This work uses temporal logic as a calculus for expressing integrity constraints that specify admissibility of dynamic database behaviour. Formulas are interpreted in state sequences representing dynamic behaviour. Our approach incorporates temporal quantification by "always", "sometimes", and quantifiers bounded by intervals in state sequences. Moreover, dynamically changing domains of database values are considered. We then use special kinds of formulas as a language for dynamic constraints and give some hints how to specify in typical situations. For such formulas, a frame for monitoring constraints during runtime of a database is discussed which allows to characterize admissibility operationally.

1. INTRODUCTION

Many approaches to database design are suggesting a "conceptual" (or "logical") design step where a global information structure, the conceptual schema, should be defined independently of any implementation of a data model [5,21]. This step corresponds to the specification phase of general software design. Like that it has to bridge the gap between requirement analysis and implementation in a specific database management system. Therefore, a specification language is needed to describe the contents and the behaviour of a database in an abstract and exact way.

Specifying a conceptual schema essentially means defining the properties of database objects in the course of time. A database develops dynamically according to user transactions manipulating its contents, i.e. taking the database from one state into another. Thus the dynamic behaviour of a database can be represented by sequences of states. If a sequence obeys certain integrity constraints on database objects, it becomes "admissible". Such rules may refer to each single state, but generally they refer to several states at once or to entire intervals in the state sequence. Together with the syntactic description of database objects, the constraints make up the schema specification.

Since Hammer/McLeod [11] and Stonebraker [20] (both 1975) there have been numerous papers dealing with integrity constraints for databases. Most of them, however, are only concerned with "static" constraints, i.e. criteria for admissible states, disregarding dynamic contexts. The problem of "dynamic" constraints specifying admissible state sequences has rarely been addressed. Some preliminary ideas are contained in the ISO-report [12]; [15,16] consider pairs of before/after states.

Except for designing and implementing a database, a schema specification serves as a basis for monitoring dynamic behaviour, i.e. for checking admissibility of a state sequence during runtime. The problem of enforcing (static) integrity constraints [20,22] has recently found growing interest [6,14,23].
A new approach to dynamic constraints using temporal logic \cite{13,17} has been initiated by Sernadas \cite{18}, Casanova et al.\cite{3,4} and Golshani et al.\cite{10}. We take up these ideas here, because temporal logic is, in our opinion, a convenient tool for modelling time-dependent database aspects (cf.\cite{1}). By concentrating on the problem of specifying dynamic constraints without considering the problem of specifying and verifying transactions \cite{4,24}, we restrict ourselves to a simpler form of temporal logic incorporating only the temporal formulas "always" and "sometime" as well as their bounded versions "always... until" and "sometime... before". We, however, generalize the work mentioned above since we are able to handle dynamically changing domains; we distinguish between "possible" and "actual" values by providing corresponding kinds of \( \forall, \exists \) quantifiers.

Such a logic calculus for formulating constraints must be based on specific concepts for modelling objects. Like Golshani et al.\cite{10} we adopt a functional approach in the spirit of Buneman/Frankel \cite{2} and Shipman \cite{19}. Accordingly, a schema consists of sorts of entities, functions taking arguments and delivering results of specified sorts, and constraints expressed in temporal logic. It should be noted, however, that temporal logic is not bound to this approach to object modelling; it can be used with other approaches as well, e.g. relational ones as in \cite{4}.

In contrast to previous functional approaches, we introduce two syntactic levels, the data level and the object level. The data level comprises specifications of basic data types like \texttt{BOOL}, \texttt{INT}, etc. The data level has a fixed interpretation that does not vary in time and is often the same for large classes of database schemata. The object level, on the other side, contains sorts like \texttt{Person}, \texttt{Project}, etc. and functions on them whose interpretation varies in time, dependent on the database state. The object level will most probably be different for different database schemata.

In the next section, we present our version of temporal logic, based on our approach to object modelling. We define the syntax and semantics of temporal formulas including temporal quantification. The semantics is explained by interpreting formulas in infinite state sequences.

In section 3, conceptual schema specifications are restricted to special kinds of formulas which probably are sufficient to express dynamic database constraints in most applications. We use a suggestive syntax for constraints referring to intervals in a state sequence and we give some hints how to build constraints depending on the existence of objects. As an example, we specify parts of a database for an automobile market (cf. \cite{12}).

Section 4 discusses principles for monitoring dynamic constraints during runtime of a database. For the constraint language chosen, we obtain an operational characterization of admissibility. This makes clear on which conditions a current state sequence will become admissible in the future, if it has already been accepted up to a present state.

