# Schema Integration Framed in Modal Logic

Veselka Boeva

***Abstract:*** *In this paper standard modal logic is suggested as a formal framework for modelling and analysing different aspects of schema integration. Schemata are represented by sets of modal first-order formulae and interpreted in terms of standard models of modal logic. An approach based on correspondence assertions, i.e. expressing relationships between different constructs in the schemata, is used. It is demonstrated that freeness of conflicts is a condition ensuring that the schemata can be meaningfully integrated. In other words, then the schemata can be merged into an integrated schema that represents as much information as the original ones. In view of this, we first show how the model determining the schema that is a combination of the original ones can be restricted to a model representing the schema integrating the individual ones together with the correspondence assertions. Then the latter schema is proven to be dominating with respect to original ones.*
***Key words:*** *Conflict Detection, Modal Logic, Schema Integration*

## INTRODUCTION

Conflict detection is the key issue for software system design and use. For instance, conflict is a major complication in any schema integration process. Schemata are developed by different user groups or designers therefore, some constructs in the integrated schemata, while the modelled reality may be equivalent, may be incompatible and they must be modified before integration may take place. The latter is a subject of conflict resolution methods, whereas our interests in this paper are concerning conflict detection problems and especially, schema integration analyses in relation to freeness of conflict.

Many different approaches to schema integration have been presented. For instance, [9, 10, 13, 14] discussing various aspects of schema integration process in the framework of first-order logic. An approach using integration assertions, i.e. relating equivalent constructs in the schemata, is chosen in these works to analyse different features of this process with respect to freeness of conflicts. Later the results are applied to integration of multi-agent architecture designs [4, 7, 11] for handling problematic cases of global inconsistency in distributed information systems. However, due to complexity considerations in particular domains, these approaches introduced an unnecessarily complicated framework that can be reduced by the introduction of modal operators.

Similarly, different varieties of temporal logic and BDI logic are considered as a formal framework in [5, 8, 15], but also here, unnecessarily complicated machineries are introduced. Thus [1-3] suggest an approach using standard modal logic for expressing properties of specifications or schemata. To enable the methods for determining conflicts in specifications, and to allow analysis of integrated models, relevant sets of specifications or schemata are represented by sets of modal first-order formulae.This approach has several advantages. First, a quite weak language is sufficient for the purpose of modelling and analysing important aspects of schema dynamic. Then, due to considering together static and dynamic features of specifications, namely as a theory over the predicate modal logic KT, various procedures for detecting dynamic conflicts can be proposed, e.g. see [1].

In view of the above, our consideration herein, are based on the modal logic framework for conceptual schemata. We discuss schema integration process and suggest a procedure constructing a standard model of modal logic that determines the schema integrating the individual ones together with the correspondence assertions. Then it is proven that this schema weakly dominates the original ones when they are free of conflicts with respect to the correspondence assertions.

# CONFLICTFREENESS AND SCHEMA INTEGRATION

## 1. PRELIMINARIES

We first provide a brief background to modal logic. Then in section 1.2 we shall introduce a modal logic framework for conceptual schemata, developed in [1-3].

### 1.1 MODAL LOGIC

Modal logic can be seen as a generalization of classical propositional logic. It has been developed for formalizing arguments involving the notions of possibility and necessity.[1] The language of propositional modal logic consists of a set of atomic formulae, logical connectives, e.g., $\wedge$, $\vee$, $\neg$, $\rightarrow$, $\leftrightarrow$ as well as modal operators of *possibility* $\Diamond$ and *necessity* $\Box$.[2] The formulae of the language are of the following form: (i) atomic formulae; (ii) if p and q are formulae, so are: $\neg p$, $p \wedge q$, $p \vee q$, $p \rightarrow q$, $p \leftrightarrow q$, $\Diamond p$, $\Box p$.

