Theoretical Computer Science 29 (1984) 27–39
North-Holland

REPRESENTATION OF NONDETERMINISTIC INFORMATION

Ewa ORŁOWSKA and Zdzisław PAWLAK
Institute of Computer Science, Polish Academy of Sciences, 00-901 Warsaw, PKiN, Poland

Communicated by E. Engeler
Received October 1982
Revised March 1983

Abstract. In this paper we develop a method of dealing with nondeterministic information. We introduce the concept of knowledge representation system of nondeterministic information and we define a language providing a means for defining nondeterministic information. We also develop deduction methods for the language.

1. Introduction

The origin of knowledge representation methods has to do with the need to collect and process data related to a certain part of the reality, referred to as a universe of discourse. We assume that the universe of discourse consists of discrete objects. An object is anything which can be spoken of in the subject position of a natural language sentence (e.g., book, company). Objects need not be atomic or undivisible. They can be composed or structured, but are treated as a whole. Furthermore, we assume that we know a priori some interesting characteristics or properties which are meaningful for these objects. A property is denoted by a verb phrase in a natural language sentence (e.g., is interesting, is big). To express properties we use the notions of attribute (e.g., colour, height) and an attribute value (e.g., blue, tall). In general, information about values of attributes for objects is incomplete and therefore to some extent ambiguous. For example, we usually do not know precisely person's age, we can give its possible values only. In recent years several approaches have been taken with regard to representation of incomplete information [2, 3, 4, 5, 8].

The present paper is a contribution to the work in logical formalisms for representing incomplete knowledge. The notion of knowledge representation system of nondeterministic information introduced in the paper is a generalization of the notions of attribute based information systems introduced in [7, 8]. Information system presented in [7] consists of a set OB of objects, a set AT of attributes, a family $\{\mathrm{VAL}_a\}_{a\in \mathrm{AT}}$ of sets of values of attributes and an information function

$$f:\mathrm{OB}\times\mathrm{AT}\rightarrow\mathrm{VAL}=\bigcup_{a\in\mathrm{AT}}\mathrm{VAL}_a$$

such that for each $o \in \mathrm{OB}$ and each $a \in \mathrm{AT}$, $f(o, a) \in \mathrm{VAL}_a$. The data definition language for such systems was introduced in [6]. The generalization of these systems is a many-valued information system [8]. In many-valued systems we assume that $f \subseteq \mathrm{OB} \times \mathrm{AT} \times \mathrm{VAL}$ is not necessarily a function but an arbitrary relation such that if $(o, a, v) \in f$, then $v \in \mathrm{VAL}_a$. However, there are situations when the characteristics of given objects is determined neither by an information function, nor by an information relation. It might be a case that the only information we have for an object o and an attribute a is a set of possible values of a for o. To deal with such cases, Pawlak [8] introduced the notion of an approximate information system. In an approximate information system we consider information function f to be a function from set $\mathrm{OB} \times \mathrm{AT}$ into set $P(\mathrm{VAL})$ of all the subsets of set VAL such that $f(o, a) \in \mathrm{VAL}_a$, and, moreover, we assume that there exists a unique $v \in \mathrm{VAL}_a$ such that $f(o, a) = v$. In this paper we consider a generalization of many-valued and approximate information systems, called nondeterministic information systems.

2. System of nondeterministic information

By a system of nondeterministic information we mean a quadruple

$$S = (\mathrm{OB}, \mathrm{AT}, \{\mathrm{VAL}_a\}_{a \in \mathrm{AT}}, f)$$

where OB, AT and VAL_a, for each $a \in \mathrm{AT}$, are nonempty sets of objects, attributes and attribute values, respectively,

$$f : \mathrm{OB} \times \mathrm{AT} \to P(\mathrm{VAL}) \quad \text{where } \mathrm{VAL} = \bigcup_{a \in \mathrm{AT}} \mathrm{VAL}_a$$

is a total function such that $f(o, a) \subseteq \mathrm{VAL}_a$ for every $o \in \mathrm{OB}$ and $a \in \mathrm{AT}$.

The information function f does not specify a single value of an attribute for an object. With each object there is associated a set of possible values of every attribute. We do not specify how many values an attribute may take for a given object. Sets $f(o, a)$ are said to be generalized values of attribute a.

