Calculus-based Runtime Verification

Xuan Qi, Changzhi Zhao

Abstract—In this paper, a uniform calculus-based approach for synthesizing monitors checking correctness properties specified by a large variety of logics at runtime is provided, including future and past time logics, interval logics, state machine and parameterized temporal logics. We present a calculus mechanism to synthesize monitors from the logical specification for the incremental analysis of execution traces during test and real run. The monitor detects both good and bad prefix of a particular kind, namely those that are informative for the property under investigation. We elaborate the procedure of calculus as monitors.

Keywords—calculus, eagle logic, monitor synthesis, runtime verification

I. INTRODUCTION

RUNTIME verification (RV) is an emerging lightweight verification technique in which executions of systems under scrutiny are checked for satisfaction or violation of given correctness properties, while it complements verification techniques such as model checking and testing, it also paves the way for not-only detecting incorrect behavior of a software system but also for reacting and potentially healing the system when a correctness violation is encountered.

Typically, monitors are generated automatically from some high-level specification. Runtime verification, which has its roots in model checking, often employs some variant of linear temporal logic, such as Pnueli’s LTL [2]. Meanwhile some more formal methods are required to describe different requirement, for example, future and past time logics, interval logics, state machine and parameterized temporal logics. Accordingly, a lot of monitor construction method has been studied. [3] presents a rewriting algorithm for efficiently testing future time linear Temporal logic formulae. [4] makes use of the characterization that finite trace LTL can be defined recursively, both on the structure of the formulae and on the size of the executing trace, and presents that an efficient dynamic programming algorithms can be generated from any LTL formulae. The commercial tool Temporal Rover (TR) [5,6] supports a fixed future and past time LTL, with the possibility of specifying real-time and data constraints as annotations on the temporal operators. It implementation is based on alternating automata. Algorithms using alternating automaton to monitor LTL properties are also proposed in [7].

In [8], the approach consists of translating LTL formulae to finite-state automaton, the translation algorithms modifies standard LTL to Büchi automaton conversion techniques to generate automata that check finite program traces.

First, it’s clearly that above monitor construction methods orient to special monitor logic. Different monitor construction methods have been given for different specification logic. Second, the above methods are all based on finite trace semantics. The standard semantics of above specification logic are often based on infinite state sequences, but during the running, you can only see the finite prefix of the infinite run at any time. So it is rational for the monitor to give the verdict of the infinite run based on the finite prefix. But the above finite trace semantics are all inconsistencies. In other words, the monitor semantics (finite trace semantics) often are not consistent with the standard inconsistencies. A logic L is called consistent if there exists not a model M and a formula ϕ∈L such that M |= ϕ and M |= ¬ϕ. For example, consider the LTL formulae ϕ=Xp and ϕ’=X¬p, u∈Σ* be a finite trace consisting of only a single element, such that u=a with a∈{p}. Theoretically, in a finite trace interpretation of LTL, we could have both u|=ϕ and u|=¬ϕ, since there is no successor action available to satisfy either case. But sometimes the consistency between finite trace semantics and standard semantics of the specification logic is vital important. We call the consistency of the monitor logic as impartiality.

How to design a unifying logic in which all these logics can be modeled and a unifying framework in which different monitor constructions are as the same way. And how to make the monitor satisfied the property of impartiality are two critical questions which will be solved in this paper. [9,10] present a rule-based framework, called EAGLE, that has been shown to be capable to defining and implementing a range of finite trace monitoring logics, including future and past time temporal logic, real-time and metric temporal logics, forms of quantified temporal logics, and so on. Our work is significantly influenced by idea and implementation of EAGLE. But the monitor semantics does obey impartiality and the corresponding monitor identifies the a particular kind of good and bad prefix, namely informative good/bad prefix for the property under investigation.

The remainder of this paper is organized as follows. Section 2 discusses two foundation works: informative prefix and Eagle logic. Section 3 introduces the calculus process for varieties of primitive operators, next-time operator (○), previous-state operator (○), and rule definitions. Then shows the workings of the calculus algorithm through two examples, finally gives the concise proof of correctness. Section 4 closes the paper with discussion and conclusions.

