Automatic Verification of Combined Specifications: An Overview

Ernst-Rüdiger Olderog

Department of Computing Science
University of Oldenburg
Oldenburg, Germany

Abstract

This paper gives an overview of results of the project “Beyond Timed Automata” carried out in the Collaborative Research Center AVACS (Automatic Verification and Analysis of Complex Systems) of the Universities of Oldenburg, Freiburg, and Saarbrücken. We discuss how properties of high-level specifications of real-time systems combining the dimensions of process behaviour, data, and time can be automatically verified, exploiting recent advances in semantics, constraint-based model checking, and decision procedures for complex data.

As specification language we consider CS-OZ-DC, which integrates concepts from Communicating Sequential Processes (CSP), Object-Z (OZ), and Duration Calculus (DC). Our approach to automatic verification of CSP-OZ-DC rests on a compositional semantics of this languages in terms of Phase-Event-Automata. These can be translated into Transition Constraint Systems which serve as an input language of an abstract refinement model checker called ARMC which can handle constraints covering both real-time and infinite data. This is demonstrated by a case study concerning emergency messages in the European Train Control System (ETCS). For CSP-OZ-DC we also discuss a UML profile and tool support.

Keywords: Real-time systems, complex data, CSP, Object-Z, Duration Calculus, model checking, abstraction refinement, UML profile, tool support

1 Introduction

Computers are more and more used to control the behavior of complex systems, for instance in the traffic domain. Such applications are typically safety critical, i.e., a malfunction of the computers is costly and dangerous. Think of assistance systems that should guarantee the collision freedom of traffic agents such as cars, trains, and planes. Such applications necessitate the use of formal models of the

1 This work was partly supported by the German Research Council (DFG) as part of the Transregional Collaborative Research Center “Automatic Verification and Analysis of Complex Systems” (SFB/TR 14 AVACS, http://www.avacs.org/).
2 Email: olderog@informatik.uni-oldenburg.de
overall system and of formal verification for establishing the relevant safety properties. The models must be able to represent various aspects of the systems such as state spaces and their transformation, communication between system components, real-time constraints, interfaces to a continuously evolving physical environment, and dynamically changing system structures. To cope with such models in a manageable way, combined specification techniques have been proposed, integrating well researched specification techniques for individual system aspects. It is a major research challenge to develop methods for the automatic verification and analysis of such combined specifications modeling complex real-life systems.

To address this challenge the research center AVACS (Automatic Verification and Analysis of Complex Systems) was initiated in 2004. In AVACS, researchers of the Universities of Oldenburg, Freiburg and Saarbrücken as well as the Max-Planck-Institute for Informatics in Saarbrücken collaborate [2]. The idea of AVACS is to bring experts in semantic modeling and specification together with experts in verification and analysis techniques. Research in AVACS is organized in four layers:

(i) Complex Systems:
   e.g., the European Train Control System (ETCS)

(ii) Models of Complex Systems:
   real-time – hybrid – systems of systems

(iii) Combining Verification and Analysis Technologies:
   combine technologies $t_1, ..., t_n$ for system $s$

(iv) Verification and Analysis Kernel Technologies:

At the top layer (i) are complex systems like the European Train Control System (ETCS). In ETCS trains communicate wireless with radio block centers (RBCs) that control the traffic in certain areas (see Fig. 1). The RBCs grant movement authorities for trains up to a position closely behind the preceding train. In case of an emergency incident of the first train, the RBC has to ensure that this train and all successive trains will stop safely in order to avoid collisions.

At the bottom layer (iv) there are various individual verification and analysis technologies like “Abstraction” or “Heuristic Search”. The idea is to combine at layer (iii) such technologies in a suitable novel way so that particular system classes can be verified and analyzed. To be successful with such a combination, AVACS pursues a divide and conquer strategy whereby (in the first phase of the project) systems are classified into real-time systems, hybrid systems, and systems of systems. The corresponding research areas are called R, H, and S, each organized into three subprojects.

In this paper we give an overview of one subproject on real-time systems. These are systems that interact with their environment in such a way that for certain inputs the corresponding outputs have to occur within given time bounds. Many embedded systems, in particular those in safety critical applications like the ETCS, are of this
The subproject presented here is called “R1: Beyond Timed Automata” and it is coordinated by the author. R1 is motivated by the observation that currently the behavioral verification of specifications of real-time systems is based on their representation as timed automata and relies on model checkers like UPPAAL [20]. A limitation of this approach is that model checking with timed automata is restricted to real-time systems with finite data only. However, in applications systems often exhibit both real-time and complex, infinite data structures.