Some preliminary ideas and results along these lines have been given in our previous paper \cite{8}.

2. TEMPORAL LOGIC

Temporal logic can be based on different approaches to object modelling, e.g. relational \cite{4} or functional \cite{10} ones. Here, we adopt the functional approach as described in \cite{2,19}, since it conforms in spirit with the well developed methodology and theory of data type specifications, especially the algebraic approach \cite{7,9}, and also has strong relationship with Chen's ER-model.

A conceptual schema specification is given in two levels, the data level and the object level.
On the data level, a data specification \( D\text{-SPEC} = \langle S_D, \Omega_D, \xi_D \rangle \) consists of data sorts \( S_D \), data functions \( \Omega_D \) between data sorts, and axioms \( \xi_D \) (most often equations) defining the behaviour of the data functions. The "syntactic" part of \( D\text{-SPEC} \) is called the data signature \( \xi_D = \langle S_D, \Omega_D \rangle \). For the purpose of this paper, we assume that the data specification is fixed, specifying standard data types like BOOL, INT, CHAR, TEXT, DATE etc.

The object level extends the data specification by giving an object specification \( O\text{-SPEC} = \langle \langle S_0, \xi_0 \rangle, \phi_0, c_0 \rangle \), consisting of a partially ordered set of object sorts \( \langle S_0, \xi_0 \rangle \), object functions \( \phi_0 \) between object and data sorts, and constraints \( c_0 \) determining admissible states and state sequences. The partial ordering \( \preceq \) on \( S_0 \) represents a generalization relationship between object sorts. The "syntactic" part of \( O\text{-SPEC} \) is called the object signature \( \Sigma_0 = \langle \langle S_0, \xi_0 \rangle, \phi_0 \rangle \).

A conceptual schema specification \( CS\text{-SPEC} = D\text{-SPEC} + O\text{-SPEC} \) consists of a data specification \( D\text{-SPEC} \) and an object specification \( O\text{-SPEC} \) based upon \( D\text{-SPEC} \). The syntactic part \( \xi_{CS} = \xi_D + \xi_0 \) is called a conceptual schema signature. In what follows, we will concentrate on \( O\text{-SPEC} \) assuming \( D\text{-SPEC} \) to be fixed. In order to illustrate the syntactic part of an object specification, we give an example. We make use of a set constructor as a parameterized object type: for each sort \( s \in S_0 \), "set \( s \)" denotes a new sort (of sets of objects of sort \( s \)), on which the predicate \( \in : s \times \text{set } s \rightarrow \text{BOOL} \) for set membership is defined.

Example: We give part of the object signature of a simple automobile market database.

\[
\begin{align*}
S_0 & : \text{Manufacturer, Garage, Person, Car_Owner, Car, Model} \\
\preceq & : \text{Manufacturer } \preceq \text{ Car_Owner} \\
& \text{Garage } \preceq \text{ Car_Owner} \\
& \text{Person } \preceq \text{ Car_Owner} \\
\phi_0 & : \text{name : Car_Owner } \rightarrow \text{ TEXT} \\
& \text{property : Car_Owner } \rightarrow \text{ set } \text{ Car} \\
& \text{models : Manufacturer } \rightarrow \text{ set } \text{ Model} \\
& \text{serialno : Car } \rightarrow \text{ INT} \\
& \text{model : Car } \rightarrow \text{ Model} \\
& \text{manuf : Car } \rightarrow \text{ Manufacturer} \\
& \text{owner : Car } \rightarrow \text{ set } \text{ Car_Owner} \\
\end{align*}
\]

In order to complete this object signature up to an object specification, constraints \( c_0 \) have to be added. Following [3,4,10,18] we propose temporal logic as an appropriate language for specifying static and dynamic database constraints. To introduce our version of temporal logic we first discuss the semantics, i.e. the interpretation, of object signatures.

Let \( \Sigma = \Sigma_D + \Sigma_0 \) be a conceptual schema signature. A \( \Sigma \)-universe is a mapping \( \pi \)
associating a set $\pi(s)$ of possible values to each object sort $s \in S_0$. A $L$-
instance or $L$-state $\sigma$ wrt $\pi$ is a mapping associating a set $\sigma(s)$ of actual
values to each object sort $s \in S_0$ and associating an actual function $\sigma(f)$:
$\sigma(s_1) \times \ldots \times \sigma(s_n) \rightarrow \sigma(s_0)$ to each object function $f: s_1 \ldots s_n \rightarrow s_0$.
These mappings are subject to the following conditions:

