MTA SZTAKI Közlemények 39/1988 p. 53-66.

A DEDUCTIVE LANGUAGE FOR THE REPRESENTATION OF INCOMPLETE KNOWLEDGE

HA HOANG HOP

Institute of Computer Science and Cybernetics
Hanoi - Vietnam

Abstract We give here a formal treament of the problem of representing incomplete knowledge by a new approach in which formal knowledge bases are described in terms of abstract interactive operations between them and expert systems /ESs/. We then setle up an original knowledge representation language which allows the explicit representation of nonmonotonicity in the framework of beliefs and knowledges. The correctness of this language is proved.

1. Introduction Our formalisation of RL - the representation language - is started with McDermott and Doyle /1980/, Moore /1983/, Levesque /1984/. Levesque s paper provides much technical inspirations. We generalise Levesque s knowledge representation language by two foundamental aspects: all interactive operations are abstract instead of only as "Tell" and "Ask", and beliefs are more explicitly represented by means of simultaneous two algebraic operators K "It is known" and L "It is believed". The first aspect allows a high level of conceptual modeling, since with two operations "Tell" and "Ask", formal knowledge bases would have the risk of colliding to traditional databases in which well-formed formulae are only considered in part; the second aspect enables a more explicit representation of beliefs, as we explicitly specify in RL the second algebraic operator L - we distinguish belief and knowledge, this distinction seems to be essential when dealing with the Generalised Closed-World Assumption /Reiter -1978, Bossu and Siegel-1985, Przymusinska and Gelfond-1986/

We give an approach to above problem by a new manner: turning reasoning steps into so-called knowledge matrices and then showing the correctness of RL.

2. The knowledge representation language RL

2.1. Concepts, definitions

We buil up RL on the basis of the following sets:

W, KB, F, R

where

- W is a non-empty set of possible worlds,
- KB is a set of preliminary knowledges about W. In the following, KB is treated as an abstract knowledge base.
- F is the set of all formulae of FOL and all rmulae of the form:

/1/ $p \wedge Kq_1 \wedge \ldots \wedge Kq_n \supset r$ or simpler $Kq_1 \wedge \ldots \wedge Kq_n \supset r$

/2/ $p \wedge Lq1 \wedge \ldots \wedge Lq_n \supset r$ or simpler $Lq_1 \wedge \ldots \wedge Lq_n \supset r$

where p, q_i $(i=1,\ldots,n)$, r belong to FOL. K is quoted as "it is known" and L is as "it is believed".

- R is a set of specific representations so that each $r \in R$ transforms an arbitrary pair (w,d) where $w \in W$ and $d \in F$ into an element of

$$C = \{ \text{yes, no, unknown, known, believed} \}$$

Definition 2.1.1. Axioms for FOL proof theory:

/a1/ $p \supset (q \supset r)$.

/a2/ $(p \supset (q \supset r)) \supset ((p \supset q) \supset r$.

/a3/ $(\sim q \supset \sim p) \supset ((\sim q \supset \sim p) \supset q$.

/a4/ $\forall x\,(p \supset q) \supset (\forall x p \supset \forall x q)$.

/a5/ $\forall x p \supset p^x_t$.

Axiom of Equality:

/ae/ $(i=i) \wedge (i \neq j)$ for all distinct i, j p, q, r are formulae of FOL, x is free variable, t is closed term and i, j are indexes.

Comment 2.1.2. Since KB is defined as preliminary knowledges from W, it puts equivalence to that KB is incomplete.

With the algebraic operator K, we have Levesque s query language in which a formula of the form Kp is read as

The KB knows p /ak/

or

KB $\Vdash$ p

where $\Vdash$ is a specified provability relation and p is any formula of FOL possibly containing K's.

The second algebraic operator L is as

The KB believes p /ab/

or

KB $\VDash$ p

where $\VDash$ is a new form of provability /or query evaluation/

While the first query evaluation $\Vdash$ is understood as usually, the second needs an exact semantics.