Xuan Qi is with Beijing Institute of System Engineering, P.R.China (Beijing 9702 Box, 100101, P.R.China ; e-mail: qixuanhappy@126.com).
Changzhi Zhao is with School of Computer, National University of Defense Technology, P.R.China.
II. PRELIMINARIES

A. Informative prefix

Various approaches to runtime verification and reasoning about systems based on truncated paths have been based upon a seminal paper by Kupferman and Vardi[1]. In [1], safety formulas are classified into three kinds, the intentionally safe the accidentally safe and the pathologically safe, depending on the kinds of prefixes their properties possess. A prefix σ is called informative for a formula if it “tells the whole story”[1] of why the formula holds for every infinite state sequence of which σ is a prefix. Intentionally safe formulas are formulas of which every bad prefix is informative (e.g. $\square p$), an accidentally safe formula is a safety formula of which all state sequences that violate it, do have some informative bad prefix (e.g. $\square (p \lor O\neg q)$, examples from [1]) . Pathologically safe formulas are formulas that have computations that violate it without any informative bad prefix.

Formally, the definition of informative prefix can be defined as follow:

Definition I Let $\sigma = a_0 a_1 \ldots a_n \in \Sigma^*$ be a finite state sequence, $\sigma$ is informative for $\phi$ iff there exists a finite sequence $Tr \in (2^{LTL})^*$ of sets of formulas, for each $n \leq |Tr|$, such that $\sigma \in Tr(0)$,

$Tr(n+1)$ is empty,

for all $0 \leq i \leq n$ and $\psi \in Tr(i)$, the following hold.

- if $\psi$ is an atomic proposition, then $\psi \in a_i$.
- if $\psi = \psi_1 \lor \psi_2$, then $\psi_1 \in Tr(i)$ or $\psi_2 \in Tr(i)$.
- if $\psi = \psi_1 \lor \psi_2$, then $(1)$ $\psi_1 \in Tr(i)$ or $(2)$ $\psi_1 \in Tr(i)$ and $\psi_2 \in Tr(i + 1)$.

We call such a sequence Tr an informative sequence. If such an informative sequence exists, it tells us why $\phi$ holds for any extension of the prefix $\sigma$. It indicates what formulas hold at what moment of the prefix and why. Since $Tr(i)$ is at some point empty, this reasoning is complete and thus applies to any extension of the prefix. For example, if $\varphi_1 \land \varphi_2 \in Tr(i)$, according to the definition, $\varphi_1 \in Tr(i)$ and $\varphi_2 \in Tr(i)$, which tells us that $\varphi_1 \land \varphi_2$ holds for any extension of the prefix. The informative bad prefixes can be considered as the only proper counterexamples, since they demonstrate why the formula does not hold or hold. So it is helpful to fault diagnosis and fault localization. To find the informative prefix of the property under investigation is one of the main goal for runtime verification. [11] elaborates on the construction of the monitor for temporal logic properties in which the automaton forms the basis of a monitor that detects both good and bad informative prefix for the property under investigation. In this paper, we will give a monitor construction method based on calculus which support much more formal property specification, including future and past time logics, interval logics, state machine and parameterized temporal logics.

B. Eagle

The Eagle logic is designed to support infinite trace monitoring, and contains a small set of powerful operators, which allow on to define new logics on top. Eagle essentially supports recursive parameterized equations, with a minimal/maximal fix-point semantics, together with three temporal operators: next-time, previous-time and concatenation. The equations are also called as rules. Rules can be parameterized with formulas, supporting the definition of new temporal operators, and they can be parameterized with values, thus supporting logics that can reason about data and as a special case of data, real-time. Here we assume boolean expressions over individual states as automatic propositions which comprise the finite trace. The expressiveness of the logic system is rich. Actually, any linear-time temporal logic, whose temporal modalities can be recursively defined over the next, past or concatenation modalities, can be embedded within it. Meanwhile the logic has supported a limited form of quantification. Interesting reader can refer to [9, 10] for details. We present the syntax and semantics below:

Syntax The syntax of EAGLE is shown in figure 1. A specification is consists of a declaration part D and an observer part O. D comprises zero or more rule definitions R, and O comprises zero or more monitor definitions M, which specify what is to be monitored. Rules and monitors are both named (N). Each rule definition is preceded by one of the keywords min or max, indicating at the end of the trace how to interpret the semantic of the rules. A parameter type can either be form, representing formulas, or a primitive type int, long, float, etc. The body of a rule/monitor is a Boolean valued formula of the syntactic category Form. Any recursive call on a rule must be strictly guarded by a temporal operator. The propositions of the logic are Boolean expressions over an observer state. Formulas are composed using standard proposition logic operators together with a next-time operator (\(\circ\) F), a previous-state operator ( F) and a concatenation operator ( F) and a concatenation operator $F_1 \bullet F_2$.