(i) $\forall s \in S_0 : \sigma(s) \subseteq \pi(s)$
(ii) $\forall s, t \in S_0 : s \leq t \Rightarrow \pi(s) \subseteq \pi(t)$
(iii) $\forall s, t \in S_0 : s \leq t \Rightarrow \sigma(s) \subseteq \sigma(t)$

For data sorts $s \in S_D$, we assume a fixed interpretation where all possible val-
ues are actual, i.e. $\sigma(s) = \pi(s)$. Moreover, for the set constructor, we have
(iv) $\forall s \in S_0 : \pi(\text{set } s) = 2^\pi(s)$
(v) $\forall s \in S_0 : \sigma(\text{set } s) \subseteq 2^{\sigma}(s)$

Let $X = \{x_1, x_2, \ldots\}$ be a set of variables and let each variable $x \in X$ have a sort
$\text{sort}(x) \in S = S_D + S_0$. We assume that there is a countably infinite number of vari-
ables of each sort. We define the language constructs of temporal logic along
with their semantics. Terms and atomic formulas are interpreted in a given
state $\sigma$ and a given local substitution $\alpha$ for a set $Y \subseteq X$ of variables, i.e. a
mapping $\alpha: Y \rightarrow \bigcup_{s \in S} \sigma(s)$ such that, for each $x \in Y$, $\alpha(x) \in \sigma(\text{sort}(x))$. Formulas
are interpreted in a given state sequence $\sigma = (\sigma_0, \sigma_1, \ldots)$ and a given global
substitution $\alpha$ defined like a local substitution with $\sigma$ replaced by $\pi$ and
where all states are based on a fixed given universe $\pi$. We use the following
notations: $[\sigma, \alpha](t)$ denotes the value of a term $t$ in a state $\sigma$ with a local
substitution $\alpha$; $[\sigma, \alpha] \models \phi$ denotes the validity of a temporal formula $\phi$ in
a state sequence $\sigma$ with a global substitution $\alpha$.

(1) Terms

(1.1) Syntax: Each variable $x \in X$ is a term of sort $\text{sort}(x)$.
Semantics: $[\sigma, \alpha](x) = \alpha(x)$

(1.2) Syntax: If $f: s_1 \ldots s_n \rightarrow s_0 \in S_0$, and $t_i$ is a term of sort $s_i$
for $i = 1, \ldots, n$, then $f(t_1, \ldots, t_n)$ is a term of sort $s_0$.
Semantics: $[\sigma, \alpha](f(t_1, \ldots, t_n)) = \sigma(f)([\sigma, \alpha](t_1), \ldots, [\sigma, \alpha](t_n))$

Remark: As a special case, constants $c$ of sort $s$ can be viewed as nullary
functions $c: \rightarrow s$, and their semantics is an element $\sigma(c) \in \sigma(s)$.

(2) Atomic Formulas

(2.1) Syntax: Each boolean term $t$ (i.e. $\text{sort}(t) = \text{BOOL}$) is an atomic formula.
Semantics: $[\sigma, \alpha] \models t$ iff $[\sigma, \alpha](t) = \text{true}$

(2.2) Syntax: If $t, u$ are terms of the same sort, $t = u$ is an atomic formula.
Semantics: $[\sigma, \alpha] \models t = u$ iff $[\sigma, \alpha](t) = [\sigma, \alpha](u)$
(2.3) Syntax: If \( t \) is a term and \( s \in \mathcal{S}_0 \) such that \( s \in \text{sort}(t) \), then \( t \) is \( s \) is an atomic formula.
Semantics: \( [\sigma, a] \models t \) \( s \) iff \( [\sigma, a](t) \in \sigma(s) \)

(3) Formulas

(3.1) Atomic formulas
Syntax: Each atomic formula \( \phi \) is a formula. Each occurrence of a variable in \( \phi \) is free.
Semantics: \( [\sigma, a] \models \phi \) iff \( \alpha \) is local in \( \sigma_0 \) and \( [\sigma_0, \alpha] \models \phi \).
Here, \( \alpha \) is local in \( \sigma_0 \) iff \( \alpha(x) \in \sigma_0(\text{sort}(x)) \) for each \( x \) occurring in \( \phi \), i.e. all objects substituted by \( \alpha \) for occurring variables exist in the state \( \sigma_0 \).

(3.2) Boolean connectives
Syntax: If \( \phi, \psi \) are formulas, then so are \( \neg \phi \), \( \phi \land \psi \), etc. All free occurrences of variables in \( \phi \) or \( \psi \) remain free in these formulas.
Semantics: \( [\sigma, a] \models \neg \phi \) iff not \( [\sigma, a] \models \phi \),
\( [\sigma, a] \models \phi \land \psi \) iff \( [\sigma, a] \models \phi \) and \( [\sigma, a] \models \psi \), etc.