A *system of modal logic* is any set of formulae closed with respect to all propositionally correct modes of inference. Any system of modal logic contains the axiom Df$\Diamond$ ($\Diamond p \leftrightarrow \Box p \Box$), and various different systems of modal logic can be obtained by imposing additional axioms.

The semantic analysis of a system of modal logic is performed using the notion of a *model of modal logic*, which is usually viewed as a structure of the form [6], M = <W, R, V>, where W denotes a *set of possible worlds* (or *states*), R is a binary relation on W called *accessibility relation* and V is a multivalued mapping from the set of atomic formulae into W called *value assignment function*.

The interpretation of the accessibility relation in a model of modal logic can vary significantly, but in general it may be thought as expressing the fact that some things may be possible from the standpoint of one world and impossible from the standpoint of another. Imposing various conditions on the accessibility relation, we obtain different classes of models of modal logic that determine different systems of modal logic. For instance, a system corresponding to a class of reflexive models is known as the *normal system of modal logic*, KT [6].

The value assignment function V associates to each atomic formula p the set V(p) of those possible worlds in which p is true. We use $||p||^M$ to denote the set of all worlds in which p is true. This set is known as the *truth set* of a formula p. Further a formula p is *true* in a model M if p is true in any world of M. The truth set is inductively extended to all non-modal formulae in the standard way. The truth conditions of modal formulae are defined using the accessibility relation R, i.e. for any formula p and any world $w \in W$ it holds:

$$w \in ||\Diamond p||^M \Leftrightarrow (\exists v \in W)(v \in R(w) \wedge v \in ||p||^M)$$
$$w \in ||\Box p||^M \Leftrightarrow (\forall v \in W)(v \in R(w) \Rightarrow v \in ||p||^M).$$

### 1.2 SCHEMA REPRESENTATION

In the sequel, we give a short overview of the approach considered in [1-3].

**Definition 1** Let L be a finite first-order language extended with the modal operators of possibility and necessity.[3]

[1] As was demonstrated in [12] modal logic (propositional and predicate) can be seen as a sub-logic of predicate logic (possibly with many sorts and generalized quantifiers).
[2] Predicate logic adds to propositional syntax quantifiers and names for variables, predicates, functions and constants.
[3] In most cases, a reasonable assumption is that there is a finite number of relevant objects in the considered reality, i.e. a language with a finite number of constants is used.

(i) A *schema* S is a first-order theory over the predicate modal logic KT, consisting of finitely many closed formulae in language L;

(ii) L(S) = {p | p$\in$L and p is a sub-formula of a formula of S}, i.e. L(S) is *closed under sub-formulae* of the formulae in S;

(iii) $P(z)\rightarrow\Diamond Q(z)$ describes possible *transitions*[4] between different states of S, where P(z) and Q(z) are first-order formulae in L, and z is a vector of variables in the alphabet of L[5];

(iv) An *integration assertion* expressing the schema $S_2$ in the schema $S_1$ is a closed first-order formula: $\forall(p(x)\leftrightarrow F(x))$, where p is a predicate symbol in $L(S_2)$ and F(x)[6] is a formula in $L(S_1)$.

Based on the fact that a normal system of modal logic is determined by each of its *canonical standard models* [6], the description of a schema in [1-3], was interpreted in terms of standard model of modal logic. A canonical model for a system of modal logic is a model which verifies just those formulae that are the theorems of the system.

**Definition 2** The *description* of a schema S is a standard model of modal logic M = <W, R, V>, such that

(i) W = {w | w is S-maximal set of formulae}[7];

(ii) For every $(w, v)\in W^2$, $v\in R(w)$ iff $\{\Diamond F(x) | F(x)\in v\}\subseteq w$;

(iii) For every atomic formula $F(x)\in L(S)$, V(F(x)) is the proof set[8] of the formula F(x).