$$\begin{align*}
S & := DO \\
D & := R^* \\
O & := M \\
R & := max | min(N(T_1 x_1, ... , T_n x_n) = F) \\
M & := mon N = F \\
T & := Form | primitive type \\
F & := expression \mid true \mid false \mid \neg F \mid F_1 \land F_2 \mid F_1 \lor F_2 \mid F_1 \rightarrow F_2 \mid F_1 \bullet F_2 \mid (N(F_1, \ldots , F_n)) x_i
\end{align*}$$

Fig. 1 Syntax of EAGLE

Semantics The model of EAGLE logic are execution traces. An execution trace $\sigma$ is a finite sequence of program states $\sigma = s_0 s_1 \ldots s_n$ where $|N| = n$ is the length of the trace. The ith state $s_i$ is denoted by $\sigma(i)$. The term $\sigma(i)$ denotes the sub-trace of $\sigma$ from position i to position j, both position is included. The semantics of the logic is defined in terms of a satisfaction relation between execution traces and specifications. That is, given a trace $\sigma$ and a specification D O, satisfaction is defined as follows: $\sigma|D \sqsubseteq O$ iff $\forall n \in F, \sigma_1|D \sqsubseteq O, l|D \sqsubseteq F$. That is to say, if the trace, observed from position 1 (the first state) satisfied each monitored formula in a specification, the trace
satisfied the specification. The definition of the satisfaction relation \( \sigma_0 \subseteq \langle \text{Trace} \times \text{nat} \rangle \times \text{Form} \), for a set of rule definitions \( D \), is presented in Figure 2.

\[
\sigma, i || \sigma, i + 1 \models \text{update}(F, \sigma(i)) \quad \text{if} \quad \sigma, i || \sigma, i + 1 \models F.
\]

\( i \) is the current index of the observation. \( \sigma, i \) is the monitored formula. \( F \) is the formula to be evaluated. The result of \( \text{update}(\sigma, i)||\sigma, i + 1 \models F \) is either \( \text{true} \) or \( \text{false} \). The definition of \( \text{update}(\sigma, i)||\sigma, i + 1 \models F \) is given in Figure 2.

The definition of the satisfaction relation \( \sigma_0 \subseteq \langle \text{Trace} \times \text{nat} \rangle \times \text{Form} \), for a set of rule definitions \( D \), is presented in Figure 2.

\[
\sigma, i || \sigma, i + 1 \models \text{update}(F, \sigma(i)) \quad \text{if} \quad \sigma, i || \sigma, i + 1 \models F.
\]

\( i \) is the current index of the observation. \( \sigma, i \) is the monitored formula. \( F \) is the formula to be evaluated. The result of \( \text{update}(\sigma, i)||\sigma, i + 1 \models F \) is either \( \text{true} \) or \( \text{false} \). The definition of \( \text{update}(\sigma, i)||\sigma, i + 1 \models F \) is given in Figure 2.

\[
\sigma, i || \sigma, i + 1 \models \text{update}(F, \sigma(i)) \quad \text{if} \quad \sigma, i || \sigma, i + 1 \models F.
\]

\( i \) is the current index of the observation. \( \sigma, i \) is the monitored formula. \( F \) is the formula to be evaluated. The result of \( \text{update}(\sigma, i)||\sigma, i + 1 \models F \) is either \( \text{true} \) or \( \text{false} \). The definition of \( \text{update}(\sigma, i)||\sigma, i + 1 \models F \) is given in Figure 2.

\[
\sigma, i || \sigma, i + 1 \models \text{update}(F, \sigma(i)) \quad \text{if} \quad \sigma, i || \sigma, i + 1 \models F.
\]

\( i \) is the current index of the observation. \( \sigma, i \) is the monitored formula. \( F \) is the formula to be evaluated. The result of \( \text{update}(\sigma, i)||\sigma, i + 1 \models F \) is either \( \text{true} \) or \( \text{false} \). The definition of \( \text{update}(\sigma, i)||\sigma, i + 1 \models F \) is given in Figure 2.

\[
\sigma, i || \sigma, i + 1 \models \text{update}(F, \sigma(i)) \quad \text{if} \quad \sigma, i || \sigma, i + 1 \models F.
\]

