cap \alpha_1 (b +_{21} a, 0, 0)$, and in view of (2_{12})

$$a +_{12} b = b +_{21} a, \text{ and similarly } a +_{ij} b = b +_{ji} a.$$

Again, by the definition and (10_{23})

$$(a +_{12} b, 0, 0)\ \varepsilon_{12}(a, b, 0)\ \varepsilon_{23}(a, 0, b)\ \varepsilon_{13}(a +_{13} b, 0, 0),$$

whence $(a +_{12} b, 0, 0)\ \varepsilon \cap \alpha_1 (a +_{13} b, 0, 0)$, and by (6_1)

$$a +_{12} b = a +_{13} b, \qquad \text{and similarly} \quad a +_{ij} b = a +_{ik} b.$$

Now, one easily gets that addition is commutative and does not depend on the subscripts, e.g., $a +_{12} b = a +_{13} b = b +_{31} a = b +_{32} a = a +_{23} b = a +_{21} b = b +_{12} a$. Also, 0 is neutral by definition. To get an inverse of a observe that by (3_{21}) there is (b, c, d) with $0 \varepsilon_{12}(b, c, d)\ \alpha_2 (a, 0, 0)$ where necessarily $b = a$ and $d = 0$ which then yields $a + c = 0$ according to the definition of addition. Finally, by (12) and (10)

$$(((a + b) + c), 0, 0)\ \varepsilon_{13}(a + b, 0, c)\ \varepsilon_{12}(a, b, c)$$
$$\varepsilon_{23}(a, 0, b + c)\ \varepsilon_{13}(a + (b + c)), 0, 0),$$

whence $(((a + b) + c), 0, 0)\ \varepsilon \cap \alpha (a + (b + c), 0, 0)$, which results in the associative law $(a + b) + c = a + (b + c)$ by another use of (6_1). ∎

LEMMA 31. *Let α_i, ε_{ij} be a 0-nice frame on A^3, where A is an abelian group with addition according to* (12) *(e.g., a canonical frame of congruences). Then an equivalence relation ρ of E corresponds to an ij-graph congruence of the group A^3 if and only if it permutes with α_j and all α_k, ε_{ik} and ε_{jk} where $k \neq i, j$ and if one has*

$$\rho \cap \alpha_j = id \tag{13}$$

$$\rho \circ \alpha_j = \alpha_i \circ \alpha_j. \tag{14}$$

In particular, all α_i and ε_{ij} are congruences of A^3.

Proof. Consider $\rho \subseteq \alpha_i \circ \alpha_j$. Then from the Modular Law and (1), (2) it follows that

$$\rho = (\rho \circ \alpha_k) \cap (\alpha_i \circ \alpha_j) \tag{15}$$

$$\rho = (\rho \circ \varepsilon_{ik}) \cap (\alpha_i \circ \alpha_j), \tag{16}$$

$$\rho = (\rho \circ \varepsilon_{jk}) \cap (\alpha_i \circ \alpha_j). \tag{17}$$

For example, (16) is obtained from $(\rho \circ \varepsilon_{ik}) \cap \alpha_k' = \rho \circ (\varepsilon_{ik} \cap \alpha_k')$ and $\varepsilon_{ik} \cap \alpha_k' = \varepsilon_{ik} \cap \alpha_k' \cap \alpha_j' = \varepsilon_{ik} \cap \alpha_i = id$ where we used (2) and the fact that α_i' is a projection kernel.

