Temporal-Logic Based Runtime Observer Pairs for System Health Management of Real-Time Systems*

Thomas Reinbacher[1], Kristin Y. Rozier[2], and Johann Schumann[3]

[1] Vienna University of Technology, Austria, treinbacher@ecs.tuwien.ac.at
[2] NASA Ames Research Center, Moffett Field, CA, USA, Kristin.Y.Rozier@nasa.gov
[3] SGT, Inc., NASA Ames, Moffett Field, CA, USA, Johann.M.Schumann@nasa.gov

Abstract. We propose a real-time, Realizable, Responsive, Unobtrusive Unit (rt-R2U2) to meet the emerging needs for System Health Management (SHM) of new safety-critical embedded systems like automated vehicles, Unmanned Aerial Systems (UAS), or small satellites. SHM for these systems must be able to handle unexpected situations and adapt specifications quickly during flight testing between closely-timed consecutive missions, not mid-mission, necessitating fast reconfiguration. They must enable more advanced probabilistic reasoning for diagnostics and prognostics while running aboard limited hardware without affecting the certified on-board software. We define and prove correct translations of two real-time projections of Linear Temporal Logic to two types of efficient observer algorithms to continuously assess the status of the system. A *synchronous* observer yields an instant abstraction of the satisfaction check, whereas an *asynchronous* observer concretizes this abstraction at a later, a priori known, time. By feeding the system's real-time status into a statistical reasoning unit, e.g., based on Bayesian networks, we enable advanced health estimation and diagnosis. We experimentally demonstrate our novel framework on real flight data from NASA's Swift UAS. By on-boarding rt-R2U2 aboard an existing FPGA already built into the standard UAS design and seamlessly intercepting sensor values through read-only observations of the system bus, we avoid system integration problems of software instrumentation or added hardware. The flexibility of our approach with regard to changes in the monitored specification is not due to the reconfigurability offered by FPGAs; it is a benefit of the modularity of our observers and would also be available on non-reconfigurable hardware platforms such as ASICs.

1 Introduction

Autonomous and automated systems, including Unmanned Aerial Systems (UAS), rovers, and satellites, have a large number of components, e.g., sensors, actuators, and software, that must function together reliably at mission time. System Health Management (SHM) [17] can detect, isolate, and diagnose faults and possibly initiate recovery activities on such real-time systems. Effective SHM requires assessing the status of the system with respect to its specifications and estimating system health during mission time. Johnson et al. [17, Ch.1] recently highlighted the need for new, formal-methods based capabilities for modeling complex relationships among different sensor data and reasoning about timing-related requirements; computational expense prevents the current best methods for SHM from meeting operational needs.

* A full version with appendices containing full proofs of correctness for all observer algorithms is available at http://research.kristinrozier.com/TACAS14.html. This work was supported in part by the Austrian Research Agency FFG, grant 825891, and NASA grant NNX08AY50A.

We need a new SHM framework for real-time systems like the Swift [16] electric UAS (see Fig. 1), developed at NASA Ames. SHM for such systems requires:
RESPONSIVENESS: the SHM framework must continuously monitor the system. Deviations from the monitored specifications must be detected within a tight and a priori known time bound, enabling mitigation or rescue measures, e.g., a controlled emergency landing to avoid damage on the ground. Reporting intermediate status and satisfaction of timed requirements as early as possible is required for enabling responsive decision-making.
UNOBTRUSIVENESS: the SHM framework must not alter crucial properties of the system including *functionality* (not change behavior), *certifiability* (avoid re-certification of flight software/hardware), *timing* (not interfere with timing guarantees), and *tolerances* (not violate size, weight, power, or telemetry bandwidth constraints). Utilizing commercial-off-the-shelf (COTS) and previously proven system components is absolutely required to meet today's tight time and budget constraints; adding the SHM framework to the system must not alter these components as changes that require them to be re-certified cancel out the benefits of their use. Our goal is to create the most effective SHM capability with the limitation of read-only access to the data from COTS components.
REALIZABILITY: the SHM framework must be usable in a plug-and-play manner by providing a generic interface to connect to a wide variety of systems. The specification language must be easily understood and expressive enough to encode e.g. temporal relationships and flight rules. The framework must adapt to new specifications without a lengthy re-compilation. We must be able to efficiently monitor different requirements during different mission stages, like takeoff, approach, measurement, and return.