The goal of R1 is to advance the state of the art in automatic verification of high-level specifications of systems with the three dimensions of process behaviour, data, and real-time — beyond the capabilities of timed automata. In the first phase of R1, the core activities comprised the development of a system specification language, an approach to the automatic verification of real-time properties, and the application to the case study ETCS. As system specification language, CSP-OZ-DC (combining subsets from Communicating Sequential Processes, Object-Z, and Duration Calculus) was developed [18,16]. A key result in this development was a compositional semantics on the basis of Phase Event Automata (PEA), an extension of timed automata to represent data [16]. It involves a translation of the DC subsets of counterexample formulas (with events) and so-called test formulas into equivalent PEA. It was shown that PEA can be translated into Transition Constraint Systems (TCS), which serve as input for the abstraction refinement model checker ARMC [29] and the deductive slicing abstraction model checker SLAB [4]. While ARMC is based on predicate abstraction, SLAB is a combination of deductive model checking (based on Craig interpolation) and slicing. Both tools call decision procedures when checking entailment of constraints [13,36,37] as well as methods for computing interpolants [35,32]. By combining CSP-OZ-DC with ARMC (or SLAB) and decision procedures, properties of systems with both real-time constraints and (certain) infinite data types can be verified automatically, as demonstrated by case studies [17]. In particular, real-time properties of emergency messages in the ETCS case study were verified [23,11,22].

These core activities were complemented by research into reducing the size of the state spaces of specifications with the help of slicing techniques. This approach has been applied both at the level of CSP-OZ-DC [5,3] and at the level of TCS [4].
This paper is organized as follows. In Section 2 the combined specification language CSP-OZ-DC is outlined. In Section 3 an approach to automatic verification of real-time properties of CSP-OZ-DC specification is presented. Section 4 reports on case studies performed in this setting. Section 5 describes the tool support available for the approach, and Section 6 concludes the paper.

2 Combined Specifications

To specify real-time systems with a rich data part, we developed a high-level system specification language called CSP-OZ-DC [18,16] which separates the aspects of process behavior, data, and time. CSP-OZ-DC combines concepts from Communicating Sequential Processes (CSP) [14,31], Object-Z (OZ) [7,33], and the Duration Calculus (DC) [41,40]. The central notion is that of a class, consisting of an interface, a CSP part, an OZ part, and a DC part. An example of (part of) a class is shown Fig. 2. It models the rear train in the ETCS for the case of two trains.

The idea is that the rear train measures its position periodically and adjusts its speed so that it is always able to brake safely before reaching the limit of authority (LOA) along the track by applying its service brake. For this purpose, it periodically computes the service brake intervention limit (SBI), which represents the last position at which the train can apply the service brake in order to stop before the current end of authority.

The class has three parameters: the identity ID of the train, the position StartPos of the train at its start, and the position StartSBI of the SBI at its start. The interface declares channel names and types used by the class. Here we see the channels updPos for updating the train’s position and compSBI for computing the next SBI.

The CSP part constrains the sequencing of events (communication) along the interface channels using CSP process notation. It may consist of multiple processes defined by CSP process equations, one of which is a distinguished process named main, which denotes the initial process. Here we see that the CSP part consists of two subprocesses working in parallel (denoted by \parallel). The subprocess Running is taking care of the normal operation of speed control and the subprocess HandleEM takes care of emergency situations. The subprocess Running first inputs the current position pos of the train and the current limit of authority loa and then computes the SBI in the state variable sbi. If the position pos of the train has got beyond the value sbi the service brake has to be applied.

The OZ part specifies the state space and operations upon it. It comprises a nameless state schema describing the state space, a schema Init constraining the initial state, and communication schemas com_c describing the transformation of the state space induced by communicating along an interface channel c. Here we see that the state contains two variables sbi (for the current SBI) and curPos (for the current position) of type Position and curSpd (for the current speed) of type Speed. Of the communication schemas we exhibit the one for the channel compSBI which specifies how the new value of the variable sbi is defined. The notation \Delta(sbi)
of Object-Z defines that \( sbi \) is the only state variable changed by the schema.

The DC part constraints the timing of states and events. Here we see a so-called counterexample formula that states a lower time bound: any two successive communication events on the channel \( updPos \) should not be less than \( updBound \) seconds apart.