Let $i = 1$, $j = 2$. Now, $(x, y, c)\ \rho (a, b, c)$ implies $(x, y, d)\ \alpha_3 \circ \rho \circ \alpha_3 (a, b, d)$ whence $(x, y, d)(\rho \circ \alpha_3) \cap \alpha_3' (a, b, d)$ by permutability, and also $(x, y, d)\ \rho (a, b, d)$ by (15_{12}). This amounts to

$$(x, y, c)\ \rho (a, b, c) \qquad \text{iff} \quad (x, y, d)\ \rho (a, b, d). \tag{18}$$

Thus, $(x, y, 0)\ \rho (a, b, 0)$ implies $(x + u, y, 0)\ \varepsilon_{13}(x, y, u)$ $\rho (a, b, u)\ \varepsilon_{13}(a + u, b, 0)$. In view of (12_{13}), (10_{13}), and (16_{123}) one has

$$(x, y, 0)\ \rho (a, b, 0) \qquad \text{implies} \quad (x + u, y, 0)\ \rho (a + u, b, 0). \tag{19}$$

Now, let $(x, y, z)\ \rho (a, b, c)$. Since $\rho \subseteq \alpha_3'$ in view of (14), we have $z = c$ and $(x, y, 0)\ \rho (a, b, 0)$ by (18). Then, given (u, v, w) application of (19) yields $(x + u, y, 0)\ \rho (a + u, b, 0)$. The analogue of (19) for addition in the second component leads to $(x + u, y + v, 0)\ \rho (a + u, b + v, 0)$. Finally, (18) (and $z = c$) results in $(x + u, y + v, z + w)\ \rho (a + u, b + v, c + w)$. Summarizing,

$$(x, y, z)\ \rho (a, b, c)$$
$$\text{implies} \quad (x + u, y + v, z + w)\ \rho (a + u, b + v, c + w).$$

But this means that ρ is a congruence of the abelian group A^3. Thus, the lemma is shown referring to Lemma 20 and Proposition 23. ∎

Proof of Theorem 27. Let E be a set with arbitrary permuting frame α_i, ε_{ij} of equivalences. By Lemma 29 we may assume that $E = A^3$ and that the frame is 0-nice. Then Lemmas 30 and 31 provide us with an abelian group structure on A such that the frame equivalences are congruences of A^3. By (8), α_1 corresponds to the subgroup $A_1 = A \times 0^2$. For ε_{12} we have $(a, b, c)\ \varepsilon_{12} 0$ iff $(a + b, 0, 0)\ \varepsilon_{12} 0$ and $c = 0$ (by (12)) iff $a + b = c = 0$ (by (10)). So ε_{12} corresponds to E_{12}, similarly for the other indices, showing that we indeed have the canonical frame of congruences. ∎

8. UNSOLVABILITY

Considering the language $L_{\Pi s}$ for systems (E, h) of equivalence relations, assume the set $\Xi_\varkappa$ to be structured as in Section 6. In particular, we have the notion of a "closed set" defined there. Let Π^c be the subclass of Π consisting of all (E, h) with h having closed domain.

Observe that the relational product $\circ$ is not an operation symbol in our language—otherwise no interpretation into databases and emvd's would be possible. But the coordinatization results involve some products of the basic equivalence relations. The necessity to speak about these is the very reason that we introduced the auxiliary names and why we work with closed subsets of names: within such, for any basic name we also have the associated auxiliary names available.

Given $X \subseteq X_\omega$, let H_X be the conjunction of the following atomic sentences (splitting an identity into two inclusions, if needed) where $x, y \in X$, $I \subseteq \underline{3}$, i, j, k a permutation of $\underline{3}$, and $l \neq m$ in $\underline{3}$:

$$\alpha_I \circ \alpha_J = \alpha_{I \cup J}, \qquad \alpha_I \cap \alpha_J = \alpha_{I \cap J} \tag{20}$$

$$\varepsilon_{ij} \cap \alpha_i = \alpha_\varnothing, \qquad \varepsilon_{ij} \circ \alpha_i = \alpha_{ij}, \qquad \varepsilon_{ij} \circ \alpha_k = \varepsilon_{ijk} \tag{21}$$

$$\varepsilon_{ij} \circ \varepsilon_{jk} = \varepsilon, \qquad \varepsilon_{ik} = \varepsilon_{ki} = \varepsilon \cap \alpha_{ik} \tag{22}$$