2.2. Semantics for RL

The language RL has all formation rules of FOL and the following two rules:

If p $\in$ RL then Kp $\in$ RL /ak/
If p $\in$ RL then Lp $\in$ RL /ab/

So, in F there is three kinds of formula:

/i/ $p \wedge q$
/ii/ $K(p \wedge q)$
/iii/ $L(p \wedge q)$

/i/ will be true or false depending on the interpretation of predicate symbols.
/ii/ will be true or false depending on KB and on what is known or unknown.
/iii/ will be true or false depending on KB and on what is believed.

Semantics of RL will be depended on the set of possible worlds W. We use here Kripke s interpretation. Kripke /1963/ uses the concept of possible worlds to create a formal semantics for modal logic. Later, mathematical logicians, e.g., Chang and Keisler /1973/, Kalish et al./1980/, equate the concept of possible worlds with the model for a formal language of FOL. In another development, linguists and philosophical logicians,e.g., Cresswell /1973/, Rescher /1975/, seem to regard possible worlds more broadly, as a kind of Gestalt experiments, not limited by the vocabulary of the language of FOL or any others. Our usage of possible worlds here will be more on mathematical side,i.e., that a possible world is an alternative model, but the philosophical side remains valid.

Definition 2.2.1. Kp is true in KB iff p is true in every possible world of W.
Definition 2.2.2. Lp is true in KB iff p is true in every possible world of W.

In agreement with McDermott /1982/ s argument, we need two inference rules: modus ponens and necessitation. We come to the axiom schemata for RL.

Definition 2.2.3. Axiom Schemata for The Knowledge Representation Language RL

. The axioms of FOL

- Kp where p is an axiom of FOL
- Lp where p is an axiom of FOL
- $K(p \supset q) \supset (Kp \supset Kq)$
- $(\forall x)\, Kp \supset K(\forall x)\, p$
- $L(p \supset q) \supset (Lp \supset Lq)$
- $(\forall x)\, Lp \supset L(\forall x)\, p$
- $p \equiv Kp$ where p is pure.

  / a formula is pure when it is known /

Definition 2.2.4. Monotonic inference rules for RL

- $p,\ p \supset q \vdash q$ /modus ponens/
- $p \vdash Kp$ /K-necessitation/
- $p \vdash Lp$ /L-necessitation/

Comment 2.2.5. On the basis of the notion of possible worlds, K and L are treated in an unified way: definitions 2.2.1 and 2.2.2 have pointed out this unification, thus in the above definition, we may use $\vdash$ instead of $\Vdash$ and $\models$ without confusions. The semantics for RL is both sound and complete since RL is both sound and complete with respect to Levesque /1984/ and Hop /1987/. Due to the limited space, we do not quote the proofs here.

Theorem 2.2.6. RL is both sound and complete

3. The correctness of RL

In this section, we consider more concretely what RL will be with abstract interactive oprations. We define exactly interactive operations by terms of specific representation, then gradually, the so-called knowledge matrix is constructed that serves for the proof of the correctness of RL - this proof in turns, is strongly not only supports for theorem 2.2.6, but opens new possibilities for further investigations.

3.1. Definitions

Definition 3.1.1. $r \in R$ is defined as follows:

$$r: KB \times RL \longrightarrow \{\text{yes, no, known, unknown, believed}\}$$

so that

$$r(k, p) = \begin{cases} \text{yes, if } Kp \text{ true on } k \\ \text{no, if } K\ p \text{ true on } k \\ \text{believed, if } Lp \text{ true on } k \\ \text{known, if } p \text{ is pure} \\ \text{unknown, otherwise} \end{cases}$$

Definition 3.1.2. The quantification of R:

$$Q\big(r(k,p)\big) \rightarrow \{0, \Delta_1, \Delta_2, \Delta_3, 1\}$$

so that

$$Q(r) = \begin{cases} 1 & \text{if } r(k,p) = \text{yes} \\ 0 & \text{if } r(k,p) = \text{no} \\ \Delta_1 & \text{if } r(k,p) = \text{believed} \\ \Delta_2 & \text{if } r(k,p) = \text{known} \\ \Delta_3 & \text{if } r(k,p) = \text{unknown} \end{cases}$$

where $0 \leqslant \Delta_3 < \Delta_2 \leqslant \Delta_1 \leqslant 1.$

Definition 3.1.3. The classification of algebraic/modal formulae:

/i/ if p is formula, then so Kp, Lp.