The above model M can be interpreted as a L(S)-filtration of the canonical model for the system of modal logic KT, i.e. its worlds are the equivalence classes of worlds in the canonical model. A representative member of each equivalence class can be chosen. Thus, W can be considered as the set of all such worlds, which are S-maximal sets of formulae. Further in M just those atomic formulae are true at a world as are contained by it. Moreover, R is defined so that a world collects all the possibilitations of formulae occurring in its alternatives [6].

**Example 1** Let us consider a schema $S_1 = \{\neg r(a)\vee r(b), r(c)\leftrightarrow r(a), r(a)\rightarrow\Diamond r(b), \neg r(a)\rightarrow \Diamond(\neg r(b)\wedge\neg r(c))\}$. The description of $S_1$ is a model $M_1 = \langle W_1, R_1, V_1\rangle$, where $W_1 = \{w_1, w_2, w_3\}$ and

$w_1 = \{r(a), r(b), r(c), \Diamond r(b), \ldots\}$
$w_2 = \{\neg r(a), \neg r(b), \neg r(c), \Diamond(\neg r(b) \wedge \neg r(c)), \ldots\}$
$w_3 = \{\neg r(a), r(b), \neg r(c), \Diamond(\neg r(b) \wedge \neg r(c)), \ldots\}$. [9]

Then, due to Definition 2, the accessibility relation $R_1$ is given by $\{(w_1, w_1), (w_1, w_3), (w_2, w_2), (w_3, w_3), (w_3, w_2)\}$.

## 2. CONFLICTFREENESS

Next we recall the concept of conflictfreeness defined in [2]. Consider schemata $S_1$ and $S_2$ determined by the standard models of modal logic $M_1 = \langle W_1, R_1, V_1\rangle$ and $M_2 = \langle W_2, R_2,$

---

[4] In the considered context the dynamic part of a schema is represented by modal first-order logic formulae.
[5] The notation A(x) means that x is free in A(x).
[6] F(x) does not contain modal operators of possibility and necessity.
[7] A set of formulae w is a S-maximal set of formulae when it is S-consistent and has only S-inconsistent proper extensions [6].
[8] A proof set of a formula F(x), is the set of S-maximal sets of formulae containing F(x).
[9] In the description of the worlds in Example 1 are included only atomic formulae and formulae that are necessary to build the accessibility relation $R_1$.

$V_2$>, respectively. Let IA be a set of integration assertions expressing $S_2$ in $S_1$. Moreover, assume that $L(S_1) \cap L(S_2) = \emptyset$. Further, we regard a model M = < $W_1 \times W_2$, R, V>, where R is defined by:

$$(\forall (w_1, w_2) \in W_1 \times W_2)(R((w_1, w_2)) = R_1(w_1) \times W_2 \cup W_1 \times R_2(w_2)), \quad (1)$$

and for any atomic formula $F(x) \in L(S_1) \cup L(S_2)$, V is given by:

$$V(F(x)) = \begin{cases} V_1(F(x)) \times W_2 \text{ if } F(x) \in L(S_1) \\ W_1 \times V_2(F(x)) \text{ if } F(x) \in L(S_2) \end{cases}.$$

Note that the definition of R is different from that given in [2]. Namely, R is constructed in such a way that it models any transition possible in models $M_1$ and $M_2$. For instance, if $v_1 \in R_1(w_1)$ then for any $(w_2, v_2) \in W_2^2$ we have that $(v_1, v_2) \in R((w_1, w_2))$, since according to (1) each world in $R_1(w_1)$ is paired with each world in $W_2$. The same relationship holds for $R_2$ and R. Due to the above construction of M, it is clear that it is a model for the schema $S_1 \cup S_2$, i.e. the formula $\bigwedge_{F(x) \in S_1 \cup S_2} F(x)$ is valid in any world $(w_1, w_2) \in W_1 \times W_2$.

Now, let us consider the set of all worlds $(w_1, w_2) \in W_1 \times W_2$, that are models for the set of integration assertions IA, i.e. the formula

$$F_{IA} = \bigwedge_{F(x) \in IA} F(x)$$

holds in such worlds (states). Then in [2], the concept of freeness of conflicts is given in terms of these worlds, called secure states.