$$\rho_{xij} \cap \alpha_i = \rho_{xij} \cap \alpha_j = \alpha_\varnothing, \qquad \rho_{xij} \circ \alpha_j = \alpha_{ij} \tag{23}$$

$$\rho_{xijk} = \rho_{xij} \circ \alpha_k, \qquad \rho_{xijlm} = \rho_{xij} \circ \varepsilon_{lm} \tag{24}$$

$$\rho_{xijk} \cap \alpha_{ik} = \rho_{xik}, \qquad \rho_{xijk} \cap \alpha_{kj} = \rho_{xkj} \tag{25}$$

$$\rho_{xyijk} = \rho_{xij} \circ \rho_{yjk} \tag{26}$$

LEMMA 32. *(E, h) in Π^c is isomorphic to (A^3, f) for some abelian congruence system (A^3, f, r, X) in $\mathscr{A}G$ if and only if it satisfies H_X, $h(\alpha_\varnothing) = id$ and $h(\alpha_{\underline{3}}) = \nabla$.*

Proof. An abelian congruence system satisfies all the claimed relations, obviously. Conversely, let (E, h) be given. For convenience, we write ξ in place of $h(\xi)$. Equations (20)–(22) together with $\alpha_\varnothing = id$ and $\alpha_{\underline{3}} \nabla$ ensure that we have a permuting frame α_i, ε_{ij} on E. (In more detail: in view of $\alpha_\varnothing = id$, (1) and (2) are special cases of (20) and (21). From (20) we have $\alpha_{ij} = \alpha_i \circ \alpha_j$, so (3) follows from (21). Moreover, $\varepsilon = \varepsilon_{ij} \circ \varepsilon_{jk}$ from (22) and we obtain (4) from the second part of (22). Finally, all of the α_i and ε_{ij} permute since their products are again equivalence relations; cf. Prop.5: $\alpha_i \circ \alpha_j = \alpha_{ij}$ by (20), $\alpha_i \circ \varepsilon_{ij} = \alpha_{ij}$ by (21), $\varepsilon_{ij} \circ \alpha_k = \varepsilon_{ijk}$ by (21), $\varepsilon_{ij} \circ \varepsilon_{jk} = \varepsilon$ by (22).) So Theorem 27 allows us to assume that the given frame is the canonical frame of congruences of some abelian group A^3.

By (23), (24), and Lemma 31 the ρ_{xij}, $x \in X$, are graph congruences. (Namely, (13) and (14) are special cases of (23)—we already know $\alpha_\varnothing = id$ and $\alpha_i \circ \alpha_j = \alpha_{ij}$. Moreover, ρ_{xij} permutes with α_j, α_k, ε_{ik}, and ε_{jk} since the product with each of these is again an equivalence relation: $\rho_{xij} \circ \alpha_j = \alpha_{ij}$ by (23), $\rho_{xij} \circ \alpha_k = \rho_{xijk}$ by (24), $\rho_{xij} \circ \varepsilon_{lm} = \rho_{xijlm}$ by (24).) Thus, by definition, there are endomorphisms $r(i, j, x)$ of A such that ρ_{xij} is the congruence associated with $\Gamma_{r(i,j,x)ij}$.