1.1 Related Work

Existing methods for Runtime Verification (RV) [4] assess system status by automatically generating, mainly software-based, observers to check the state of the system against a formal specification. Observations in RV are usually made accessible via software instrumentation [15]; they report only when a specification has passed or failed. Such instrumentation violates our requirements as it may make re-certification of the system onerous, alter the original timing behavior, or increase resource consumption [23]. Also, reporting only the outcomes of specifications violates our responsiveness requirement.

Systems in our applications domain often need to adhere to timing-related rules like: *after receiving the command 'takeoff' reach an altitude of* $600\,ft$ *within five minutes*. These flight rules can be easily expressed in temporal logics; often in some flavor of linear temporal logic (LTL), as studied in [7]. Mainly due to promising complexity results [6, 11], restrictions of LTL to its past-time fragment have most often been used for RV. Though specifications including past time operators may be natural for some other domains [19], flight rules require future-time reasoning. To enable more intuitive specifications, others have studied monitoring of future-time claims; see [22] for a survey and [5, 11, 14, 21, 27, 28] for algorithms and frameworks. Most of these observer algorithms, however, were designed with a software implementation in mind and require a powerful computer. There are many hardware alternatives, e.g. [12], however all either resynthesize monitors from scratch or exclude checking real-time properties [2]. Our unique approach runs the logic synthesis tool once to synthesize as many real-time observer blocks as we can fit on our platform, e.g., FPGA or ASIC; our Sec. 4.1 only interconnects these blocks. Others have proposed using Bayesian inference

techniques [10] to estimate the health of a system. However, modeling timing-related behavior with dynamic Bayesian networks is very complex and quickly renders practical implementations infeasible.

1.2 Approach and Contributions

We propose a new paired-observer SHM framework allowing systems like the Swift UAS to assess their status against a temporal logic specification while enabling advanced health estimation, e.g., via discrete Bayesian networks (BN) [10] based reasoning. This novel combination of two approaches, often seen as orthogonal to each other, enables us to check timing-related aspects with our paired observers while keeping BN health models free of timing information, and thus computationally attractive. Essentially, we can enable better real-time SHM by utilizing paired temporal observers to optimize BN-based decision making. Following our requirements, we call our new SHM framework for real-time systems a rt-R2U2 (real-time, Realizable, Responsive, Unobtrusive Unit).

Our rt-R2U2 synthesizes a pair of observers for a real-time specification φ given in Metric Temporal Logic (MTL) [1] or a specialization of LTL for mission-time bounded characteristics, which we define in Sec. 2. To ensure RESPONSIVENESS of our rt-R2U2, we design two kinds of observer algorithms in Sec. 3 that verify whether φ holds at a discrete time and run them in parallel. *Synchronous* observers have small hardware footprints (max. eleven two-input gates per operator; see Theorem 3 in Sec. 4) and return an instant, three-valued abstraction $\{\textbf{true}, \textbf{false}, \textbf{maybe}\}$) of the satisfaction check of φ with every new tick of the Real Time Clock (RTC) while their *asynchronous* counterparts concretize this abstraction at a later, a priori known time. This unique approach allows us to signal early failure *and acceptance* of every specification whenever possible via the asynchronous observer. Note that previous approaches to runtime monitoring signal only specification failures; signaling *acceptance*, and particularly *early acceptance* is unique to our approach and required for supporting other system components such as prognostics engines or decision making units. Meanwhile, our synchronous observer's three-valued output gives intermediate information that a specification has not yet passed/failed, enabling probabilistic decision making via a Bayesian Network as described in [26].

We implement the rt-R2U2 in hardware as a self-contained unit, which runs externally to the system, to support UNOBTRUSIVENESS; see Sec. 4. Safety-critical embedded systems often use industrial, vehicle bus systems, such as CAN and PCI, interconnecting hardware and software components, see Fig 1. Our rt-R2U2 provides generic read-only interfaces to these bus systems supporting our UNOBTRUSIVENESS requirement and sidestepping instrumentation. Events collected on these interfaces are time stamped by a RTC; progress of time is derived from the observed clock signal, resulting in a discrete time base $\mathbb{N}_0$. Events are then processed by our runtime observer pairs that check whether a specification holds on a sequence of collected events. Other RV approaches for on-the-fly observers exhibit high overhead [13, 20, 24] or use powerful database systems [3], thus, violate our requirements.