**Definition 3** The schemata $S_2$ and $S_1$ are *free of conflicts* w.r.t. IA iff

$$(\forall w_1 \in W_1)((\exists w_2 \in W_2)((w_1, w_2) \in ||F_{IA}||^M)).$$

According to the above definition, two schemata are free of conflicts with respect to a set of integration assertions iff for each world in the first model, there exists a world in the second one such that their union is a model for the set of integration assertions. Basically, this means that a schema is not allowed to restrict another schema when they are integrated, i.e., that all states that was possible to access before the integration, still are accessible after the integration.

**Example 2** Consider the schema $S_1$ from Example 1 and a schema $S_2 = \{p(a) \rightarrow p(b), p(b) \vee \neg p(c), (p(a) \wedge p(b)) \rightarrow \Diamond \neg p(b)\}$. Further, let $\forall x(p(x) \leftrightarrow r(x))$ be a possible integration assertion for the schemata. The description of schema $S_2$ is a model $M_2$ = <$W_2$, $R_2$, $V_2$>, where $W_2 = \{v_1, v_2, v_3, v_4, v_5\}$ and

$v_1 = \{p(a), p(b), p(c), \Diamond \neg p(b), \ldots\}$ $v_4 = \{p(a), p(b), \neg p(c), \Diamond \neg p(b), \ldots\}$
$v_2 = \{\neg p(a), \neg p(b), \neg p(c), \ldots\}$ $v_5 = \{\neg p(a), p(b), \neg p(c), \ldots\}$.
$v_3 = \{\neg p(a), p(b), p(c), \ldots\}$

$R_2$ is given by $\{(v_1, v_1), (v_1, v_2), (v_2, v_2), (v_3, v_3), (v_4, v_4), (v_4, v_2), (v_5, v_5)\}$. It can be easily checked that $||F_{IA}||^M = \{(w_1, v_1), (w_2, v_2), (w_3, v_5)\}$ and obviously the conditions for conflictfreeness, given in Definition 3, are fulfilled.

## 3. SCHEMA INTEGRATION

As was mentioned in the introduction, conflict detection is particularly useful when merging schemata. A usual requirement when integrating a set of schemata is that the combined schema does not restrict the intension behind the individual ones. This has been formalized in [3] by defining the concept of weak dominance in terms of standard modal logic, recalled in the next definition. This concept intuitively expresses that the dominating schema, in some sense, contains at least as much information as the dominated ones.

**Definition 4** $S_2$ *weakly dominates* $S_1$ w.r.t. the set of integration assertions IA, expressing $S_2$ in $S_1$, iff there is a total injective function ζ such that for each $(v_1, v_2) \in W_1^2$ and $v_2 \in R_1(v_1)$, $(\zeta(v_1), \zeta(v_2)) \in W_2^2$ and $\zeta(v_2) \in R_2(\zeta(v_1))$, and moreover $v_1 \cup \zeta(v_1)$ and $v_2 \cup \zeta(v_2)$ are IA-consistent sets of formulae.

First, we shall show that when $S_2$ and $S_1$ are free of conflicts w.r.t. IA then a transition to a secure state in the model M, determining the schema $S_1 \cup S_2$, is always possible, i.e. a transition to a world in the set $||F_{IA}||^M$. Thus, let us assume that schemata $S_2$ and $S_1$ are free of conflicts w.r.t. the set of integration assertions IA. Obviously, it must be proven that

$$(\forall (w_1, w_2) \in W_1 \times W_2)\ (R((w_1, w_2)) \cap ||F_{IA}||^M \neq \varnothing). \qquad (2)$$