/ii/ Classification listing for all formulae / with or without algebraic/modal operators /

α	α_1	α_2
$(p \wedge q, 1)$	$(p, 1)$	$(q, 1)$
$(p \vee q, 0)$	$(p, 0)$	$(q, 0)$
$(p \supset q, 0)$	$(p, 1)$	$(q, 0)$
$(\sim p, 1)$	$(p, 1)$	$(q, 1)$
$(\sim p, 0)$	$(p, 0)$	$(q, 0)$

β	β_1	β_2
$(p \wedge q, 0)$	$(p, 0)$	$(q, 0)$
$(p \vee q, 1)$	$(p, 1)$	$(q, 1)$
$(p \supset q, 1)$	$(p, 0)$	$(q, 1)$

ν	ν_0
Kp, 1	p, 1
Lp, 0	p, 0

π	π_0
Kp, 0	p, 0
Lp, 1	p, 1

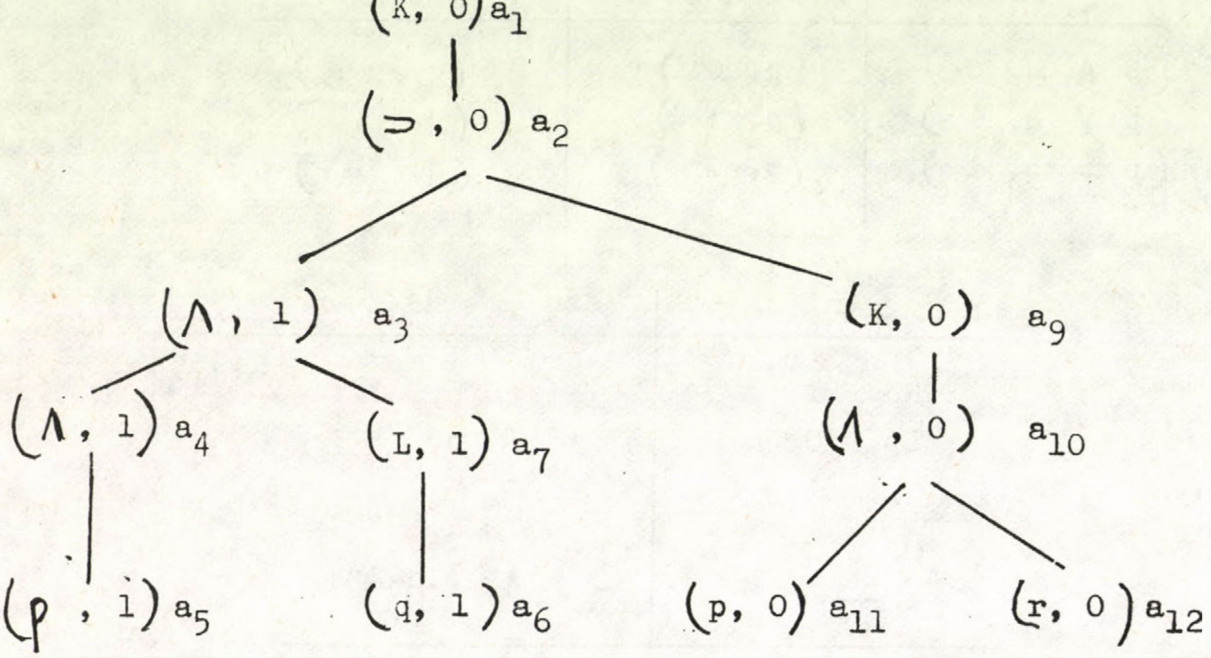
where $\alpha, \alpha_1, \beta, \beta_1, \beta_2, \alpha_2$ --- are used to denote signed formulae and their components of respective types. This classification is a modification of Hop /1986/ for temporal case, however, we give here only a classificatio which relates to classical quantification with two values 0 and 1 - this treatment is enough for the matrix representation that we will present belows.