To meet our REALIZABILITY requirement, we design an efficient, highly parallel hardware architecture, yet keep it programmable to adapt to changes in the specification. Unlike existing approaches, our observers are designed with an efficient hardware implementation in mind, therefore, avoid recursion and expensive search through memory and aim at maximizing the benefits of the parallel nature of hardware. We synthesize

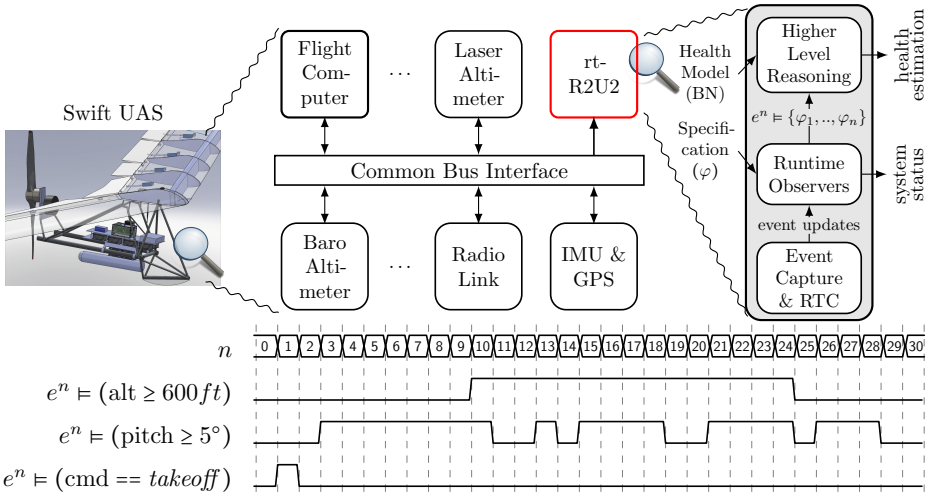


Fig. 1. rt-R2U2: An instance of our SHM framework rt-R2U2 for the NASA Swift UAS. Swift subsystems (top): The laser altimeter maps terrain and determines elevation above ground by measuring the time for a laser pulse to echo back to the UAS. The barometric altimeter determines altitude above sea level via atmospheric pressure. The inertial measurement unit (IMU) reports velocity, orientation (yaw, pitch, and roll), and gravitational forces using accelerometers, gyroscopes, and magnetometers. Running example (bottom): predicates over Swift UAS sensor data on execution e; ranging over the readings of the barometric altimeter, the pitch sensor, and the takeoff command received from the ground station; n is the time stamp as issued by the Real-Time-Clock.

rt-R2U2 *once* and generate a configuration, similar to machine code, to interconnect and configure the static hardware observer blocks of rt-R2U2, adapting to new specifications without running CAD or compilation tools like previous approaches. UAS have very limited bandwidth constraints; transferring a lightweight configuration is preferable to transferring a new image for the whole hardware design. The checks computed by these runtime observers represent the system's status and can be utilized by a higher level reasoner, such as a human operator, Bayesian network, or otherwise, to compute a health estimation, i.e., a conditional probability expressing the belief that a certain subsystem is healthy, given the status of the system. In this paper, we compute these health estimations by adapting the BN-based inference algorithms of [10] in hardware. Our contributions include synthesis and integration of the synchronous/asynchronous observer pairs, a modular hardware implementation, and execution of a proof-of-concept rt-R2U2 running on a self-contained Field Programmable Gate Array (FPGA) (Sec. 5).

2 Real-time projections of LTL

MTL replaces the temporal operators of LTL with operators that respect time bounds [1].

Definition 1 (Discrete-Time MTL). *For atomic proposition $\sigma \in \Sigma$, σ is a formula. Let time bound $J = [t, t']$ with $t, t' \in \mathbb{N}_0$. If φ and ψ are formulas, then so are:*

$$\neg\varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \rightarrow \psi \mid \mathcal{X}\varphi \mid \varphi\, \mathcal{U}_J\, \psi \mid \Box_J\, \varphi \mid \Diamond_J\, \varphi.$$

Time bounds are specified as intervals: for $t, t' \in \mathbb{N}_0$, we write $[t, t']$ for the set $\{i \in \mathbb{N}_0 \mid t \leq i \leq t'\}$. We use the functions $\min, \max, \mathtt{dur}$, to extract the lower time bound (t), the

upper time bound (t'), and the duration ($t' - t$) of J. We define the satisfaction relation of an MTL formula as follows: an execution $e = (s_n)$ for $n \geq 0$ is an infinite sequence of states. For an MTL formula φ, time $n \in \mathbb{N}_0$ and execution e, we define φ *holds at time* n *of execution* e, denoted $e^n \vDash \varphi$, inductively as follows:

$e^n \vDash true$ is **true**, $e^n \vDash \sigma$ iff σ holds in s_n, $e^n \vDash \neg\varphi$ iff $e^n \nvDash \varphi$,
$e^n \vDash \varphi \wedge \psi$ iff $e^n \vDash \varphi$ and $e^n \vDash \psi$, $e^n \vDash \mathcal{X}\,\varphi$ iff $e^{n+1} \vDash \varphi$,
$e^n \vDash \varphi\,\mathcal{U}_J\,\psi$ iff $\exists i (i \geq n) : (i - n \in J \wedge e^i \vDash \psi \wedge \forall j (n \leq j < i) :\ e^j \vDash \varphi)$.