It can be easily checked that M is a reflexive model. Namely, taking into account (1) and the fact that $R_1$ and $R_2$ are reflexive relations we obtain that $(\forall (w_1, w_2) \in W_1 \times W_2)((w_1, w_2) \in R((w_1, w_2)))$. Due to this and the freeness of conflicts between $S_2$ and $S_1$ w.r.t. IA, it holds that for any $w_1$ in $W_1$ there exists a world $w_2$ in $W_2$, such that $(w_1, w_2) \in ||F_{IA}||^M$ and moreover $(w_1, w_2) \in R((w_1, w_2))$, i.e. $R((w_1, w_2))$ has a non-empty intersection with $||F_{IA}||^M$. Clearly, the above assertion is valid and furthermore, in view of the structure of the truth sets of modal formulae, it is equivalent to $W_1 \times W_2 = ||\Diamond F_{IA}||^M$.

Now, let us restrict the model $M = \langle W_1 \times W_2, R, V \rangle$ to a model in which the only possible transitions are those to secure states. Hence, $R((w_1, w_2)) \subseteq ||F_{IA}||^M$ must be true for each world $(w_1, w_2)$ in such a model. Thus, we consider only the worlds in the set $||F_{IA}||^M$ and restrict M to $M_{IA} = \langle W_{IA}, R_{IA}, V_{IA} \rangle$, in such a way that $W_{IA} = ||F_{IA}||^M$ and $R_{IA}$ is defined by:

$$(\forall (w_1, w_2) \in W_{IA})(R_{IA}((w_1, w_2)) = R((w_1, w_2)) \cap ||F_{IA}||^M).$$

Moreover, $V_{IA}$ is the restriction of V to $W_{IA}$. The definition of $R_{IA}$ and (2) imply the reflexivity of model $M_{IA}$. Observe also, that any transition possible in $M_1$ is represented in $M_{IA}$. Then it can be easily seen that $M_{IA}$ is a model for the schema $S_1 \cup S_2 \cup IA$, i.e. the formula $F_{S1 \cup S2 \cup IA}$ holds in any world of this model.

Next, we shall prove that the integrated schema $S_1 \cup S_2 \cup IA$ weakly dominates $S_1$ w.r.t. IA. In view of Definition 4, it is required that there should be a total injective function from the dominated schema to the dominating one. Let us consider $(w_1, v_1) \in W_1^2$ and $v_1 \in R_1(w_1)$ in the model $M_1 = \langle W_1, R_1, V_1 \rangle$, determining the schema $S_1$. Since $S_2$ and $S_1$ are conflictfree w.r.t. IA then for $(w_1, v_1) \in W_1^2$ there exists $(w_2, v_2) \in W_2^2$, such that $((w_1, w_2), (v_1, v_2)) \in W_{IA}^2$. Moreover, we have that $(v_1, v_2) \in R_{IA}((w_1, w_2))$. The latter follows from the definitions of R and $R_{IA}$. Consequently, we can define the function ζ by letting $\zeta(w_1)$ be equal to the corresponding $(w_1, w_2)$, for any $w_1 \in W_1$. Clearly, ζ defined in this way is injective. It is also obvious that the function ζ has the property that for each $w_1 \in W_1$, $w_1 \cup \zeta(w_1)$ is an IA-consistent set of formulae. Thus, we obtain that the schema $S_1 \cup S_2 \cup IA$ weakly dominates $S_1$ w.r.t. the given set of integration assertions IA.

**Example 3** Consider the schemata $S_1$ and $S_2$ from Example 2. They are free of conflicts w.r.t. IA, defined by $\forall x(p(x) \leftrightarrow r(x))$. Thus we can build a model $M_{IA} = \langle W_{IA}, R_{IA}, V_{IA} \rangle$, where $W_{IA} = \{(w_1, v_1), (w_2, v_2), (w_3, v_5)\}$ and $R_{IA}$ is given by $\{((w_1, v_1), (w_1, v_1)), ((w_1, v_1), (w_2, v_2)), ((w_1, v_1), (w_3, v_5)), ((w_2, v_2), (w_2, v_2)), ((w_3, v_5), (w_3, v_5)), ((w_3, v_5), (w_2, v_2))\}$.