(3.3) Quantification over actual values
Syntax: If \( \phi \) is a formula and \( x \in X \) is a variable, then \( (\forall x) \phi \) and \( (\exists x) \phi \) are formulas. Each free occurrence of \( x \) in \( \phi \) is no longer free in these formulas, whereas free occurrences of other variables remain free.
Semantics: \( [\sigma, a] \models (\forall x) \phi \) iff for all \( v \in \sigma_0(\text{sort}(x)) \), we have \( [\sigma, a(x(v))] \models \phi \). Here, \( a(x(v)) \) denotes \( a \) changed for the argument \( x \) to yield the value \( v \). Similarly, \( [\sigma, a] \models (\exists x) \phi \) iff there is a \( v \in \sigma_0(\text{sort}(x)) \) such that \( [\sigma, a(x(v))] \models \phi \).

(3.4) Quantification over possible values
Syntax: Like (3.3), with \( \forall \) and \( \exists \) instead of \( \forall \) and \( \exists \).
Semantics: Like (3.3), with \( \pi \) instead of \( \sigma_0 \).

(3.5) Temporal quantification
Syntax: If \( \phi \) is a formula, then always \( \phi \) and sometime \( \phi \) are formulas. Free occurrences of variables in \( \phi \) remain free in these formulas.
Semantics: \( [\sigma, a] \models \text{always} \ \phi \) iff for all \( i \geq 0 \), \( [\sigma_{i+1}, a] \models \phi \).
Here, \( \sigma_i = (\sigma_{i-1}, \sigma_{i-2}, \ldots) \) denotes the tail of \( \sigma \) from the position \( i \).
Similarly, \( [\sigma, a] \models \text{sometime} \ \phi \) iff there is an \( i \geq 0 \) such that \( [\sigma_i, a] \models \phi \).

Remark: We will also talk about interpretation in finite state sequences \( (\sigma_0 \ldots \sigma_n) \). By this we mean the interpretation in the infinite state sequence \( (\sigma_0 \ldots \sigma_n \sigma_n \sigma_n \ldots) \) remaining constant from position \( n \) on. Accordingly, by an interpretation of a temporal formula in a single state \( \sigma \) we mean the interpretation in the state sequence \( (\sigma \sigma \sigma \ldots) \).
The well known duality of universal and existential quantification under negation, \( \neg(\forall x)\phi \iff (\exists x)\neg\phi \), holds for both forms of quantification introduced here, over actual and possible values. Moreover, a corresponding duality principle holds for temporal quantification, too:

\[ \neg \text{always } \phi \iff \text{sometime } \neg \phi \]
or, equivalently,

\[ \neg \text{sometime } \phi \iff \text{always } \neg \phi \]

The language of temporal logic gives an appropriate means to specifying database constraints; for an example see section 3. In many applications, however, situations have to be expressed where a temporal quantifier is not required to hold indefinitely, but is bounded in time by some other condition. Typically such conditions can be represented by the first state in a sequence where a given formula \( \tau \) becomes true. In order to express such situations conveniently, two additional bounded temporal quantifiers are introduced:

- \( \text{always } \phi \text{ until } \tau \)
- \( \text{some time } \phi \text{ before } \tau \)

Here, \( \phi \) and \( \tau \) are arbitrary temporal formulas. The semantics of bounded temporal quantification in a state sequence \( \sigma \) and a global substitution \( a \) is as follows. Let \( m = \min(\{ j \geq 0 \mid [\sigma^j, a] \models \phi \} + \{ \infty \}) \).

\[ [\sigma, a] = \text{always } \phi \text{ until } \tau \text{ iff for all } i, 0 < i < m, \text{ we have } [\sigma^i, a] \models \phi \]

\[ [\sigma, a] = \text{some time } \phi \text{ before } \tau \text{ iff there exists an } i, 0 < i < m, \text{ such that } [\sigma^i, a] \models \phi \]

These definitions do not imply that \( \tau \) ever becomes true. If wanted, this must be specified additionally by \( \text{some time } \tau \). The bounded quantifiers again behave dually:

\[ \neg \text{always } \phi \text{ until } \tau \iff \text{some time } \neg \phi \text{ before } \tau \]

\[ \neg \text{some time } \phi \text{ before } \tau \iff \text{always } \neg \phi \text{ until } \tau \]

3. CONCEPTUAL SCHEMA SPECIFICATION

For most database applications, we feel it is sufficient to restrict the general form of temporal formulas given in the previous section to a simpler one. Additionally, we are interested to have constraint formulas that can be monitored during database operation; this aspect will be discussed in section 4.