Consider, for example, a system of medical information. Let set OB of objects be a set of diseases, set AT of attributes be the set of some parameters of patient's body, e.g., temperature, blood tension, state of throat etc. Set VAL_a of values of parameter a is a set of possible values of that parameter. For example, $\mathrm{VAL}_{\mathrm{temperature}}$ is the set of elements of the interval 35°–42°. For a disease o and a parameter a the set $f(o, a)$ is the set of values of a which may occur during disease o.

Given a system S of nondeterministic information, we define binary relations of informational inclusion ($\mathrm{in}(S)$) and informational connection ($\mathrm{con}(S)$) in the set OB as follows:

$$(o, o') \in \mathrm{in}(S) \quad \text{iff} \quad f(o, a) \subseteq f(o', a) \quad \text{for all } a \in \mathrm{AT},$$

$$(o, o') \in \mathrm{con}(S) \quad \text{iff} \quad f(o, a) \cap f(o', a) \neq \emptyset \quad \text{for all } a \in \mathrm{AT}.$$

Hence, an object o is informationally included in object o' whenever for every attribute $a \in \mathrm{AT}$ the possible values of a for o are among the possible values of a for o'. For example, a disease o is informationally included in a disease o' if the symptoms of o occur during o', or, loosely speaking, if disease o' is accompanied by disease o, or if o' may be caused by o. Objects o and o' are informationally connected if for every attribute $a \in \mathrm{AT}$ the generalized values of a for o and o' have an element in common. Such objects can be considered to be similar with respect to the attributes of the given system.

The following properties of the relations $\mathrm{in}(S)$ and $\mathrm{con}(S)$ immediately follow from the definition.

Theorem 2.1. (a) *Relation* $\mathrm{in}(S)$ *is reflexive and transitive.*
(b) *Relation* $\mathrm{con}(S)$ *is reflexive and symmetric.*

In the next section we present a formal language whose formulas are schemes of sentences expressing properties of objects in systems of nondeterministic information. We develop a deductive system for the language based on axiomatization of propositional modal logics [1].

3. Logic NIL of nondeterministic information

To define formulas of the language of logic NIL we admit the following nonempty, at least denumerable, and pairwise disjoint sets of symbols:
- a set CONAT of constants representing attributes,
- a set CONGVAL of constants representing generalized values of attributes,
- a set $\{\neg, \vee, \wedge, \rightarrow, \leftrightarrow\}$ of classical sentential operations of negation, disjunction, conjunction, implication and equivalence, respectively,
- a set $\{\langle,\rangle, \Diamond, [,], \Box\}$ of unary modal sentential operations,
- a set $\{(,)\}$ of brackets.

FORNIL, the set of all formulae, is the least set satisfying the following conditions:

$(aV) \in$ FORNIL for any $a \in$ CONAT and $V \in$ CONGVAL
if $A, B \in$ FORNIL, then $\neg A, A \vee B, A \wedge B, A \rightarrow B, A \leftrightarrow B \in$ FORNIL
if $A \in$ FORNIL, then $\langle A,\rangle A, \Diamond A, [A,]A, \Box A \in$ FORNIL.

Formulae of the form (aV) are called nondeterministic descriptors. Let DESNIL denote the set of all nondeterministic descriptors.

Formulae are intended to be schemes of sentences providing definitions of sets of objects. For example, a formula of the form (aV) represents the set of those objects for which the set of possible values of attribute denoted by a coincides with the set corresponding to V. Modal operations enable us to express facts connected with informational inclusion and informational connection of objects. They provide a means for considering Boolean structure of families of generalized values of attributes. Formula $\langle(aV)$ represents the set of those objects which informationally include at least one object assuming V, as a value of a. In particular if we consider a system with the single attribute a, then this set coincides with the set of those objects o for which V is included in $f(o, a)$. Similarly, formula $\rangle(aV)$ corresponds to the set of objects which are informationally included in objects assuming V as a value of a. If a is the only attribute of a system, then this set coincides with the set of those objects o for which $f(o, a)$ is included in V. Formula $\Diamond(a, V)$ represents the set of objects which are informationally connected with some objects assuming V for a.

The semantics of the given language is defined by means of notions of model and satisfiability of formulas in a model. By a model we mean a system

$$M = (\mathrm{OB}, R, Q, m)$$

where

- OB is a nonempty set of objects,
- R is a reflexive and transitive relation in set OB,
- Q is a reflexive and symmetric relation in set OB,
- $m : \text{DESNIL} \to P(\mathrm{OB})$ is a meaning function assigning sets of objects to nondeterministic descriptors.