The ETCS with two trains consists of several classes. Besides \( RearTrain \) there are the classes for the \( LeadingTrain \), the RBC, the CommunicationNetwork (between trains and RBC), the Track, and the Driver [22]. Objects of classes may be combined into systems using the CSP operators of parallel composition and renaming.
2.1 UML Profile

Often specifiers use diagrams to support their understanding of a system. To facilitate this, a UML profile in the notations of UML 2.0 [39] has been developed for CSP-OZ [25] and extended to CSP-OZ-DC. The profile comprises class diagrams, protocol state machines, and component diagrams. These diagrams are annotated by suitable tags to represent the contents of classes in the form of Z and DC expressions. The semantics of the UML profile is given by a translation of the profile into CSP-OZ-DC as illustrated by Fig. 3. For details of the CSP-OZ part we refer to the paper [25].

![Fig. 3. Semantics of the UML profile for CSP-OZ-DC](image)

2.2 Operational Semantics

The key for an automatic verification of CSP-OZ-DC is an operational semantics of this language defined by J. Hoenicke on the basis of Phase Event Automata (PEA) [16]. PEA extend timed automata [1] such that the parallel composition synchronizes on both phases (state formulae) and events. This permits the a compositional semantics definition for CSP-OZ-DC, i.e., one satisfying the equation

\[
\mathcal{A}(\text{CSP-OZ-DC}) = \mathcal{A}(\text{CSP}) \parallel \mathcal{A}(\text{OZ}) \parallel \mathcal{A}(\text{DC})
\]

where \(\parallel\) denotes the (synchronous) parallel composition of PEA \(\mathcal{A}(\ldots)\). In fact, \(\mathcal{A}(\text{DC})\) decomposes even further into a parallel composition of PEA for each individual timing constraint in the DC part. An important property of this semantics is that whenever a subset of PEA in a parallel composition satisfies a requirement (represented as a DC formula) then also the full parallel composition does. This allows for a cone-of-influence verification technique.

For the DC part the class of counterexample formulae (with facilities to constrain the occurrences of both state changes and communication events) was introduced, extending the well-known class of “DC implementables” by A.P. Ravn [30]. The main theorem proved by J. Hoenicke in [16] is that every counterexample formula \(F\) has an operational semantics in form of a deterministic PEA \(\mathcal{A}(F)\) such that the runs of \(\mathcal{A}(F)\) are equivalent to the DC interpretations of \(F\). The proof of this theorem uses a so-called powerset construction to cope with the nondeterminism arising from overlapping phases in \(F\). Overlapping phases allow for concise specifications. The determinism of \(\mathcal{A}(F)\) permits an easy treatment of negation, which underlies
the specification with counterexample formulae as well as an automata-theoretic approach to model checking DC. In the latter approach, the desired property is negated and then represented as a PEA running in parallel to the system.

For the real-time requirements, R. Meyer extended the class of translatable formulae even further to so-called test formulae [21]. Whereas counterexample formulae are negated traces (of timed phases), test formulae contain arbitrary Boolean combinations of such traces and are closed under disjunction, conjunction, and the DC chop. Using so-called sync events, test formulae can be brought into a disjunctive normal form over traces and their negations, which facilitates their translation into PEA. To date, test formulae are the largest class of DC formulae that have an equivalent operational semantics in terms of automata [24].

Figure 4 shows the automaton PEA_{OZ} representing the semantics of the OZ part of the class in Fig. 2. Here updPos and comSBI are Boolean variables representing the presence or absence of a communication event on the corresponding channel in the CSP part.

Note that for each communication schema in the OZ part there is a corresponding transition labeled with the Boolean event variable and the formula of the communication schema, here com_{updPos} and com_{comSBI}. The idling transition $\phi_{idle}$ is taken if none of the communications in the OZ part of the class is enabled. Here $\phi_{idle}$ abbreviates the formula

$$\phi_{idle} \iff \neg \text{updPos} \land \neg \text{comSBI} \land \ldots \land \text{sbi} = \text{sbi}^{\prime} \land \text{curPos} = \text{curPos}^{\prime} \land \ldots$$

Figure 5 shows the automaton PEA_{DC} representing the semantics of the counterexample formula

$$\neg \Diamond ( \uparrow \text{updPos} ; \; \ell < 5 ; \; \text{updPos} )$$

in the DC part of the class in Fig. 2. Here $c$ is a clock that is used to measure the duration of 5 seconds.