With the dualities $\Diamond_J \varphi \equiv$ **true** $\mathcal{U}_J\ \varphi$ and $\neg \Diamond_J \neg\varphi \equiv \Box_J\ \varphi$ we arrive at two additional operators: $\Box_J\ \varphi$ (φ *is an invariant within the future interval* J) and $\Diamond_J \varphi$ (φ *holds eventually within the future interval* J). In order to efficiently encode specifications in practice, we introduce two special cases of $\Box_J\ \varphi$ and $\Diamond_J \varphi$: $\blacksquare_\tau \varphi \equiv \Box_{[0,\tau]}\ \varphi$ (φ *is an invariant within the next* τ *time units*) and $\blacklozenge_\tau \varphi \equiv \Diamond_{[0,\tau]} \varphi$ (φ *holds eventually within the next* τ *time units*). For example, the flight rule from Sec. 1, "After receiving the takeoff command reach an altitude of $600ft$ within five minutes," is efficiently captured in MTL by $(\text{cmd} == \textit{takeoff}) \rightarrow \blacklozenge_5 (\text{alt} \geq 600ft)$, assuming a time-base of one minute and the atomic propositions $(\text{alt} \geq 600ft)$ and $(\text{cmd} == \textit{takeoff})$ as in Fig. 1.

Systems in our application domain are usually bounded to a certain mission time. For example, the Swift UAS has a limited air-time, depending on the available battery capacity and predefined waypoints. We capitalize on this property to intuitively monitor standard LTL requirements using a mission-time bounded projection of LTL.

Definition 2 (Mission-Time LTL). *For a given LTL formula* ξ *and a mission time* $t_m \in \mathbb{N}_0$*, we denote by* ξ_m *the mission-time bounded equivalent of* ξ*, where* ξ_m *is obtained by replacing every* $\Box\varphi$*,* $\Diamond\varphi$*, and* $\varphi\,\mathcal{U}\,\psi$ *operator in* ξ *by the* $\blacksquare_\tau \varphi$*,* $\blacklozenge_\tau \varphi$*, and* $\varphi\,\mathcal{U}_J\,\psi$ *operators of MTL, where* $J = [0, t_m]$ *and* $\tau = t_m$*.*

Inputs to rt-R2U2 are time-stamped events, collected incrementally from the system.

Definition 3 (Execution Sequence). *An execution sequence for an MTL formula* φ*, denoted by* $\langle T_\varphi \rangle$*, is a sequence of tuples* $T_\varphi = (v, \tau_e)$ *where* $\tau_e \in \mathbb{N}_0$ *is a time stamp and* $v \in \{\textbf{\textit{true}}, \textbf{\textit{false}}, \textbf{\textit{maybe}}\}$ *is a verdict.*

We use a superscript integer to access a particular element in $\langle T_\varphi \rangle$, e.g., $\langle T_\varphi^0 \rangle$ is the first element in execution sequence $\langle T_\varphi \rangle$. We write $T_\varphi.\tau_e$ to access τ_e and $T_\varphi.v$ to access v of such an element. We say T_φ holds if $T_\varphi.v$ is **true** and T_φ does not hold if $T_\varphi.v$ is **false**. For a given execution sequence $\langle T_\varphi \rangle = \langle T_\varphi^0 \rangle, \langle T_\varphi^1 \rangle, \langle T_\varphi^2 \rangle, \langle T_\varphi^3 \rangle, \ldots$, the tuple accessed by $\langle T_\varphi^i \rangle$ corresponds to a section of an execution e as follows: for all times $n \in [\langle T_\varphi^{i-1} \rangle.\tau_e + 1, \langle T_\varphi^i \rangle.\tau_e]$, $e^n \vDash \varphi$ in case $\langle T_\varphi^i \rangle.v$ is **true** and $e^n \nvDash \varphi$ in case $\langle T_\varphi^i \rangle.v$ is **false**. In case $\langle T_\varphi^i \rangle$ is **maybe**, neither $e^n \vDash \varphi$ nor $e^n \nvDash \varphi$ is defined.