We say that an object $o \in \mathrm{OB}$ satisfies a formula A in a model M (M, o sat A) iff the following conditions are satisfied:

M, o sat (aV) iff $o \in m(aV)$,

M, o sat $\neg A$ iff not M, o sat A,

M, o sat $A \vee B$ iff M, o sat A or M, o sat B,

M, o sat $A \wedge B$ iff M, o sat A and M, o sat B,

M, o sat $A \to B$ iff M, o sat $\neg A \vee B$,

M, o sat $A \leftrightarrow B$ iff M, o sat $(A \to B) \wedge (B \to A)$,

M, o sat $\langle A$ iff there is an $o' \in \mathrm{OB}$ such that $(o', o) \in R$ and $M, o' \in$ sat A,

M, o sat $\rangle A$ iff there is an $o' \in \mathrm{OB}$ such that $(o, o') \in R$ and M, o' sat A,

M, o sat $\Diamond A$ iff there is an $o' \in \mathrm{OB}$ such that $(o, o') \in Q$ and M, o' sat A,

M, o sat $[A$ iff for all $o' \in \mathrm{OB}$ if $(o', o) \in R$, then M, o' sat A,

M, o sat $]A$ iff for all $o' \in \mathrm{OB}$ if $(o, o') \in R$, then M, o' sat A,

M, o sat $\Box A$ iff for all $o' \in \mathrm{OB}$ if $(o, o') \in Q$, then M, o' sat A.

Operations [,] and $\Box$ are dual with respect to $\langle, \rangle$ and $\Diamond$, respectively. They correspond to necessity operators in modal logics.

To each formula A of the language we assign the set $\mathrm{ext}_M A$ (extension of A in M) of those objects which satisfy a formula in a model:

$$\mathrm{ext}_M A = \{o \in \mathrm{OB}: M, o \text{ sat } A\}.$$

Theorem 3.1

(a) $\mathrm{ext}_M(aV) = m(aV)$.

(b) $\mathrm{ext}_M A = -\mathrm{ext}_M A$.

(c) $\mathrm{ext}_M A \vee B = \mathrm{ext}_M A \cup \mathrm{ext}_M B$.

(d) $\mathrm{ext}_M A \wedge B = \mathrm{ext}_M A \cap \mathrm{ext}_M B$.

(e) $\mathrm{ext}_M A \to B = -\mathrm{ext}_M A \cup \mathrm{ext}_M B$.

(f) $\mathrm{ext}_M A \leftrightarrow B = \mathrm{ext}_M A \cap \mathrm{ext}_M B \cup (-\mathrm{ext}_M A) \cap (-\mathrm{ext}_M B)$.

(g) $\mathrm{ext}_M \langle A = \{o \in \mathrm{OB}$: there is an $o' \in \mathrm{OB}$ such that $(o', o) \in R$ and $o' \in \mathrm{ext}_M A\}$.

(h) $\mathrm{ext}_M \rangle A = \{o \in \mathrm{OB}$: there is an $o' \in \mathrm{OB}$ such that $(o, o') \in R$ and $o' \in \mathrm{ext}_M A\}$.

(i) $\mathrm{ext}_M \Diamond A = \{o \in \mathrm{OB}$: there is an $o' \in \mathrm{OB}$ such that $(o, o') \in Q$ and $o' \in \mathrm{ext}_M A\}$.

(j) $\mathrm{ext}_M [A = \mathrm{ext}_M \neg \langle \neg A$.

(k) $\mathrm{ext}_M]A = \mathrm{ext}_M \neg \rangle \neg A$.

(ℓ) $\mathrm{ext}_M \Box A = \mathrm{ext}_M \neg \Diamond \neg A$.

We say that a formula A is true in a model M ($\models_M A$) iff $\mathrm{ext}_M A = \mathrm{OB}$. A formula A is valid ($\models A$) iff it is true in every model. A set T of formulas is satisfied by an object o in a model M (M, o sat T) iff M, o sat A for every formula $A \in T$. A set T is satisfiable iff there exists a model M and an object o such that M, o sat T. A formula A is a semantical consequence of a set T of formulas ($T \models A$) iff M, o sat A whenever M, o sat T for every model M and for every object o from the set of objects of M.

We admit the following axioms and inference rules for the logic NIL.

Axioms

(A1) All formulas having the form of tautologies of the classical propositional calculus.

(A2) $[(A \to B) \to ([A \to [B)$.

(A3) $](A \to B) \to (]A \to]B)$.

(A4) $\Box(A \to B) \to (\Box A \to \Box B)$.