3 Automatic Verification

We consider the problem whether a given specification CSP-OZ-DC satisfies a real-time requirement expressed by a DC formula. The aim is an automatic verification method. Our approach is illustrated by the following scheme:
CSP-OZ-DC satisfies DC?

\[
\begin{align*}
P EA: & \quad A(CSP) \parallel A(OZ) \parallel A(DC) \parallel A_{test}(DC) \\
&
\text{Is the bad state of } A_{test}(DC) \text{ reachable?}
\end{align*}
\]

In order to check whether a specification CSP-OZ-DC satisfies a real-time property, represented by a test formula DC, both the specification and the property are translated to Phase Event Automata running in parallel. The property DC is translated to a so-called test automaton \( A_{test}(DC) \), which has a distinguished bad state such that specification CSP-OZ-DC satisfies the test formula DC if and only if at the PEA level the bad state is reachable in \( A_{test}(DC) \) as part of the overall parallel composition.

To check for reachability we apply the abstraction refinement model checker ARMC developed by A. Podelski and A. Rybalchenko [28, 29]. ARMC takes as input Transition Constraint Systems (TCS). The PEA semantics of CSP-OZ-DC is very well suited as an intermediate language in the translation process from CSP-OZ-DC down to TCS. At the level of TCS, the clocks of PEA are represented as real-valued data variables, following the “old-fashioned recipe” advocated by L. Lamport. As an example consider the transition constraint system \( T(\text{PEA}_{DZ}) \) for the automaton PEA\(_{DZ}\) for the DC constraint shown in Fig. 5:

\[
\begin{align*}
T(\text{PEA}_{DZ}) &\iff ph = 0 \land \neg \text{updPos} \land c' = c + \text{len} \land ph' = 0 \\
&\lor ph = 0 \land \text{updPos} \land c' = \text{len} \land c' \leq 5 \land ph' = 1 \\
&\lor ph = 1 \land \neg \text{updPos} \land c' = c + \text{len} \land c' \leq 5 \land ph' = 1 \\
&\lor ph = 1 \land \text{updPos} \land c = 5 \land c' = \text{len} \land c' \leq 5 \land ph' = 1 \\
&\lor ph = 1 \land \neg \text{updPos} \land c = 5 \land c' = c + \text{len} \land ph' = 0
\end{align*}
\]

Here \( c \) is a real-valued variable representing the corresponding clock of PEA\(_{DZ}\) and \( \text{len} \) is a real-valued variable with the constraint \( \text{len} > 0 \) that represents time progress. The variables \( ph \) represents the current phase of PEA\(_{DZ}\).
### 3.1 Abstraction Refinement

Verification of temporal safety and liveness properties can be effectively automated by applying a reduction to least fixpoint computation [6,27]. Such a fixpoint computation engine serves as a basis for the verification tool ARMC [28,29]. ARMC is a model checking tool that applies abstraction refinement to efficiently handle the high complexity of verification tasks envisaged in the AVACS project. Its distinguishing characteristics lie in the way it applies logical reasoning to deal with abstraction [29]. ARMC is implemented in a Prolog system together with Constraint Logic Programming extensions. Interpolation is an important component of the abstraction refinement algorithm used by ARMC. It provides an effective means for computing the separation between the sets of ‘good’ and ‘bad’ states. ARMC uses an algorithm for the generation of interpolants for the combined theory of linear arithmetic and uninterpreted function symbols [32]. It uses a reduction of the problem to constraint solving in linear arithmetic, which allows for the application of existing highly optimized Linear Programming solvers in black-box fashion.

Note that the PEA and hence the TCS representing the semantics of CSP-OZ-DC specifications are in general infinite state systems due to both clocks and data values. So reachability is in general not decidable. Thus the fixpoint computation of ARMC need not terminate. However, as our case studies demonstrate, ARMC can be applied successfully to various examples.

### 4 Case Studies

A first application of this approach to verification dealt with a parametric elevator by J. Hoenicke and P. Maier [17]. In this example the number of floors are treated as parameters. A safety property that depended on all parts of the specification (i.e., communication, data, and time) was verified automatically with ARMC. The specification of the elevator in CSP-OZ-DC comprised both infinite data (i.e., integers representing an arbitrary number of floors) and continuous real-time.