\( i \) is the current index of the observation. \( \sigma, i \) is the monitored formula. \( F \) is the formula to be evaluated. The result of \( \text{update}(\sigma, i)||\sigma, i + 1 \models F \) is either \( \text{true} \) or \( \text{false} \). The definition of \( \text{update}(\sigma, i)||\sigma, i + 1 \models F \) is given in Figure 2.
we simply propagate the arguments to the subformula. The only difference from the counterparts in the traditional EAGLE logic [9][10] is the final-value definition on formulas except true and false. At the end of finite trace, if the result formula is true, it shows that the finite trace is the informative good sequence for the property under investigation, if the result formula is false, it shows that the finite trace is the informative bad sequence for the property. According to the definition 1, the calculus process identifies the informative prefix for the property under investigation. Otherwise it shows that the property has no informative prefix or the current finite trace is only the proper prefix of its informative prefix. So we can not given the monitoring result only based on the current observed finite trace.

The functions transform, eval, update are defined in a special way for operators ○ and −. For the operator ○ we introduce the operator Next: Form → Form. Then we define transform, eval, update as follows:

\[ \text{transform}(\circ F, Z, b) = \text{Next}(\text{transform}(F, Z, b)) \]

\[ \text{eval}(\text{Next}(F, s)) = \text{update}(F, s, \text{null, null}) \]

\[ \text{update}(\text{Next}(F), s, Z, b) = \text{Next}(\text{update}(F, s, Z, b)) \]

The operator requires special attention. If a formula F is guarded by a previous operator then we evaluate F at every event and use the result of this evaluation in the next state. Thus, the result of evaluating F is required to be stored in some temporary placeholder so that it can be used in the next state. To allocate a placeholder for a operator, we introduce the operator Previous: Form → Form. We define transform, eval, update for as follows:

\[ \text{transform}(F, Z, b) = \text{Pr eval}(Y, \text{eval}(Y)) \]

\[ \text{where } Y = \text{transform}(F, Z, b) \]

\[ \text{eval(Pr eval}(F, \text{past}, s)) = \text{eval(Pr eval}(F, \text{past}, s)) \]

\[ \text{update(Pr eval}(F, \text{past}, s, Z, b) = \text{Pr eval}(\text{update}(F, s, Z, b)) \]

Here, the definitions of the three functions on operator are as same as the two-valued EAGLE logic. In update function we not only update the first argument F but also evaluate F and pass is as the second argument of Previous. Note that the final-eval function only is used at the end of the finite trace, it only concerns whether the result formula is true/false or not. So it does not need to be defined on ○ and − more. The reason is same for rule application below.

B. Monitor Synthesis for Rules

In this paper, we will give different forms of rule definitions. In traditional Eagle logic [9][10], without loss of generality, the standard form of a rule is \[ \text{R}(\text{Form } f_1, \ldots, \text{Form } f_n, \text{T}_i, p_1, \ldots, T_n, p_n) = B \] where \( f_1, \ldots, f_n \) are arguments of type Form and \( p_1, \ldots, p_n \) are arguments of primitive type. There the rule definition is divided into two styles: max rules and min rules. But in this paper, it is not needed any more, because final-eval function is not dependent on the rule types.

Without loss of generality, in this paper, the standard form of a rule is \( \text{R}(\text{Form } f_1, \ldots, \text{Form } f_n, \text{T}_i, p_1, \ldots, T_n, p_n) = B \) where \( f_1, \ldots, f_n \) are arguments of type Form and \( p_1, \ldots, p_n \) are arguments of primitive type. Such a rule can be written in short as:

\[ \text{R}(\text{Form } \bar{f}, \bar{T}, \bar{p}) = B \]

Where \( \bar{f} \) and \( \bar{p} \) represent tuples of type \( \text{Form} \) and \( \bar{T} \) respectively. For such a rule we introduce an operator \( \overline{R} : \text{Form} \times \bar{T} \rightarrow \text{Form} \). Informally, the first argument of \( \overline{R} \) represents the transformed right hand side of the rule.

For the rule \( \overline{R}(\text{Form } \bar{f}, \bar{T}, \bar{p}) = B \) the definitions of transform, eval, update are synthesized as follows:

\[ \text{transform}(\overline{R}(\text{Form } \bar{f}, \bar{T}, \bar{p}), \text{R}(\text{Form } \bar{f}, \bar{T}, \bar{p}), b) = \overline{R}(b, \bar{T}) \]

\[ \text{transform}(\overline{R}(\text{Form } \bar{f}, \bar{T}, \bar{p}), Z, b) = \overline{R}(\text{pb}, \text{transform}(\text{Form } \bar{f} \rightarrow \text{Form } \bar{T}, \text{R}(\text{Form } \bar{f}, \bar{T}, b'), \bar{T}) \]