In the remainder of this paper, we will frequently refer to execution sequences collected from the Swift UAS as shown in Fig. 1. The predicates shown are atomic propositions over sensor data in our specifications and are sampled with every new time stamp n issued by the RTC. For example, $\langle T_{\text{pitch} \geq 5^\circ} \rangle$ = ((**false**, 0), (**false**, 1), (**false**, 2), (**true**, 3), ..., (**true**, 17), (**true**, 18)) describes $e^n \vDash (\text{pitch} \geq 5^\circ)$ sampled over $n \in [0, 18]$ and $\langle T_{\text{pitch} \geq 5^\circ} \rangle$ holds 19 elements.

3 Asynchronous and Synchronous Observers

The problem of monitoring a real-time specification has been studied extensively in the past; see [8, 22] for an overview. Solutions include: (a) translating the temporal formula into a finite-state automaton that accepts all the models of the specification [11,12,14,28], (b) restricting MTL to its *safety* fragment and waiting until the operators' time bounds have elapsed to decide the truth value afterwards [5, 21], and (c) restricting LTL to its past-time fragment [6, 11, 24]. Compiling new observers to automata as in (a) requires re-running the logic synthesis tool to yield a new hardware observer, in automaton or autogenerated VHDL code format as described in [12], which may take dozens of minutes to complete, violating the REALIZABILITY requirement. Observers generated by (b) are in conflict with the RESPONSIVENESS requirement and (c) do not natively support flight rules. Our observers provide UNOBTRUSIVENESS via a self-contained hardware implementation. To enable such an implementation, our design needs to refrain from dynamic memory, linked lists, and recursion – commonly used in existing software-based observers, however, not natively available in hardware.

Our two types of runtime observers differ in the times when new outputs are generated and in the resource footprints required to implement them. A *synchronous* (time-triggered) observer is trimmed towards a minimalistic hardware footprint and computes a three-valued abstraction of the satisfaction check for the specification with each tick of the RTC, without considering events happening after the current time. An *asynchronous* (event-triggered) observer concretizes this abstraction at a later, a priori known, time and makes use of synchronization queues to take events into account that occur after the current time.[1] Our novel parallel composition of these two observers updates the status of the system at every tick of the RTC, yielding great responsiveness. An inconclusive answer when we can't yet know **true**/**false** is still beneficial as the higher-level reasoning part of our rt-R2U2 supports reasoning with inconclusive inputs. This allows us to derive an intermediate estimation of system health with the option to initiate fault mitigation actions even without explicitly knowing all inputs. If exact reasoning is required, we can re-evaluate system health when the *asynchronous* observer provides exact answers.

In the remainder of this section, we discuss[2] both *asynchronous* and *synchronous* observers for the operators $\neg\,\varphi$, $\varphi \wedge \psi$, $\blacksquare_{\tau}\,\varphi$, $\square_{J}\,\varphi$, and $\varphi\,\mathcal{U}_{J}\,\psi$. Informally, an MTL observer is an algorithm that takes execution sequences as input and produces another execution sequence as output. For a given unary operator $\bullet$, we say that an observer algorithm implements $e^n \vDash \bullet\,\varphi$, iff for all execution sequences $\langle T_{\varphi} \rangle$ as input, it produces an execution sequence as output that evaluates $e^n \vDash \bullet\,\varphi$ (analogous for binary operators).

3.1 Asynchronous Observers

The main characteristic of our asynchronous observers is that they are *evaluated with every new input tuple* and that for every generated output tuple T we have that $T.v \in \{\textbf{true}, \textbf{false}\}$ and $T.\tau_e \in [0, n]$. Since verdicts are exact evaluations of a future-time specification φ for each clock tick they may resolve φ for clock ticks prior to the current time n if the information required for this resolution was not available until n.

[1] Similar terms have been used by others [9] to refer to monitoring with pairs of observers that do not update with the RTC, incur delays dangerous to a UAS, and require system interaction that violates our requirements (Sec. 1).

[2] Proofs of correctness for every observer algorithm appear in the Appendix.