Switching to subgroups via Proposition 23, by Lemma 21 we get that $(\Gamma_{r(x,i,j)ij} + E_{jk}) \cap (A_i + A_k)$ equals $\Gamma_{r(x,i,j)ik}$ and by (24),(25) that it equals $\Gamma_{r(x,i,k)ik}$. (Namely, $\rho_{xij} \circ \varepsilon_{jk} = \rho_{xijk}$ by (24) and $\rho_{xijk} \cap \alpha_{ik} = \rho_{xik}$ by (25).) With Lemma 20 it follows that $r(x, i, j) = r(x, i, k)$. Similarly, we get that $r(x, i, j)$ is independent of i, too. Then (24) and (26) show that we indeed have an abelian congruence system (A^3, h, r, X) associated with the map $x \mapsto r(x) = r(x, i, j)$ of X into $End(A)$. (Namely, denoting by $f(\xi)$ the subgroup associated with $\xi = h(\xi)$ we have $f(\rho_{xij}) = \Gamma_{r(x)ij}$ by construction and $f(\rho_{xijk}) = \Gamma_{r(x)ij} + A_k$ by (24), $f(\rho_{xijlm}) = \Gamma_{r(x)ij} + E_{lm}$ by (24), $f(\rho_{xyijk}) = \Gamma_{r(x)ij} + \Gamma_{r(y)jk}$ by (26).) ∎

Now, we are ready for the crucial

THEOREM 33. *There is an interpretation of $\mathscr{A}G$ and L_{IIs} into Π^c and L_{IIs}.*

Proof. Given $H \Rightarrow F$ in L_{IIs} let $F' = F$ and

$$H' = H \wedge H_X \wedge \bigwedge_{\xi \in \Psi^c} \alpha_\varnothing \subseteq \xi \subseteq \alpha_{\underline{3}},$$

where Ψ is the set of all ξ occurring in $H \Rightarrow F$ and X a finite subset of X_ω such that $\Psi^c := \Xi(X) \supseteq \Psi$.

For any structure (A^3, h, r, X) in $\mathscr{A}G$ we have (A^3, h) in Π^c and (by Lemma 32) H_X satisfied as well as $h(\alpha_\varnothing) = id$ and $h(\alpha_{\underline{3}}) = \nabla$. Therefore, $(A^3, h, r, X) \models H \Rightarrow F$ if and only if $(A^3, h) \models H' \Rightarrow F'$.

Conversely, let (E, h) in Π^c with domain $\Xi \supseteq \Psi^c$ a model of H' not satisfying F'. Redefine $\hat{h}(\xi) = h(\xi)$ for ξ occurring in $H' \Rightarrow F'$ and $\hat{h}(\xi) = h(\alpha_\varnothing)$, else. In $(E, \hat{h})$ the sentence H' is valid but F' is not—since the values of the constants occurring in these sentences have not been changed. Observe that due to H' we have $\hat{h}(\alpha_\varnothing) \subseteq \hat{h}(\xi) \subseteq \hat{h}(\alpha_{\{1,2,3\}})$ for all $\xi \in \Xi$. Hence by Proposition 7 we first may form the quotient $(E', h') = (E, \hat{h})/f(\alpha_\varnothing)$ and still have H' valid but not F'. By the same proposition, there is a class E'' of $h'(\alpha_{bf3})$ such that in the restriction (E'', h'') we have H' valid but not F'. But now we have $h''(\alpha_\varnothing) = id$ and $h''(\alpha_{\underline{3}}) = \nabla$.

Moreover, having H_X satisfied, Lemma 32 applies and there is (A^3, f, r, X) in $\mathscr{A}G$ such that (E'', h'') is isomorphic to (A^3, f). In particular, this abelian congruence system satisfies H' but not F'. ∎

Proof of Theorem 1. Propositions 24, 33, 9, and 12 combine to an interpretation of $\mathscr{S}$ and L_{SG} into $\mathscr{S}$ and L_s via the structure classes $\mathscr{A}G$, Π^c, and Π and the language L_{IIs}. The crucial Theorem 16 of Beeri and Vardi (1981) then carries the interpretation on to $\mathscr{S}$ and L_{emvd}. Therefore, it suffices to show that (finite) validity is algorithmically undecidable for L_{SG} and $\mathscr{S}$. Given a finite collection of semigroup words over X, choose a new generator symbol x_w for each subword $w(x_1, \ldots, x_n)$ and form a conjunction H_w of atomic sentences such that for any map σ of X_ω into a semigroup S one has $w(\sigma(x_1), \ldots, \sigma(x_n)) = \sigma(x_w)$ if and only if $(S, \sigma) \models H_w$—this is easily achieved inducting on the complexity of w. Now, translate an implication