\[ \text{where } Z = \text{transform}(F, Z, b) \text{ and } Z \text{ not match } R(F, ?) \]

update(\overline{R}(\text{pb}, H(b), \bar{T}), s, \overline{R}(\text{pb}, H(b), \bar{p}, b') = \overline{R}(b', \bar{T}) \]

update(\overline{R}(\text{pb}, H(b), \bar{T}), s, Z, b) = \overline{R}(\text{pb}, \text{update}(H(b), \bar{T}), s, \overline{R}(\text{pb}, H(b), \bar{T}), \bar{T}) \]

\[ \text{where } Z \text{ not match } \overline{R}(\text{pb}, H(b), ?) \]

eval(\overline{R}(\text{pb}, H(b), \bar{T}), s) = \overline{R}(H(b), \bar{T}) \rightarrow eval(\bar{T}, s), s \]

Note that the result of \( \text{eval}(P, s) \), where \( P \) is an expression, may be a partially evaluated expression if some of the variables referred to by the expressions are partially evaluated. The expression gets fully evaluated once all the variables referred to by the expressions are fully evaluated. The reader can refer to [9][10] for the detail.

C. Examples

We provide one example to show the workings of above calculus process which identifies the informative prefix for the property under investigation. In order to compare with traditional EAGLE logic, we use the example in [9] but different result will give.

Example the property under investigation which is in modified Eagle form is:

\[ \text{Ep(Form } f) = f \rightarrow \text{Ep}(f) \]

\[ \text{mon } M = \circ \text{Ep}(q) \]

The finite trace is \( \sigma = [q] \}. \)

First, transform function is applied:

\[ \text{transform}(\circ \text{Ep}(q), \text{null, null}) \]

\[ \text{Next(\text{Ep}(\text{pb}(q \rightarrow \text{pr evious(\text{Ep}(b), false)))), ))} \]

Second, eval function is applied on state \( \sigma_1 = \{ q \}: \)

\[ \text{eval(\text{transform}(\circ \text{Ep}(q), \text{null, null}), \sigma_1) = \text{Ep}(\text{pb}(q \rightarrow \text{pr evious(\text{Ep}(b), true)))} \]

Third, eval function is applied again on state \( \sigma_2 = \{ \}: \)

\[ \text{eval(\text{Ep}(\text{pb}(q \rightarrow \text{pr evious(\text{Ep}(b', true)))), \sigma_2) \]

\[ 1. = \text{eval}(q \rightarrow \text{pr evious(\text{Ep}(\text{pb}(q \rightarrow \text{pr evious(\text{Ep}(b'), true))), true), \sigma_2) \]

\[ 2. = \text{eval}(q, \sigma_2) \rightarrow \text{eval(Pr evious(\text{Ep(\text{pb}(q \rightarrow \text{pr evious(\text{Ep}(b'), true))), true), \sigma_2) \]

\[ 3. = q \rightarrow \text{eval(\text{true}, \sigma_2) \]

\[ 4. = q \rightarrow \text{true } = \text{true} \]

Finally, at the end of the trace, the final-eval function is applied on the result formula:

\[ \text{final = eval(\text{true}) = true} \]

It is easy to see that \( \sigma = [q] \} \) is the informative good prefix of the property and the calculus process identifies the
informative prefix and gives the verdict result: true. If \( \sigma = \{q\} \), in traditional Eagle logic, the result of calculus will be false, because in [9], at the end of the trace, if the result formula is min rule, the verdict is false, if the result formula is max rule, the verdict is true. There the definition of Ep rule will be:

\[
\min(Form\ f) = f \lor \ Ep(f)
\]

Which is min rule, So at the end of the trace (second step), the calculus process of traditional Eagle will give the verdict: false. But based on our calculus, the calculus process will give verdict: ?. it shows that the current trace \( \{q\} \) is only the proper prefix of informative prefix \( \{p\} \).

IV. CONCLUSION

In this paper, a calculus-based approach for synthesizing monitors checking correctness properties specified in multiple kinds of logic which can be represented by modified Eagle logic. Different from the traditional Eagle logic, the rule definition in modified Eagle logic does not distinguish with the max and min rule. So at the end of the trace, the verdict result is not dependent on the rule type, but only concerns with whether the result formula is true, false or otherwise. Indeed the calculus process provides a informative sequence for the property under investigation. So the calculus process identifies the informative good or bad prefix for the property under investigation.

REFERENCES