Therefore, from now on we only consider formulas of the form

- \( \forall x_1 \ldots \forall x_n \text{ always } (\phi \implies \text{always } \psi \text{ until } \tau) \)
- \( \forall x_1 \ldots \forall x_n \text{ always } (\phi \implies \text{some time } \psi \text{ before } \tau) \)

where \( x_1, \ldots, x_n \) are the free variables of the bracketed parts and \( \phi, \psi \) and \( \tau \) are nontemporal(!) formulas. Furthermore we will drop the first \( \text{always} \) and all \( \forall \)-quantifications and use a more suggestive syntax:

\[ \text{from } \phi \text{ always } \psi \text{ until } \tau \]
\[ \text{after } \phi \text{ some time } \psi \text{ before } \tau \]

Formulas involving "always" are called universal constraints, whereas for formulas involving "some time" we use the notion existential constraint. A state sequence \( \sigma \) (finite or infinite) is called admissible wrt a set \( C \) of con-
straints, iff each $\gamma \in C$ (in its original form) is valid in $\sigma$ with an arbitrary substitution. In this case we will also say that the state sequence $\sigma$ satisfies the constraints $C$.

The from/after- and until/before-bounds may be left out liberately and then the formulas

$$
\begin{align*}
\text{always } \varphi & \text{ until } \tau \\
\text{from } \varphi & \text{ always } \psi \\
\text{always } \varphi & \\
\end{align*}
$$

are understood to be semantically equivalent to:

$$
\begin{align*}
\text{from } \text{true always } \psi & \text{ until } \tau \\
\text{from } \varphi & \text{ always } \psi & \text{ until } \text{false} \\
\text{from } \text{true always } \psi & \text{ until } \text{false} \\
\end{align*}
$$

Of course the same applies to existential constraints by analogy.

Considering the time-variant nature of our constraints, it should be noted that the from/after-bounds establish a "less-equal" relation in time, whereas the until/before-bounds express a "less" relation in time.

In our approach it is possible to generally specify the existence of objects (of sort $s$) as a boolean function

$$
\text{exists: } s \rightarrow \text{BOOL}
$$

by giving the (defining) constraint

$$
\forall x : \text{exists}(x)
$$

with $x$ variable of sort $s$. Please note that we have to use the quantification $\forall$ over actual values (and not $\forall$ over possible values). Thus $\text{exists}(x)$ will be true exactly for those objects which are existent in the present state of the database. For the rest of the paper we assume "exists" predicates for all object sorts.

Example: We give a few constraints for our automobile market database whose signature has been specified in section 2 and is now enlarged by the following boolean object functions:

$$
\begin{align*}
\text{registered} & \quad \text{deregistered} : \text{Car} \rightarrow \text{BOOL} \\
\end{align*}
$$

The variables used here have the following sorts:

$$
\begin{align*}
\text{c, c': Car} & \quad \text{co, co', co'': Car_Owner} \\
\text{m: Model} & \quad \text{mf: Manufacturer} \\
\text{ms: set Model} & \\
\end{align*}
$$

(A) $\text{after exists(c) sometime registered(c) before deregistered(c)}$

(B) $\text{from registered(c) always registered(c) until deregistered(c)}$

(C) $\text{from deregistered(c) always deregistered(c) until } \neg \text{exists(c)}$

(D) $\text{after } c \in \text{property(co)} \land co \text{ is Manufacturer} \\
\text{ sometime } \exists \text{co': } c \in \text{property(co')} \land co' \text{ is Garage} \\
\text{ before } \exists \text{co'': } c \in \text{property(co'')} \land co'' \text{ is Person}$

(E) $\text{from } c \in \text{property(co)} \land \neg (co \text{ is Manufacturer})$

$$
\begin{align*}
\text{always } \forall \text{co': } c \in \text{property(co')} \Rightarrow \neg (co' \text{ is Manufacturer}) \\
\end{align*}
$$

(F) $\text{sometime } \neg \text{exists(m)}$

(G) $\text{after ms=models(mf)}$

$$
\begin{align*}
\text{ sometime } \exists \text{exists(m)} \land m \in \text{models(mf)} \land \neg m \in \text{ms}
\end{align*}
$$
(H) $c \in \text{property}(co) \iff co \in \text{owner}(c)$

(I) $m=\text{model}(c) \Rightarrow m \in \text{models}(\text{manuf}(c))$

(J) $\text{serialno}(c) = \text{serialno}(c') \Rightarrow c=c'$

Constraints (A)-(C) require that, if a car exists in the database, it sometime must be registered and then will always be registered until it is deregistered. Once it is deregistered, it cannot be registered again. (D) and (E) express that cars are sold from manufacturers to garages and after this cannot be owned by manufacturers again. (F) and (G) demand that "old" models will disappear from the market and "new" models have to be produced. Constraints (H)-(J) refer to each single state of a given state sequence since they do not contain temporal quantifiers. Such constraints are called static. (H) for example expresses that the functions "property" and "owner" contain the information. The first constraints (A)-(G), in contrast, are called dynamic, since they only make sense when considering state sequences.