$$w_1 = v_1 \wedge \cdots \wedge w_m = v_m \Rightarrow w_0 = v_0$$

into

$$H_{w_1} \wedge H_{v_1} \wedge x_{w_1} = x_{v_1} \wedge \cdots \wedge H_{w_m} \wedge H_{v_m} \wedge x_{w_m} = x_{v_m} \wedge H_{w_0} \wedge H_{v_0} \Rightarrow x_{w_0} = x_{v_0}.$$

Shortly, there is an interpretation of semigroups as algebraic structures into semigroups as ternary relational

structures with regard to universal Horn formulas cf. Beeri and Vardi (1981). Thus, a solution of the (finite) validity problem for L_{SG} and $\mathscr{S}$ would provide a decision procedure for the proper implications in the universal Horn theory of (finite) semigroups, which is also referred to as the word problem for (finite) semigroups. The problem and its finite variant are known to be unsolvable according to the classical result of Markov (1947) and Post (1947), respectively to Gurevich (1966). ▮

Proof of Corollary 2. The set of emvd-implications falsified in some finite database is recursively enumerable. So its complement cannot be recursively enumerable, in particular not via an axiomatization. Also, implication and finite implication cannot coincide in view of Theorem 10. ▮

9. DISCUSSION

We have carefully avoided mentioning lattices so far. Of course, the system of all equivalence relations on a set forms a lattice (with respect to $\cap$ and $\vee$) and so does the system of all congruences of an abelian group. Also, our coordinatization result arose in the context of von Neumann's (1960) coordinatization of complemented modular lattices, and interpreting semigroups into modular lattices via frames is the basic idea of Lipshitz (1974).

Yet it must be pointed out that in the present interpretation we do not refer to lattices of equivalence relations nor to abstract lattices which would require a global modularity assumption hardly to be satisfied by structures derived from databases where the only restriction allowed is a finite set of fd's and emvd's.

In contrast, we relied on local modularity which can be finitely encoded. For a base set with the 15 equivalence relations given by a frame and the products to be formed from it, this yielded the abelian group A and its endomorphism monoid $End(A)$. Then, for each element of $End(A)$ we wanted to speak about, we threw in another 30 and for each pair of them another 12 equivalence relations and one more for the conclusion. Thus, finitely many equivalence relations and atomic statements about them sufficed to deal with the subwords involved in a single finite semigroup presentation and identity to be tested.

Our result carries over to representable relation algebras having at least meet, product, and inversion among their operations and the identity relation as a constant. For a vector space V let $R(V)$ denote the relation algebra generated by the system of all congruences of V. For a class $\mathscr{R}$ of algebras let $\mathscr{R}_f$ denote the class of its finite members.

COROLLARY 34. *Let $\mathscr{R}$ a class of representable relation algebras containing $R(V)$ for some infinite dimensional vector space resp. finite $R(V)$'s of arbitrarily large dimension. Then the quasivariety generated by $\mathscr{R}$ and $\mathscr{R}_f$, resp., contains a finitely presented algebra with unsolvable word problem. In both case, the universal Horn theories of $\mathscr{R}$ and $\mathscr{R}_f$ are recursively inseparable.*

For relation algebras in the sense of Tarski in this situation even the equational theory is undecidable—this is a recent result of Andréka *et al.* (1993).