(A5) $A \to]\langle A$.

(A6) $A \to [\rangle A$.

(A7) $]A \to A$.

(A8) $\Box A \to A$.

(A9) $]A \to]]A$.

(A10) $A \to \Box\Diamond A$.

Axioms (A2), (A3) and (A4) assure that logic NIL is a normal modal logic. Axioms (A5) and (A6) show that operation $\langle$ is inverse with respect to operation $\rangle$. Axioms (A7) and (A8) provide reflexivity of relations R and Q, respectively. Axioms (A9) and (A10) provide transitivity of relation R and symmetry of relation Q, respectively.

Rules of inference

(R1) $\dfrac{A, A \to B}{B}$ (R3) $\dfrac{A}{]A}$

(R2) $\dfrac{A}{[A}$ (R4) $\dfrac{A}{\Box A}$

Rules (R2), (R3) and (R4) are counterparts of the necessity rule in modal logics.

The given axioms and rules characterize the operations $\neg$, $\to$, $[$, $]$ and $\Box$ only, but it is sufficient due to Theorem 3.1(f), (g), (k), (l), and the following theorem.

Theorem 3.2

(a) $\text{ext}_M\, A \vee B = \text{ext}_M\, \neg A \to B$.

(b) $\text{ext}_M\, A \wedge B = \text{ext}_M\, \neg(A \to \neg B)$.

We say that a formula A is derivable from a set T of formulas ($T \vdash A$) iff it is obtainable from the axioms and the formulas from T by repeated application of inference rules. A formula A is said to be a theorem of logic NIL ($\vdash A$) iff it is derivable merely from the axioms. A set T of formulas is consistent if a formula of the form $A \wedge \neg A$ is not derivable from T.

Theorem 3.3. (Soundness theorem). (a) $\vdash A$ *implies* $\models A$.

(b) $T \vdash A$ *implies* $T \models A$.

(c) *T satisfiable implies T consistent.*

Proof. The axioms of NIL are easily seen to be valid, and rules clearly preserve validity. This proves (a) from which (b) and (c) immediately follow.

Examples of theorems of logic NIL are presented below.

Theorem 3.4

(a) $\vdash A \to \langle A$.

(b) $\vdash \langle (A \vee B) \leftrightarrow (\langle A \vee \langle B)$.

(c) $\vdash \langle (A \wedge B) \to (\langle A \wedge \langle B)$.

(d) $\vdash [(A \wedge B) \leftrightarrow ([A \wedge [B)$.

(e) $\vdash [A \to \langle [A$.

(f) $\vdash \langle [A \to A$.

(g) $\vdash \neg [A \leftrightarrow \langle \neg A$.

(h) $\vdash \neg]A \leftrightarrow \rangle \neg A$.

(i) $\vdash \Diamond \Box A \to A$.

(j) $\vdash \neg \Box A \leftrightarrow \Diamond \neg A$.

(k) $\vdash]A \wedge \langle B \to \langle (A \wedge B)$.

(ℓ) $\vdash]A \wedge \rangle B \to \rangle (A \wedge B)$.

Theorems related to operations $\rangle$ and $]$ are analogous to (a)–(f).

In the following, a completeness theorem for logic NIL will be presented. Let T be a consistent set of formulas and let relation $\approx$ in set FORNIL be defined as follows:

$$A \approx B \quad \text{iff} \quad T \vdash A \leftrightarrow B.$$

Theorem 3.5. (a) *Relation $\approx$ is an equivalence on set* FORNIL.

(b) *Relation $\approx$ is a congruence with respect to $\neg$, $\vee$ and $\wedge$.*

(c) *If $A \approx B$, then $[A \approx [B$, $]A \approx]B$ and $\Box A \approx \Box B$.*

Proof. The proof of conditions (a) and (b) is the same as for the classical propositional logic [9]. Condition (c) follows from Axioms (A2), (A3), (A4) and necessity rules.

We construct the quotient algebra

$$\text{ANIL} = (\text{FORNIL}|_{\approx}, -, \cup, \cap, 1, 0)$$

where FORNIL$|_{\approx}$ is the set of the equivalence classes $[A]$ of relation $\approx$ for all formulas A,

$$-[A] = [\neg A] \qquad 1 = [A \vee \neg A]$$

$$[A] \cup [B] = [A \vee B] \qquad 0 = [A \wedge \neg A]$$

$$[A] \cap [B] = [A \wedge B]$$

Theorem 3.6. (a) *Algebra* ANIL *is a nondegenerate Boolean algebra.*

(b) $[A] \leqslant [B]$ *iff* $T \vdash [A \to B]$.