Concerning the expressive power of our constraint language, we can specify any nontemporal condition on database objects to hold in all or some states of a certain interval in a state sequence. Such an interval is bounded by two other nontemporal conditions on objects. If one wants to express conditions involving calendar dates, durations, deadlines etc. (i.e. real world time), appropriate objects have to be modelled by including additional functions into the signature. E.g., by using a function today: $\rightarrow DATE$ and a variable $d$ of sort $DATE$, the constraint

$$\text{after } d=\text{today} \land \exists(c) \text{ sometime registered}(c) \text{ before } today>d+90$$

says that a car should be registered within the first 90 days of its lifetime.

The exists predicates may be used to specify certain dependencies between the validity of constraints and the existence of objects involved therein. Let us consider some typical situations:

On the one hand, some formulas, e.g. a disjunction like "$\text{registered}(c) \lor \text{registered}(c')$", may become true, even if an object substituted for one variable does not exist. The existence of an object can be required explicitly by a constraint of the form

$$\text{from } \phi \text{ always } (\psi \land \exists(x)) \text{ until } \tau$$

where $x$ is a variable which (probably) appears in $\phi$ and $\psi$.

On the other hand, some formulas, e.g. a conjunction of atomic formulas like "$\text{registered}(c) \land \text{registered}(c')$", can only become true, if all objects involved do exist. But in general, objects may disappear from the database and perhaps sometime later re-appear. So there may be different life cycles or incarnations for one object. To restrict a constraint to the first incarnation of an object, the until (before) clause has to be modified:

$$\text{from } \phi \text{ always } \psi \text{ until } (\tau \lor \tau \exists(x))$$

Another possibility is to demand the validity of a constraint exactly for all periods when an object is present in the database:

$$\text{from } \phi \text{ always } (\exists(x) \Rightarrow \psi) \text{ until } \tau$$
4. OPERATIONAL CHARACTERIZATION OF ADMISSIBILITY

After we have introduced constructs for formulating a conceptual schema specification, we will now study principles of monitoring the specified constraints during runtime of a database. Although implementation aspects are not discussed, a framework for enforcing dynamic constraints will emerge.

Since dynamic constraints refer to sequences of database states or at least to intervals in sequences, the problem arises how to check relationships between past, present and future states in a current sequence. At each moment the present state has been reached from an initial state by performing some user transactions which induce corresponding state transitions. Whereas those past states are known, the future states cannot be predicted, if users may choose their transactions arbitrarily. Thus, any general monitoring can only deal with the database history up to the present state and check that sequence prefix for constraint violations.

But to avoid storing the entire history, the knowledge about it should be limited to a part necessary for checking any continuation. If, e.g., the start condition $\phi$ of a universal constraint "from $\phi$ always $\psi$ until $\tau$" became true in some past state for some (global) substitution $\alpha$, then $\psi$ has to be valid in that and all subsequent states until the termination condition $\tau$ occurs; so the information $\langle\phi,\tau,\alpha\rangle$ meaning "(always $\psi$ until $\tau$) with substitution $\alpha$" should be remembered for exactly that period. The key idea of monitoring is to maintain two dynamically changing sets of such notes, one set induced by universal constraints and the other one by existential constraints. These sets will be denoted as $C_u$ and $C_e$, respectively.

In the following, we present and analyse a program called DB-MONITOR that

(i) executes user transactions (which are left unspecified here)
(ii) and checks the induced state sequence for constraint violations in the actual prefix.

Then, admissible state sequences can be characterized operationally, i.e. in terms of that monitor, what can be shown by verifying some program invariants.