Our observers distinguish two types of transitions of the signals described by execution sequences. We say transition $\_\!\!\sqcap$ of execution sequence $\langle T_\varphi \rangle$ occurs at time $n = \langle T_\varphi^i \rangle.\tau_e + 1$ iff $(\langle T_\varphi^i \rangle.v \oplus \langle T_\varphi^{i+1} \rangle.v) \wedge \langle T_\varphi^{i+1} \rangle.v$ holds. Similarly, we say transition $\sqcap\!\!\_$ of execution sequence $\langle T_\varphi \rangle$ occurs at time $n = \langle T_\varphi^i \rangle.\tau_e + 1$ iff $(\langle T_\varphi^i \rangle.v \oplus \langle T_\varphi^{i+1} \rangle.v) \wedge \langle T_\varphi^i \rangle.v$ holds ($\oplus$ denotes the Boolean exclusive-or). For example, transitions $\_\!\!\sqcap$ and $\sqcap\!\!\_$ of $\langle T_{\text{pitch}\geq 5^\circ} \rangle$ in Fig. 1 occur at times 3 and 11, respectively.

Negation ($\neg\, \varphi$) The observer for $\neg\, \varphi$, as stated in Alg. 1, is straightforward: for every input T_φ we negate the truth value of $T_\varphi.v$. The observer generates $(\ldots, (\textbf{true}, 2), (\textbf{false}, 3), \ldots)$.

Invariant within the Next τ Time Stamps ($\blacksquare_\tau\, \varphi$) An observer for $\blacksquare_\tau\, \varphi$ requires registers $m_{\uparrow\varphi}$ and m_{τ_s} with domain $\mathbb{N}_0$: $m_{\uparrow\varphi}$ holds the time stamp of the latest $\_\!\!\sqcap$ transition of $\langle T_\varphi \rangle$ whereas m_{τ_s} holds the start time of the next tuple in $\langle T_\varphi \rangle$. For the observer in Alg. 2, the check $m \leq (T_\varphi.\tau_e - \tau)$ in line 8 tests whether φ held for at least the previous τ time stamps. To illustrate the algorithm, consider an observer for $\blacksquare_5$ (pitch $\geq 5^\circ$) and the execution in Fig. 1. At time $n = 0$, we have $m_{\uparrow\varphi} = 0$ and since $\langle T^0_{\text{pitch}\geq 5^\circ} \rangle$ does not hold the output is $(\textbf{false}, 0)$. Similarly, the outputs for $n \in [1, 2]$ are $(\textbf{false}, 1)$ and $(\textbf{false}, 2)$. At time $n = 3$, a $\_\!\!\sqcap$ transition of $\langle T_{\text{pitch}\geq 5^\circ} \rangle$ occurs, thus $m_{\uparrow\varphi} = 3$. Since the check in line 8 does not hold, the algorithm does not generate a new output, i.e., returns $(\_, \_)$ designating output is delayed until a later time, which repeats at times $n \in [4, 7]$. At $n = 8$, the check in line 8 holds and the algorithm returns $(\textbf{true}, 3)$. Likewise, the outputs for $n \in [9, 10]$ are $(\textbf{true}, 4)$ and $(\textbf{true}, 5)$. At $n = 11$, $\langle T^{11}_{\text{pitch}\geq 5^\circ} \rangle$ does not hold and the algorithm outputs $(\textbf{false}, 11)$. We note the ability of the observer to *re-synchronize* its output with respect to its inputs and the RTC. For $n \in [8, 10]$, outputs are given for a time prior to n, however, at $n = 11$ the observer re-synchronizes: the output $(\textbf{false}, 11)$ signifies that $e^n \not\models \blacksquare_5$ (pitch $\geq 5^\circ$) for $n \in [6, 11]$. By the equivalence $\blacklozenge_\tau\, \varphi \equiv \neg \blacksquare_\tau \neg \varphi$, we immediately arrive at an observer for $\blacklozenge_\tau\, \varphi$ from Alg. 2 by negating both the input and the output tuple.