(c) $T \vdash A$ *iff* $[A] = 1$.

(d) $[\neg A] \neq 0$ *iff not* $T \vdash A$.

Let $\mathcal{F}$ be the family of all the maximal filters in algebra ANIL. Set $\mathcal{F}$ is nonempty since the algebra is nondegenerate. We define relation $R_0 \subseteq \mathcal{F} \times \mathcal{F}$ as follows:

$(F, G) \in R_0$ iff for any formula A if $[]A] \in F$, then $[A] \in G$.

Theorem 3.7. *The following conditions are equivalent*:

(a) $(F, G) \in R_0$.

(b) *If* $[[A] \in G$, *then* $[A] \in F$.

(c) *If* $[A] \in F$, *then* $[\langle A] \in G$.

(d) *If* $[A] \in G$, *then* $[\rangle A] \in F$.

Proof. Assume condition (a), and suppose that $[]A] \in G$ and $[A] \notin F$. It follows that $[\neg A] \in F$ and, by (A5), $[]\langle \neg A] \in F$. By (a) we obtain $[\langle \neg A] \in G$. By Theorem 3.4(k) we have $[\langle (A \wedge \neg A)] \in G$, but G is a proper filter, a contradiction. Hence condition (b) holds.

Let us now assume that condition (b) holds and suppose that $[A] \in F$ and $[\langle A] \notin G$. Hence $[[\neg A] \in G$ and by (b) we have $[\neg A] \in F$, a contradiction. Hence condition (c) holds.

Assume condition (c) and suppose that $[A] \in G$ and $[\rangle A] \notin F$. Then $[\neg \rangle A] \in F$ and by (c) we have $[\langle \neg A] \in G$. By (A6), $[\neg A] \in G$, a contradiction. Hence condition (d) holds.

We also have (d) implies (a). For suppose not, then $[\rangle A] \in G$, and by (d), $[\rangle \neg A] \in F$. By Theorem 3.4(ℓ) we have $[\rangle (A \wedge \neg A)] \in F$, a contradiction.

Theorem 3.8. *Relation* R_0 *is reflexive and transitive.*

Proof. The proof follows from (A7) and (A9).

We define a relation $Q_0 \subseteq \mathcal{F} \times \mathcal{F}$ as follows:

$(F, G) \in Q_0$ iff for any formula A if $[\Box A] \in F$, then $[A] \in G$.

Theorem 3.9. *Relation* Q_0 *is reflexive and symmetric.*

Proof. The proof follows from (A8) and (A10).

Theorem 3.10. (a) *If* $[\rangle A] \in F$, *then there exists a* $G \in \mathcal{F}$ *such that* $(F, G) \in R_0$ *and* $[A] \in G$.

(b) *If* $[\langle A] \in F$, *then there exists a* $G \in \mathcal{F}$ *such that* $(G, F) \in R_0$ *and* $[A] \in G$.

(c) *If* $[\Diamond A] \in F$, *then there exists a* $G \in \mathcal{F}$ *such that* $(F, G) \in Q_0$ *and* $[A] \in G$.

Proof. Let $[\rangle A] \in F$ and consider set $X_F = \{[B] : []B] \in F\}$. Set X_F is nonempty since

$1 \in X_F$. Consider filter F' generated by set $X_F \cup \{[A]\}$. We have $F' = \{[B]$: there exist $[A_1], \ldots, [A_n] \in X_F$, $n \geq 1$, such that $[A_1] \cap \cdots \cap [A_n] \cap [A] \leq [B]\}$. We shall show that for any $[A_1], \ldots, [A_n] \in X_F$ we have $[A_1] \cap \cdots \cap [A_n] \cap [A] \neq 0$. Suppose then, conversely, $T \vdash A_1 \wedge \cdots \wedge A_n \rightarrow A$. By (A3) and (R3) we have $T \vdash](A_1 \wedge \cdots \wedge A_n) \rightarrow]\neg A$. Since $[]A_1], \ldots, []A_n] \in F$, we have $[]A_1 \wedge \cdots \wedge]A_n] \in F$. Since $\vdash]A \wedge]B \leftrightarrow](A \wedge B)$, we have $[](A_1 \wedge \cdots \wedge A_n)] \in F$. Hence $[]\neg A] \in F$, so $[\neg\rangle A] \in F$, what contradicts the assumption. Thus, filter F' is proper. Let G be the maximal filter containing F'. We clearly have $[A] \in G$ and $(F, G) \in R_0$. Hence condition (a) is satisfied. The proof of conditions (b) and (c) is similar.