Proof. That ρ is an equivalence relation can be expressed by the statements $\rho \cap id = id$, $\rho^{-1} = \rho$, and $\rho \circ \rho = \rho$. Given a semigroup presentation, again consider the system of equivalences consisting of a permuting frame and a graph congruence for each semigroup generator—which can be specified by finitely many relations in terms of the relation algebra operations. Then the sublattice generated in the partition lattice is in fact contained in the congruence lattice of the associated abelian group, in particular joins are products. So one can translate semigroup words via lattice words into relation algebra words. Finally, recall that the results of Markov, Post, and Gurevich actually provide a fixed antecedent such that the corresponding universal Horn sentences are undecidable. Concerning inseparability, let Σ_{csg0}, Σ_{g0}, and Σ_{fsg} denote the sets of universal Horn sentences valid in all cancellative semigroups with zero, all groups with zero, and all finite semigroups, respectively. Gurevich and Lewis (1984) have shown that Σ_{cgs0} and the complement $C\Sigma_{fsg}$ are recursively inseparable. Then so are Σ_{g0} and $C\Sigma_{fsg}$, obviously. But groups with zero (conceived as semigroups) can be interpreted into relation algebras, too, having with each generator also its inverse—using Lemma 26 to establish invertibility. ▮

Lattices are in the background in another respect, too. Namely, partition lattices have been introduced by Cosmadakis *et al.* (1986) in the study of database dependencies. In particular, they created the concept of a partition dependency (pd) given by an identity involving meet and join (in the terminology of presentations one would speak of a lattice relation): the pd holds for a list of attributes if and only if the identity is satisfied for the corresponding list of equivalence relations. Then the implication and the finite implication problem for pd's reduce to the uniform word problem for lattices (since the quasivariety of all lattices is generated by (finite) partition lattices). For the latter, an efficient solution is already in Skolem (1920). Unfortunately, even mvd's cannot be expressed as a conjunction of pd's (Cosmadakis *et al.*, 1986, Theorem 5).

On the other hand, Sagiv *et al.* (1981) translated fd's and mvd's into propositional logic but showed that this would not work for emvd's. Day (1993) replaced the propositional formulas by a conjunction of lattice identities and proved that this works well for fd's and mvd's with respect to any equationally defined class of lattices (reducing to the propositional case when considering distributive lattices). Yet, for emvd's, this translation has not yet succeeded—in

any case, according to Day, modularity would be a necessary requirement.

What we used here are only rudiments of the above approaches, also to be found in Rauszer (1987). In particular, we beware the general join in partition lattices. However, the work of these authors was an incentive for the present paper.

Received July 31, 1991; final manuscript received January 26, 1995

REFERENCES

Andréka, H., Givant, St., and Németi, I. (1993), Undecidable equational theories of relation algebras, preprint.

Beeri, C., Fagin, R., and Howard, J. H. (1977), A complete axiomatization for functional and multivalued dependencies in database relations, *in* "Proc. ACM SIGMOD," pp. 47–61.

Beeri, C., and Vardi, M. Y. (1981), The implication problem for data dependencies, *in* "ICALP 81" (S. Even and O. Kariv, Eds.), pp. 73–85, Lecture Notes in Computer Science, Vol. **115**, Springer, Berlin.

Chandra, A. K., Lewis, H. R., and Makowsky, J. A. (1981), Embedded implicational dependencies and their inference problem, *Proc. ACM Symp. Theory of Computing* **13**, 342–352.

Cohn, P. M. (1981), "Universal Algebra," Reidel, Dodrecht.

Cosmadakis, S. S., Kanellakis, P. C., and Spyratos, S. (1986), Partition semantics for relations, *J. Comput. System Sci.* **33**, 203–233.

Day, A. (1993), A lattice interpretation of database dependencies, *in* "Semantics of Programming Languages and Model Theory" (M. Droste, and Y. Gurevich, Eds.), Algebra, Logic, and Applications, Gordon & Breach, London.

Denes, J., and Keedwell, A. D. (1974), "Latin Squares and Their Application," Academic Press, New York, 1974.

Dubreil, P., and Dubreil-Jacotin, M.-L. (1939), Théorie algébrique des relations d'équivalences, *J. de Math.* **18**, 63–95.