Definition 3.1.4. The formula tree for a signed modal formula is a formation tree containing additional information as to the polarity of its atomic formulae.

Example 3.1.5. The formula tree for the formula

$$(K(p \wedge q) \wedge Lp \supset L(p \wedge r), 0)$$

is



the label at each node is index of the path in the tree, which provides the following matrix representation of formulae

Definition 3.1.6. A path through a formula tree is a subset of its formula tree. Denote these paths s and t, and s u for path s with an occurence of the label u .

A path through a formula tree is called atomic iff for every node k in s, either

/a/ k is a label by an atomic formula, or

/b/ k is a v in the classification tables.

The atomic paths through an quantified formula is detected by wringting components of an α-type side by side and the components of β-type one above the other to form a nested matrix.

Example 3.1.7. The matrix representation of the formula in example 3.1.5:

$$-- K - \left(--\left(-p \wedge q-\right)---\left(-Lr-\right)\right) \supset --L-- \left(\begin{pmatrix} p \\ \wedge \\ r- \end{pmatrix} -- \begin{pmatrix} p \\ \wedge \\ -r- \end{pmatrix}\right) ----$$

one of four atomic paths is a dotted line in the figure.

With $r \in R$, a pair k,p is transformed into an value in the set $\{ 0, 1, \Delta_1, \Delta_2, \Delta_3 \}$. It in turn, is transformed by quantification Q into one of the categories in classification tables, and at last, comes into the form of matrix representation above. We have the following definition.

Definition 3.1.8. Through r, Q and matrix representation, we obtain matrix $\| x_{ij} \|_{qxm}$ forming by elements: 0, 1, Δ_1, Δ_2, Δ_3, where q is cardinality /if any/ of p in RL and m is cardinality of KB /if any/. Thus, in the simplest case, KB can be represented as a matrix, and in general, it can be represented as matrices. We call this matrix /or matrices/ knowledge matrix /or matrices /.

We obtain the Theorem belows which points out the existence of specific representation for every interactive operation between KB and expert systems. This theorem actually shows the correctness of every operation between KB and their abstract expert systems.

Theorem 3.1.9. For every $r \in R$, there exists two other specific representations $r_1, r_2 \in R$ so that $r = r_1 \circ r_2$, where "o" is a consequent application of r_1 and r_2, i.e., a matrix multiply, and

$$r_1\ (k,p) = \|k_{ij}\|_{q \times m}\ ,\ r_2(\|k_{ij}\|_{q \times m}) = \|x_{ij}\|_{qxm}$$

$$x_{ij} \in \{0,\ 1,\ \Delta_1,\ \Delta_2,\ \Delta_3\}$$

Proof. Let r_3 be a specific representation that transforms $\|x_{ij}\|_{qxm}$ to a quantified matrix $\|y_{ij}\|_{qxm}$ / this situation is realistic since we can put $x_{ij} = y_{ij}$ if $x_{ij} \in \{0, 1\}$, and $y_{ij} = 1/2$ if $x_{ij} \in \{\Delta_1, \Delta_2, \Delta_3\}$ /. Thus by r_3^{-1} we denote the reverse representation from $\|y_{ij}\|_{qxm}$ to $\|x_{ij}\|_{qxm}$.

Put $r_1 = r_1 \circ r_3,\ r_2 = r_3^{-1}$ we have

$$r = r_1 \circ r_2 = r_1 \circ r_3 \circ r_3^{-1} = r.$$

The Theorem is proved.

4. Conclusion. With RL, we can resolve the following problems:

. The interaction between an abstract expert system with an incomplete knowledge base by non-monotonic manner,

. Formal semantics, a Kripke semantics-based consideration is explicitly investigated,

. For the first time, we propose the matrix representation for the problem of constructing up an knowledge representation language, this treatment shows that even with extended power provided by RL, knowledge of an incomplete knowledge base is still representable at symbol level, and even at number level by means of two modal operators in the framework of matrix representation. This means that whenever we want to have the "returns" to traditional databases, it will be quite possible.