We define a canonical model M_0 as follows:

$$M_0 = (\mathrm{OB}_0, R_0, Q_0, m_0)$$

where

- $\mathrm{OB}_0 = \mathcal{F}$,
- R_0 and Q_0 are relations defined above,
- $F \in m_0(aV)$ iff $[(aV)] \in F$.

Theorem 3.11. *The following conditions are equivalent*:

(a) M_0, F sat A.

(b) $[A] \in F$.

Proof. If A is of the form (aV), then the theorem holds by the definition of meaning function m_0 in the canonical model. If A is of the form $\neg B$ or $B \rightarrow C$, we use the definition of satisfiability and the fact that filter F is maximal and prime. If A is of the form $\langle B$ or $\rangle B$, then the theorem follows from Theorems 3.7 and 3.10(a) and (b). If A is of the form $\Diamond B$, then we use Theorem 3.10(c). Now, consider a formula of the form $]A$ and suppose that M_0, F sat $]A$ and $[]A] \notin F$. Hence $[\rangle\neg A] \in F$ and M_0, F sat $\rangle\neg A$. Thus M_0, F sat $\neg]A$, a contradiction. Now assume hat $[]A] \in F$ and consider set $X_F = \{[B]: []B] \in F\}$. We have $[A] \in X_F$. Moreover, set X_F is a filter, since we have $[B]$ and $[C] \in X_F$ iff $[B] \cap [C] = [B \wedge C] \in X_F$ for any formulas B and C. Set X_F is a proper filter, since $0 \notin X_F$. By the Kuratowski–Zorn lemma there is a maximal filter G such that $(F, G) \in R_0$ and $[A] \in G$. But X_F is contained in every filter G such that $(F, G) \in R_0$, thus $[A]$ belongs to every such filter. By the induction hypothesis we have M_0, G sat A for all G satisfying $(F, G) \in R_0$. Hence M_0, F sat $]A$. For formulas of the form $[A$ and $\Box A$ the proof is similar and uses Theorem 3.4(h) and (j).

Theorem 3.11 enables us to prove completeness and compactness of logic NIL.

Theorem 3.12 (Completeness theorem). (a) $\models A$ *implies* $\vdash A$.

(b) $T \models A$ *implies* $T \vdash A$.

(c) *T consistent implies T satisfiable.*

Proof. We now prove condition (b). Suppose that not $T \vdash A$. By Theorem 3.6(d) we have $[\neg A] \neq 0$. Thus there is a maximal filter $F_0 \in \mathcal{F}$ such that $[\neg A] \in F_0$. By Theorem 3.11 we have M_0, F_0 sat $\neg A$. For any formula $B \in T$ we have $T \vdash B$ by Theorem 3.6(c). Hence $[B] \in F_0$ and, by Theorem 3.11, M_0, F_0 sat B, a contradiction. Condition (a) follows from (b), and condition (c) follows from Theorem 3.11.

As a corollary we obtain the following theorem.

Theorem 3.13 (Compactness theorem). *The following conditions are equivalent*:
(a) *T is satisfiable.*
(b) *Every finite subset of T is satisfiable.*

Deductive methods based on logic NIL enable us to determine for any formula expressing a property of objects whether it is implied by some other formulas. In NIL all the tautologies of classical logic are valid and hence its deductive power is not less than that of the classical logic. The modal operations enable us to reason in the presence of nondeterminism understood as indefiniteness of information about objects. These operations enable us to penetrate in a sense a Boolean structure of families of generalized values of attributes. In the next section we discuss languages of systems of nondeterministic information based on the logic NIL.

4. Languages of systems of nondeterministic information

Let $S = (\mathrm{OB}, \mathrm{AT}, \mathrm{VAL}, f)$ be given, and let $\mathrm{in}(S)$ and $\mathrm{con}(S)$ be, respectively, relations of informational inclusion and informational connection determined by system S. Let n be the function

$$n : \mathrm{CONAT} \cup \mathrm{CONGVAL} \to \mathrm{AT} \cup P(\mathrm{VAL})$$

such that
- $n(\mathrm{CONAT}) = \mathrm{AT}$,
– the range of function f is included in $n(\mathrm{CONGVAL})$.