Fagin, R. (1977), Multivalued dependencies and a new normal form for relational databases, *ACM Trans. Database Systems* **2**(3), 262–278.

Fagin, R., and Vardi, M. Y. (1986), The theory of database dependencies: A survey, *in* "Mathematics of Information Processing" (M. Anshel and W. Gerwitz, Eds.), Symp. in Appl. Math., Vol. 34, pp. 19–72.

Gurevich, Y. S. (1966), The problem of equality of words for certain classes of semigroups, *Algebra i Logika* **5**, 25–35. [Russain]

Gurevich, Y. S., and Lewis, H. R. (1982), The inference problem for template dependencies, *Inform. and Control* **55**, 69–79.

Gurevich, Y. S., and Lewis, H. R. (1984), The word problem for cancellation semigroups with zero, *J. Symbolic Logic* **49**, 184–191.

Herrmann, C. (1987), Frames of permuting equivalences, *Acta. Sci. Math.* **51**, 93–101.

Hutchinson, G. (1973), Recursively unsolvable word problems for modular lattices and diagram chasing, *J. Algebra* **26**, 385–399.

Kanellakis, P. C. (1990), Elements of relational database theory, *in* "Handbook of Theoretical Computer Science, Vol. B: Formal Models and Semantics" (J. van Leeuwen, Ed.), pp. 1073–1156, Elsevier, Amsterdam.

Kurosh, A. G. (1963), "Lectures on General Algebra," Chelsea, New York.

Lipshitz, L. (1974), The undecidability of the word problem for projective geometries and modular lattices, *Trans. Amer. Math. Soc.* **194**, 171–180.

Markov, A. A. (1947), On the impossibility of certain algorithms in the theory of associative systems (Russian) I + II, *Dokl. Akad. Nauk* **55**, 587–590 and **58**, 353–356.

Ore, O. (1942), Theory of equivalence relations, *Duke Math. J.* **9**, 573–627.

Paredaens, J., de Bra, P., Gyssens, M., and van Gucht, D., (1989) "The Structure of the Relational Database Model," Springer-Verlag, Berlin.

Parker, D. S., and Parsaye-Ghomi, K. (1980), Inference involving embedded multivalued dependencies and transitive dependencies, *in* "Proc. ACM SIGMOD," pp. 52–57.

Post, E. L. (1947), Recursive unsolvability of a problem of Thue, *J. Symb. Logic* **12**, 1–11.

Rauszer, C. (1987), Algebraic and logical description of functional and multivalued dependencies, *in* "Proc. 2nd Int. Symp. Methodology of Information Systems, Charlotte."

Reidemeister, K. (1968), "Grundlagen der Geometrie," Springer, Berlin.

Sagiv, Y., Delobel, C., Parker, D. S., and Fagin, R. (1981), An equivalence between relational database dependencies and a fragment of propositional logic, *J. Assoc. Comput. Mach.* **28**, 435–453.

Sagiv, Y., and Walecka, S. (1982), Subset dependencies and a completeness result for a subclass of embedded multivalued dependencies, *J. Assoc. Comput. Mach.* **29**(1), 103–117.

Shoenfield, J. R. (1967), "Mathematical Logic," Addison-Wesley, MA.

Skolem, Th. (1920), Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze nebst einem Theoreme über dichte Mengen, *Krist. Vid. Selsk. Skr.* **1**, 1–36.

Vardi, M. Y., (1984), The implication and finite implication problems for typed template dependencies, *J. Comput. System Sci.* **28**, 3–28.

von Neumann, J. (1960), "Continuous Geometry," Princeton Univ. Press, Princeton, NJ.

Zaniolo, C. (1976), "Analysis and Design of Schemata for Database Systems," Ph.D. Dissertation, Res. Report ENG-7669, Univ. of California at Los Angeles.