Let a conceptual schema specification $CS\text{-SPEC}$ be given with a set $C$ of constraints. The frame of the program DB-MONITOR is given below. As global variables it uses $\sigma$ for the actual database state, $\alpha$ for the actual user transaction, and two sets $C_u, C_e$ consisting of triples $\langle\phi,\tau,\alpha\rangle$ s.t. $\psi, \tau$ are nontemporal formulas and $\alpha$ is a substitution. The procedure CHECK will be defined afterwards.

```
program DB-MONITOR :
    procedure EXECUTE;
    ( executes $\alpha$ and changes $\sigma$ accordingly );
    procedure CHECK;
    ...
    begin ( initialize $\sigma$ );
    CHECK;
    while true do
    begin ( input $\alpha$ );
    EXECUTE; CHECK
    end
end.
```
The following procedure CHECK manipulates the set variables $C_u$ and $C_e$: let INSERT and DELETE denote the corresponding set operations. Moreover, a boolean function $\text{VALID}(\phi, a)$ computing $([\sigma, a] = \phi)$ and a procedure ERROR that handles constraint violations, e.g. by undoing the last transaction, are used.

procedure CHECK:
var $\phi, \psi, \tau$: nontemporal formula; $a$: substitution;
begin
  (1) for each "from $\phi$ always $\psi$ until $\tau$" $\in C$ do
      for each $a$ do
        if $\text{VALID}(\phi, a)$ then INSERT($C_u$, $\langle \psi, \tau, a \rangle$);
  (2) for each $\langle \psi, \tau, a \rangle \in C_u$ do
      if $\text{VALID}(\psi, a)$ then DELETE($C_u$, $\langle \psi, \tau, a \rangle$)
      else if $\text{not VALID}(\psi, a)$ then ERROR;
  (3) for each "after $\phi$ sometime $\psi$ before $\tau$" $\in C$ do
      for each $a$ do
        if $\text{VALID}(\phi, a)$ then INSERT($C_e$, $\langle \psi, \tau, a \rangle$);
  (4) for each $\langle \psi, \tau, a \rangle \in C_e$ do
      if $\text{VALID}(\psi, a)$ then ERROR
      else if $\text{VALID}(\psi, a)$ then DELETE($C_e$, $\langle \psi, \tau, a \rangle$);
end CHECK.

It should be noted that the sequence of actions is essential within the two loop pairs (1);(2) and (3);(4), whereas these sections may be executed in parallel. The reader should verify that especially the critical situations, when start and termination conditions of a constraint coincide, are handled in accordance with former definitions.

The following theorem confirms that maintaining the sets $C_u$ and $C_e$ as above suffices to guarantee admissibility of a state sequence, if the present prefix has been accepted and if the (future) rest will satisfy the constraints in $C_u$, $C_u$ and $C_e$.

We only consider state sequences ($\sigma_0 \sigma_1 \ldots$) s.t. there are transactions inducing state transitions $\sigma_i \rightarrow \sigma_{i+1}$ (via the procedure EXECUTE). We say that, for $n \geq 0$, DB-MONITOR accepts ($\sigma_0 \ldots \sigma_n$) iff $\sigma_0, \ldots, \sigma_n$ are the states generated during the first $n$ runs of DB-MONITOR taking appropriate transactions as input and not calling ERROR. Let $C_u^n, C_e^n$ denote the contents of the variables $C_u, C_e$ after the $n$-th execution of the while-loop in DB-MONITOR.

Theorem: Let $n \geq 0$ be given. The state sequence $\sigma = (\sigma_0 \sigma_1 \ldots)$ is admissible iff the conditions (i)-(iv) hold:
(i) DB-MONITOR accepts \( \sigma_0 \ldots \sigma_n \)

(ii) \( \sigma^{n+1} \) is admissible (i.e., \( \sigma^{n+1} \) satisfies C)

(iii) \( [\sigma_{n+1}, a] \models (\text{always } \psi \text{ until } \tau) \) for each \(<\psi, \tau, a> \in C_{u}^n\)

(iv) \( [\sigma_{n+1}, a] \models (\text{sometime } \psi \text{ before } \tau) \) for each \(<\psi, \tau, a> \in C_{e}^n\)

Thus, in each present state \( \sigma_n \) successfully reached by DB-MONITOR we know that the current state sequence is admissible, if we optimistically assume every constraint to be satisfiable in the future (\( \sigma^{n+1} \)). In reality, there may be combinations of \( C, C_u \) and \( C_e \) excluding any continuation to an admissible state sequence, so that the monitor should be enhanced by some pre-computations. We will not follow this possibility here, but we have already considered a typical situation in [8].

The above theorem can be concluded from some "invariants" of the program DB-MONITOR.