We consider model

$$M = (\mathrm{OB}, \mathrm{in}(S), \mathrm{con}(S), m) \quad \text{where } m(aV) = \{o : f(o, n(a)) = n(V)\}.$$

Next, the set $\mathrm{FORNIL}(M)$ of formulae of system S is the least set containing all pairs of the form

$$(n(a)n(V)) \quad \text{for any } a \in \mathrm{CONAT} \text{ and } V \in \mathrm{CONGVAL}$$

and closed with respect to operations $\neg$, $\vee$, $\wedge$, $\to$, $\leftrightarrow$, $\langle$, $\rangle$, $\Diamond$, $[$, $]$, $\Box$.

In a natural way we define satisfiability of formulae of system S by objects of the system, and extensions of formulae, namely

$$o \operatorname{sat}(n(a)n(V)) \quad \text{iff} \quad M, o \operatorname{sat}(aV),$$

$$\operatorname{ext}(n(a)n(V)) = \operatorname{ext}_M(aV).$$

For compound formulae the respective inductive definitions are analogous to those presented in Section 3.

A formula $A \in \text{FORNIL}(M)$ is true iff ext $A = \text{OB}$.

By using formulae from set FORNIL(M) we can express many important properties of sets of objects.

Theorem 4.1. (a) *$A \rightarrow B$ is true iff* ext $A \subseteq$ ext B.

(b) *$A \leftrightarrow B$ is true iff* ext $A =$ ext B.

(c) *$\neg A$ is true iff* ext $A = \emptyset$.

In the following we present some properties specific for nondeterministic information.

Theorem 4.2. (a) *If $\langle(n(a)n(V))$ is true, then $n(V) \subseteq f(o, n(a))$ for all $o \in \text{OB}$.*

(b) *If $\rangle(n(a)n(V))$ is true, then $f(o, n(a)) \subseteq n(V)$ for all $o \in \text{OB}$.*

(c) If *$\diamond(n(a)n(V))$ is true, then $f(o, n(a)) \cap n(V) \neq \emptyset$ for all $o \in \text{OB}$.*

Formula $\langle(n(a)n(V))$ is true iff each object o in a given system has associated with it a certain object o' which is informationally included in o and assumes generalized value $n(V)$ of attribute $n(a)$. It follows that $n(V)$ is a subset of a generalized value of attribute $n(a)$ for object o. In a similar way it can be easily seen that conditions (b) and (c) hold.

Theorem 4.3. *For any system such that* $\text{AT} = \{a\}$ *the following conditions are satisfied*:

(a) *If* ext $\langle(n(a)n(V)) \neq \emptyset$, *then there is an object assuming generalized value $n(V)$ for attribute $n(a)$ and it is possible that there are objects assuming supersets of $n(V)$ for $n(a)$.*

(b) *If* ext $\rangle(n(a)n(V)) \neq \emptyset$, *then there is an object assuming $n(V)$ for $n(a)$ and it is possible that there are objects assuming subsets of $n(V)$ for $n(a)$.*

(c) *If* ext $[(n(a)n(V)) \neq \emptyset$, *then there is an object assuming $n(V)$ for $n(a)$ and there are no objects assuming supersets of $n(V)$ for $n(a)$.*

(d) *If* ext $](n(a)n(V)) \neq \emptyset$, *then there is an object assuming $n(V)$ for $n(a)$ and there are no objects assuming subsets of $n(V)$ for $n(a)$.*

Let us consider the following system of medical information:

- $\text{OB} = \{D1, \ldots, D6\}$ is a set of diseases,
- $\text{AT} = \{a1, a2\}$ is a set of symptoms occurring during diseases from OB,

– $\mathrm{VAL}_{a1} = \{v1, v2, v3, v4, v5\}$,
– $\mathrm{VAL}_{a2} = \{u1, u2, u3\}$,
– $\mathrm{VAL} = \mathrm{VAL}_{a1} \cup \mathrm{VAL}_{a2}$,
– $f: \mathrm{OB} \times \mathrm{AT} \to P(\mathrm{VAL})$ is given by the following table:

	$a1$	$a2$
$D1$	$\{v1, v3\}$	$\{u1, u2, u3\}$
$D2$	$\{v2, v5\}$	$\{u1\}$
$D3$	$\{v1, v3, v4\}$	$\{u1, u2\}$
$D4$	$\{v1\}$	$\{u1, u2\}$
$D5$	$\{v1, v3\}$	$\{u1\}$
$D6$	$\{v5\}$	$\{u1\}$

The relation of informational inclusion of the given system consists of the following pairs of diseases:

All pairs (Di, Di) for $i = 1, \ldots, 6$,
$(D4, D1)$ $(D5, D1)$ $(D6, D2)$ $(D4, D3)$ $(D5, D3)$.