Invariant within Future Interval ($\square_J\, \varphi$) The observer for $\square_J\, \varphi$, as stated in Alg. 4, builds on an observer for $\blacksquare_\tau\, \varphi$ and makes use of the equivalence $\blacksquare_\tau \varphi \equiv \square_{[0,\tau]}\, \varphi$. Intuitively, the observer for $\blacksquare_\tau\, \varphi$ returns true iff φ holds for at least the next τ time units. We can thus construct an observer for $\square_J\, \varphi$ by reusing the algorithm for $\blacksquare_\tau\, \varphi$, assigning $\tau = \texttt{dur}(J)$ and shifting the obtained output by $\min(J)$ time stamps into the past. From the equivalence $\diamondsuit_J \varphi \equiv \neg\, \square_J\, \neg \varphi$, we can immediately derive an observer for $\diamondsuit_J\, \varphi$ from the observer for $\square_J\, \varphi$. To illustrate the algorithm, consider an observer for $\square_{5,10}$ (alt $\geq 600ft$) over the execution in Fig. 1. For $n \in [0, 4]$ the algorithm returns $(\_, \_)$, since $(\langle T^{0\ldots4}_{\text{alt}\geq 600ft} \rangle.\tau_e - 5) \geq 0$ (line 3 of Alg. 4) does not hold. At $n = 5$ the underlying observer for $\blacksquare_5$ (alt $\geq 600ft$) returns $(\textbf{false}, 5)$, which is transformed (by line 4) into the output $(\textbf{false}, 0)$. For similar arguments, the outputs for $n \in [6, 9]$ are $(\textbf{false}, 1)$, $(\textbf{false}, 2)$, $(\textbf{false}, 3)$, and $(\textbf{false}, 4)$. At $n \in [10, 14]$, the observer for $\blacksquare_5$ (alt $\geq 600ft$) returns $(\_, \_)$. At $n = 15$, $\blacksquare_5$ (alt $\geq 600ft$) yields $(\textbf{true}, 10)$, which is transformed (by line 4) into the output is $(\textbf{true}, 5)$. Note also that $\mathcal{X}\varphi \equiv \square_{[1,1]}\, \varphi$.

The remaining observers for the binary operators $\varphi \wedge \psi$ and $\varphi\, \mathcal{U}_J\, \psi$ take tuples (T_φ, T_ψ) as inputs, where T_φ is from $\langle T_\varphi \rangle$ and T_ψ is from $\langle T_\psi \rangle$. Since $\langle T_\varphi \rangle$ and $\langle T_\psi \rangle$ are execution sequences produced by two different observers, the two elements of the

input tuple (T_φ, T_ψ) are not necessarily generated at the same time. Our observers for binary MTL operators thus use two FIFO-organized *synchronization queues* to buffer parts of $\langle T_\varphi \rangle$ and $\langle T_\psi \rangle$, respectively. For a synchronization queue q we denote by $q = ()$ its emptiness and by $|q|$ its size.

Algorithm 1 Observer for $\neg\,\varphi$.

At each new input T_φ:
$T_\xi \leftarrow (\neg\, T_\varphi.v, T_\varphi.\tau_e)$
return T_ξ

Algorithm 2 Observer for $\blacksquare_\tau\ \varphi$. Initially, $m_{\uparrow\varphi} = m_{\tau_s} = 0$.

At each new input T_φ:
$T_\xi \leftarrow T_\varphi$
if $\_\!\!\lceil\!\!\lceil$ transition of T_ξ occurs **then**
  $m_{\uparrow\varphi} \leftarrow m_{\tau_s}$
end if
$m_{\tau_s} \leftarrow T_\varphi.\tau_e + 1$
if T_ξ holds **then**
  if $m_{\uparrow\varphi} \le (T_\xi.\tau_e - \tau)$ holds **then**
    $T_\xi.\tau_e \leftarrow T_\xi.\tau_e - \tau$
  else
    $T_\xi \leftarrow (\_, \_)$
  end if
end if
return T_ξ

Algorithm 3 Observer for $\varphi \wedge \psi$.

At each new input (T_φ, T_ψ):
if T_φ holds and T_ψ holds and $q_\varphi \neq ()$ holds and $q_\psi \neq ()$ holds **then**
  $T_\xi \leftarrow (\textbf{true}, \min(T_\varphi.\tau_e, T_\psi.\tau_e))$
else if $\neg T_\varphi$ holds and $\neg T_\psi$ holds and $q_\varphi \neq ()$ holds and $q_\psi \neq ()$ holds **then**
  $T_\xi \leftarrow (\textbf{false}, \max(T_\varphi.\tau_e, T_\psi.\tau_e))$
else if $\neg T_\varphi$ holds and $q_\varphi \neq ()$ holds **then**
  $T_\xi \leftarrow (\textbf{false}, T_\varphi.\tau_e)$
else if $\neg T_\psi$ holds and $q_\psi \neq ()$ holds **then**
  $T_\xi \leftarrow (\textbf{false}, T_\psi.\tau_e)$
else
  $T_\xi \leftarrow (\_, \_)$
end if
dequeue$(q_\varphi, q_\psi, T_\xi.\tau_e)$
return T_ξ

Algorithm 4 Observer for $\square_J\ \varphi$.