**Lemma:** For all \( n \geq 0 \), the following properties hold after the \( n \)-th execution of the while-loop:

(i) \( \forall \phi \exists j, k \in \mathbb{N} : \langle \phi, \text{always } \psi \text{ until } \tau \rangle \in C_{u} \quad \forall j, k \leq n : [\sigma_{j, a}] \models \phi \iff U_1(n, j, \psi, \tau, a) \lor U_2(n, j, \psi, \tau, a) \)

where \( U_1(n, j, \psi, \tau, a) = (\exists k, j \leq k \leq n : [\sigma_{k, a}] \models \tau \wedge \forall k, j \leq k \leq n : [\sigma_{k, a}] \models \psi) \)

and \( U_2(n, j, \psi, \tau, a) = (\exists j, k \leq j \leq n : [\sigma_{j, a}] \models \tau \wedge \forall k, j \leq k \leq n : [\sigma_{k, a}] \models \psi) \)

(ii) \( \forall \psi, \tau \forall a : \langle \psi, \tau, a \rangle \in C_{u} \iff \exists \phi \exists j, k \in \mathbb{N} : \langle \phi, \text{always } \psi \text{ until } \tau \rangle \in C_{u} \wedge [\sigma_{j, a}] \models \phi \land U_2(n, j, \psi, \tau, a) \)

(iii) \( \forall \phi \exists j, k \in \mathbb{N} : \langle \phi, \text{sometime } \psi \text{ before } \tau \rangle \in C_{e} \quad \forall j, k \leq n : [\sigma_{j, a}] \models \phi \iff E_1(n, j, \psi, \tau, a) \lor E_2(n, j, \psi, \tau, a) \)

where \( E_1(n, j, \psi, \tau, a) = (\exists k, j \leq k \leq n : [\sigma_{k, a}] \models \psi \land \exists k, j \leq k \leq n : [\sigma_{k, a}] \models \tau) \)

and \( E_2(n, j, \psi, \tau, a) = (\exists j, k \leq j \leq n : [\sigma_{j, a}] \models \psi \land \exists j, k \leq j \leq n : [\sigma_{j, a}] \models \tau) \)

(iv) \( \forall \psi, \tau \forall a : \langle \psi, \tau, a \rangle \in C_{e} \iff \exists \phi \exists j, k \in \mathbb{N} : \langle \phi, \text{sometime } \psi \text{ before } \tau \rangle \in C_{e} \wedge [\sigma_{j, a}] \models \phi \land E_2(n, j, \psi, \tau, a) \)

Properties (i) and (iii) describe to what extent constraints have been satisfied up to the actual \( \sigma \); (ii) and (iv) characterize \( C_{u} \) and \( C_{e} \) to consist of constraints that have been only "partially" satisfied (predicates \( U_2, E_2 \)). The lemma can be proved using induction on \( n \). Additionally, (i) \( \wedge \) (iii) is a sufficient condition that ERROR has not been called.
As a corollary to the theorem, we obtain an operational admissibility criterion for finite state sequences.

**Corollary:** The state sequence $\sigma = (\sigma_0 \ldots \sigma_n)$ is admissible iff it is accepted by $DB\text{--}MONITOR$ s.t. $C^n_e = \emptyset$.

**Proof:** Since the continuation $\sigma^{n+1}$ is defined to be the constant sequence $(\sigma_n \sigma_n \ldots)$, all constraints in $C$ affecting $\sigma^{n+1}$ must have induced corresponding notes in $C_u$ or $C_e$ already during $\sigma_n$; i.e., condition (ii) of the theorem becomes irrelevant. If the prefix is accepted, constraints in $C_u$ remain satisfied and constraints in $C_e$ remain unsatisfied, so that conditions (iii) and (iv) reduce to $C^n_e = \emptyset$.

This result does not only apply to cases when the database operation really gets stopped, but it also shows the special role of the set $C_e$ in estimating the "consistency" of a database. $C_e$ contains exactly every condition $(\psi)$ for which a future state must exist s.t. it becomes valid. Only if $C_e$ is empty, the database is in a fully consistent state; otherwise some transactions disposing of those conditions still need to be executed.

5. CONCLUSIONS

We have introduced a logic calculus for specifying dynamic database behaviour not only with a modeltheoretic semantics, but also with an operational semantics. Of course, this paper could only provide principal items of a specification language and method; some further adjustment to practical needs is still missing. There also remains a gap between the monitor frame given in the last section and any monitor implementation. It can partly be filled by existing work on the enforcement of static constraints. Our emphasis is on handling the non-static aspects of constraints which seem to have been neglected before.

On the theoretical side, it would be interesting to look for a monitoring scheme discovering any future constraint violation at the earliest possible moment. Therefore criteria are needed to decide whether a continuation to an admissible state sequence does exist or not; probably some pre-computations or logical deductions checking dynamic constraints for possible contradictions have to be done by a monitor. Especially, the available transactions should be taken into account.

REFERENCES