In the following we list extensions of some formulae of the language of the system and we give their intuitive interpretation:

ext $\langle (a1\{v1\}) = \{D1, D3, D4\}$: Diseases $D1$, $D3$ and $D4$ can be caused by a disease in which symptom $a1$ assumes value $v1$; in other words if a patient suffers from one of diseases $D1$, $D3$ or $D4$, then sometime in the past he (she) possibly suffered from a disease satisfying $(a1\{v1\})$.

ext $\rangle (a1\{v1, v3, v4\}) = \{D3, D4, D5\}$: Diseases $D3$, $D4$ and $D5$ are possibly followed by a disease in which possible values of $a1$ are among $v1$, $v3$ and $v4$; or if a patient suffers from $D3$, $D4$ or $D5$, then sometime in the future he (she) will possibly suffer from a disease satisfying $(a1\{v1, v3, v4\})$.

ext $[(a2\{u1\}) = \{D2, D6\}$: Each disease causing $D2$ or $D6$ assumes value $u1$ of symptom $a2$.

ext $](a2\{u1, u2\}) = \{D3\}$: Each disease caused by $D3$ assumes $u1$ or $u2$ for symptom $a2$.

Let us observe that

$$\text{ext}\,\langle (a1\{v3\}) = \emptyset,$$

since in our system there is no object which assumes generalized value $\{v3\}$ of attribute $a1$. This means that although in our system $\{v3\}$ is a subset of generalized values of $a1$ for diseases $D1$, $D3$ and $D5$, knowledge given in the system does not enable us to point out a disease which satisfies $(a1\{v3\})$ and possibly causes diseases $D1$, $D3$ or $D5$.

The relation of informational connection of the given system consists of the following pairs:

All pairs (Di, Di) for $i = 1, \ldots, 6$
$(D1, D3)$ $(D1, D4)$ $(D1, D5)$ $(D2, D6)$ $(D3, D4)$
$(D3, D5)$ $(D4, D5)$.

All pairs (Di, Dj) for (Dj, Di) given above.

Consider, for example, the following extensions:

ext $\Diamond(a1\{v1, v3\}) = \{D1, D3, D4, D5\}$: For diseases $D1$, $D3$, $D4$ and $D5$ there are diseases informationally connected with them which may take $v1$ or $v2$ as the values of symptom $a1$.

ext $\Box(a2\{u1\}) = \{D2, D6\}$: All diseases similar to $D2$ or $D6$ in the sense of informational connection may assume value $u1$ for attribute $a2$.

By using modal operations of the language we can express those relationships between objects of a system which are determined by the algebraic structure of families of generalized values of attributes. Although these relationships are not stated explicitly in the system, they are given implicitly by the choice of generalized values of attributes. The presented language provides a means for accessing this kind of information.

References

[1] D.M. Gabbay, *Investigations in Modal and Tense Logics with Applications to Problems in Philosophy and Linguistics*, Reidel Synthese Library **92** (Reidel, Dordrecht, 1976).

[2] J. Grant, Null values in a relational data base, *Inform. Process. Lett.* **5** (1977) 156–157.

[3] P. Hajek, Automatic listing of important observational statements III, *Kybernetika* **10** (1974) 95–124.

[4] M. Jaegermann, Information storage and retrieval systems with incomplete information, Part I: Fundamenta, *Informaticae* **2** (1978) 17–41; Part II: Fundamenta, *Informaticae* **2** (1979) 141–166.

[5] W. Lipski, On semantic issues connected with incomplete information databases, *ACM Trans. Database Systems* **4** (1979) 262–296.

[6] W. Marek and Z. Pawlak, Information storage and retrieval systems, mathematical foundations, *Theoret. Comput. Sci.* **1** (1976) 331–354.

[7] Z. Pawlak, Information systems—Theoretical foundations, *Information Systems* **6** (1981) 205–218.

[8] Z. Pawlak, *Systemy Informacyjne. Podstawy Teoretyczne* (Information Systems. Theoretical Foundations), (WNT, Warsaw) (in print).

[9] H. Rasiowa and R. Sikorski, *Mathematics of Metamathematics* (PWN, Warsaw, 1970).