Recently almost all contributions to the field of knowledge representation are deductive, i.e., are based on various deductive reasoning models. It means that the logical lattice of deduction models are supposed closed. The closure property is presented inTheorem 3.1.9. Closures are abstractly and generally investigated by Khang /1978/, thus it seems to enable an unified approach to pattern recognition and knowledge representation simultaneously.

References

1. Chang, C., and Keisler, H.J. /1973/. Model Theory. North-Holland, Amsterdam.
2. Cresswell, M.J. /1973/. Logics and Languages. Methuen and Co., London.
3. Hintikka, L. /1962/. Knowledge and Belief: An Introduction to The Logic of Two values. Cornell Univ. Press.
4. Ha Hoang Hop /1986/. An Algebraic Approach to Knowledge Structuring. Proc. of KNVVT on Aotomation of Information Processes by Personal Computers. SzTAKI, Tanumányok, No. 194, pp.73-81.
5. Ha Hoang Hop /1987/ Formal Logic-based Approach to Knowledge Representation. Thesis. MTA-KFKI. Submitted to Hungarian Academy of Sciences.
6. Kalish, D., Montague , R., and Mar, G., /1980/ Logic - Technics for Formal Reasoning, Harcourt, New York.
7. Bach Hung Khang /1978/. Algebraic Closures in Pattern Recognition. Thesis of Doctor Degree. The Academy of Sciences of The USSR, 350 p.
8. Levesque, H. /1984/. Foundations of Functional Approach to Knowledge Representation, Artificial Intelligence 23, pp.155-212.
9. McDermott, D,and Doyle, J. /19 80/. Nonmonotonic Logic I, Artificial Intelligence 13/1, 2/, pp.34-41.

10. Mendelson, E. /1965/. Introduction to Mathematical Logic. Van Nostrand Reinhold, New York.

11. Moore, R.C.,/1983/. Semantic Considerations on Non-Monotonic Logic, Proc. of 8th IJCAI-83, Karsruhe, West Germany, pp.272-279.

12. Rasiowa, H, and Sikorski, R. /1970/. Mathematics of Metamathematics, PWN, Warsaw.

13. Reiter, R. On Closed-World Databases, in: Gallaire, H., and Minker, J. /Eds/, Logics and Databases, Plenum, New York, pp.55-76.

14. Rescher, N. /1975/. A Theory of Possibilities - A Constructivistic and Conceptualistic Account of Possible Individuals and Possible Worlds. Pittsburg Univ. Press.

15. Gelfond, M., and Przymusinska, H. /1986/. Negation as Failure: Careful Closure Procedure. Artificial Intelligence, 24, pp.13-63.

16. Bossu, G., and Siegel, P. /1985/. Saturation, Nonmonotonic and the Closed-World Assumption. Artificial Intelligence, 24, pp.13-63.

A DEDUCTIVE LANGUAGE FOR THE REPRESENTATION OF INCOMPLETE KNOWLEDGE

Ha Hoang Hop

Summary

In the paper a formal treatment of the problem of representing incomplete knowledge is given. By this a new approach to formal knowledge bases is described in terms of abstract interactive operations between them and expert systems (ES). An original knowledge representation language is introduced, that allows the explicite representation of non-monotonicity in the framework of beliefs and knowledges. In the paper the correctness of this language is also proved.

EGY DEDUKTIV NYELV A NEM TELJES TUDÁS REPREZENTÁLÁSÁRA

Ha Hoang Hop

Összefoglaló

A cikk a nem teljes tudás reprezentálásának egy uj formális tárgyalását adja. Ez a tárgyalás a formális tudásbázisok és a szakértő rendszerek közötti absztrakt interaktiv leképezésén alapszik. A szerző bevezet egy eredeti tudás-reprezentáló nyelvet, amely a "nem-monotonitásnak" egy explicit reprezentálását is lehetővé teszi, éspedig a tudás illetve vélemény fogalmak keretében. A szerző a nyelv korrektségét is bizonyitja.