#### 4.1 Emergency Messages

The benchmark for the project “R1: Beyond Timed Automata” was defined as the verification of timing requirements for the radio communication between trains and the radio block center (RBC) in the ETCS. Starting from a comprehensive but informal description of the ETCS in [9], J. Faber defined the case study Emergency Messages (EM) for the scenario where an RBC controls consecutive trains on a track segment (see Fig. 1). In the case of two trains, if the first train detects an emergency situation it immediately applies the emergency brake and sends an emergency message via the radio connection to the RBC, which has to inform the follower train within a predefined time interval. The train control system has to stop the follower such that no collision occurs. This property depends on several real-time requirements for the message transfer times and the reaction times of the RBC and the follower.
In [10,12,24], this case study is modeled in the specification language CSP-OZ-DC. The model involves continuous real time, real-valued variables representing train positions (on an infinite track segment) and speeds and messages transferred via CSP channels. Other quantities like the length of the train and the braking distance were treated as parameters. A part of one class of the specification is shown in Fig. 2. Then the techniques of Section 2 [17,16] were applied to translate the CSP-OZ-DC model via Phase Event Automata (PEA) into Transition Constraint Systems (TCS) that are the input of the abstraction refinement model checker ARMC [28,29].

For the case study EM, properties formalizing reaction times in the communication between trains and RBC could be verified automatically with ARMC. Thus the benchmark for R1 was fully achieved.

However, the global property of collision freedom could not be verified automatically with ARMC. The reasons are as follows. According to [16] each CSP-OZ-DC specification is represented as a parallel composition of PEA. In the case study this composition consisted of 18 automata. At present ARMC requires the parallel product of this composition to be computed. An attempt to compute the parallel product of all 18 automata of EM failed due to memory shortage. By contrast, for the verification of the reaction times in the benchmark case it is sufficient to consider only 5–7 of the automata and compute their parallel product. By the compositional semantics [16] of CSP-OZ-DC, this allows us to infer the verified property for the full parallel composition of all automata (without computing the product).

To prove collision freedom for the EM case study, a manual decomposition of this property into simpler subtasks was performed. Each of these subtasks was a variant of a reaction time property that could be verified automatically with ARMC.

Moreover, variants of the case study EM were examined which had a more sophisticated and realistic data part, but a less complicated control structure: the RBC maintains an array of consecutive trains (on an infinite track segment) where the size of the array is kept as a parameter. In case of an emergency, every train behind the emergency train has to be instructed to stop (Fig. 1). Message transfer times were not considered in this extended scenario. For this variant, collision freedom for an arbitrary number of trains [19,11] could be shown. To cope with the data type used for representing the train positions (in this case: arrays with integer elements and real numbers as elements, with a parametric dimension) methods for hierarchical reasoning in theories of complex data types developed in [34] were employed.

5 Tool Support

Tool support has been developed to handle system specifications expressed in CSP-OZ-DC, real-time requirements in form of test formulae, and their translation into Phase Event Automata (PEA) and Transition Constraint Systems (TCS).
5.1 Syspect

For CSP-OZ-DC specifications, a graphical modeling environment called Syspect (System Specification Tool) [38] has been implemented on the basis of the Eclipse platform [8]. An overview of Syspect is given in Fig. 6. The graphic modeling uses the UML profile mentioned in Subsection 2.1. See Fig. 7 for a screen shot of the class editor. The graphic model is automatically converted into an internal representation of CSP-OZ-DC, which offers the possibility to export the specification into their semantic model in terms of PEA for a subsequent verification. To this end, Syspect also permits to enter real-time requirements expressed as test formulae (see Section 2). Test formulae also serve as the slicing criterion in a Slicing Plugin of Syspect that has been implemented to perform slicing of CSP-OZ-DC specifications in order to reduce their size [5,3].

Since verification is based on the transformation of PEA into TCS, the so-called PEA toolkit [26] provides an automatic computation of the parallel product of PEA and an automatic translation of PEA into TCS, the input representation for both model checkers developed in R1, namely ARMC [29] as well as SLAB [4]. The latter integrates of slicing techniques with abstraction mechanisms. Moreover, counterexample traces produced by ARMC can be automatically traced back to the given high-level CSP-OZ-DC specification and visualized in the Syspect tool [15].
6 Conclusion

We have explained how real-time properties of systems specified in the combination CSP-OZ-DC can be automatically verified using recent advances in semantics, constraint-based model checking, and decision procedures for complex data. The verification is based on the abstraction refinement model checker ARMC that can deal with variables ranging over continuous real-time and infinite data.

A shortcoming of the current version of ARMC is that it cannot exploit the parallel composition that is present in the Phase Event Automata (PEA). Since ARMC expects as input Transition Constraint Systems (TCS) in disjunctive normal form, the parallel product of PEA, which corresponds to the conjunction of TCS, has to be computed before it can be handled by ARMC. For the full benchmark case study this leads to state spaces that are too large to be computed (see Section 4). This shortcoming will be addressed in the future work of the subproject R1.

References