It can be easily seen that $M_{IA}$ determines the integrated schema $S_1 \cup S_2 \cup IA$ and moreover, the latter one weakly dominates $S_1$ w.r.t IA, according to Definition 4.

## CONCLUSIONS AND FUTURE WORK

In this work, standard modal logic is used for expressing and analysing integrated schemata. First, we have recalled the concepts of freeness of conflict and weak dominance introduced in terms of modal logic, see [2] and [3], respectively. Then we have shown how the model determining the schema that is a combination of the individual ones can be restricted to a model representing the integration of these schemata together with the correspondence assertions. Moreover, it has been proven that the schema that is a result of this integration dominates the individual ones in case of conflictfreeness.

## REFERENCES

[1] Boeva, V. Kripke's semantics as semantics for dynamic conflict detection. In: Proc. of International Conference AUTOMATICS & INFORMATICS'2002, 61–64, 2002.

[2] Boeva, V., L. Ekenberg. A Transition Logic for Schemata Conflicts. Journal of Data and Knowledge Engineering, 2004, accepted.

[3] Boeva, V. A Modal Logic Based Approach to Schema Integration. In: Proc. of CompSysTech'2003, Int. Conf. on Computer Systems and Technologies, II:1-6, 2003.

[4] Boman, M., L. Ekenberg. Risk Constraints in Agent Conflicts. In: Proc. LUMISr2000, IEEE Computer Society Press, 2000.

[5] Carmo, J., A. Sernadas. A Temporal Logic Framework for a Layered Approach to Systems Specification and Verification. In: C. Rolland, F. Bodart, M. Leonard (ed): Temporal Aspects of Information Systems, North-Holland 31-46, 1988.

[6] Chellas, B. Modal Logic, an Introduction. Cambridge University Press, Cambridge, 1980.

[7] Davies, G., L. Ekenberg, P.Johannesson. Detecting Temporal Conflicts in Integrated Agent Specifications. In: Mueller and Dieng (ed.): Computational Conflicts. Springer Verlag 103-124, 2000.

[8] Dixon, C., M. Fisher, M. Wooldridge. Resolution for temporal logics of knowledge. Journal of Logic and Computation, 8:345-372, 1998.

[9] Ekenberg, L., P. Johannesson. Conflictfreeness as a Basis for Schema Integration. In: Proc. CISMOD'95, Lecture Notes in Computer Science Springer-Verlag, 1-13, 1995.

[10] Ekenberg, L., P. Johannesson. A Formal Basis for Dynamic Schema Integration. In: Proc. of 15th International Conference on Conceptual Modelling ER'96, Lecture Notes in Computer Science Springer-Verlag, 211-226, 1996.

[11] Ekenberg, L. The Logic of Conflicts between Decision Making Agents. Journal of Logic and Computation, 10 4:583-602, 2000.

[12] Hajek, P. Metamathematics of Fuzzy Logic, Klumer Academic Publishers, Netherlands, 1998.

[13] Johannesson, P. A Logic Based Approach to Schema Integration. In: Teorey, T. (ed.): 10th Int. Conf. on Entity-Relationship Approach. San Francisco, North Holland, 1991.

[14] Johannesson, P. A Logical Basis for Schema Integration. In: Proc. of Third International Workshop on Research Issues in Data Engineering-Interoperability in Multidatabase Systems, Ed. H. Schek, Vienna, IEEE Press, 1993.

[15] Rao., A. S., M.P. Georgeff. Decision Procedures for BDI Logics. Journal of Logic and Computation, 8:293 - 342, 1998.

## ABOUT THE AUTHOR

Ass.Prof. Veselka Boeva, PhD, Department of Computer Systems, Technical University of Plovdiv, Phone: +359 32 659 726, E-mail: veselka_boeva@hotmail.com.