At each new input T_φ:
$T_\xi \leftarrow \blacksquare_{\mathtt{dur}(J)}\ T_\varphi$
if $(T_\xi.\tau_e - \min(J) \ge 0)$ **then**
  $T_\xi.\tau_e \leftarrow T_\xi.\tau_e - \min(J)$
else
  $T_\xi \leftarrow (\_, \_)$
end if
return T_ξ

Algorithm 5 Observer for $\varphi\,\mathcal{U}_J\,\psi$. Initially, $m_{pre} = m_{\uparrow\varphi} = 0$, $m_{\downarrow\varphi} = -\infty$, and $p =$ **false**.

At each new input (T_φ, T_ψ) in lockstep mode:
if $\_\!\!\lceil\!\!\lceil$ transition of T_φ occurs **then**
  $m_{\uparrow\varphi} \leftarrow \tau_e - 1$
  $m_{pre} \leftarrow -\infty$
end if
if $\lceil\!\!\rceil\!\!\_$ transition of T_φ occurs and T_ψ holds **then**
  $T_\varphi.v, p \leftarrow \textbf{true}, \textbf{true}$
  $m_{\downarrow\varphi} \leftarrow \tau_e$
end if
if T_φ holds **then**
  if T_ψ holds **then**
    if $(m_{\uparrow\varphi} + \min(J) < \tau_e)$ holds **then**
      $m_{pre} \leftarrow \tau_e$
      return $(\textbf{true}, \tau_e - \min(J))$
    else if p holds **then**
      return $(\textbf{false}, m_{\downarrow\varphi})$
    end if
  else if $(m_{pre} + \mathtt{dur}(J) \le \tau_e)$ holds **then**
    return $(\textbf{false}, \max(m_{\uparrow\varphi}, \tau_e - \max(J)))$
  end if
else
  $p \leftarrow \textbf{false}$
  if $(\min(J) = 0)$ holds **then**
    return $(T_\psi.v, \tau_e)$
  end if
  return $(\textbf{false}, \tau_e)$
end if
return $(\_, \_)$

Conjunction ($\varphi \wedge \psi$) The observer for $\varphi \wedge \psi$, as stated in Alg. 3, reads inputs (T_φ, T_ψ) from two synchronization queues, q_φ and q_ψ. Intuitively, the algorithm follows the rules for conjunction in Boolean logic with additional emptiness checks on q_φ and q_ψ. The procedure **dequeue**$(q_\varphi, q_\psi, T_\xi.\tau_e)$ drops all entries T_φ in q_φ for which the following holds: $T_\varphi.\tau_e \le T_\xi.\tau_e$ (analogous for q_ψ). To illustrate the algorithm, consider an observer for $\blacksquare_5\ (\text{alt} \ge 600ft) \wedge (\text{pitch} \ge 5°)$ and the execution in Fig. 1.

For $n \in [0,9]$ the two observers for the involved subformulas immediately output (**false**, n). For $n \in [10,14]$, the observer for $\blacksquare_5$ (alt $\geq 600ft$) returns $(\_,\_)$, while in the meantime, the atomic proposition (pitch $\geq 5°$) toggles its truth value several times, i.e., (**true**, 10), (**false**, 11), (**false**, 12), (**true**, 13), (**false**, 14). These tuples need to be buffered in queue $q_{\text{pitch}\geq 5°}$ until the observer for $\blacksquare_5$ (alt $\geq 600ft$) generates its next output, i.e., (**true**, 10) at $n = 15$. We apply the function **aggregate**($\langle T_\varphi \rangle$), which repeatedly replaces two consecutive elements $\langle T_\varphi^i \rangle, \langle T_\varphi^{i+1} \rangle$ in $\langle T_\varphi \rangle$ by $\langle T_\varphi^{i+1} \rangle$ iff $\langle T_\varphi^i \rangle.v = \langle T_\varphi^{i+1} \rangle.v$, to the content of $q_{\text{pitch}\geq 5°}$ once every time an element is added to $q_{\text{pitch}\geq 5°}$. Therefore, at $n = 15$: $q_{\text{pitch}\geq 5°}$ = ((**true**, 10), (**false**, 12), (**true**, 13), (**false**, 14), (**true**, 15)) and $q_{\blacksquare_5\ (\text{alt}\geq 600ft)}$ = ((**true**, 10)). The observer returns (**true**, 10) (line 3) and **dequeue**($q_\varphi, q_\psi, 10$) yields: $q_{\text{pitch}\geq 5°}$ = ((**false**, 12), (**true**, 13), (**false**, 14), (**true**, 15)) and $q_{\blacksquare_5\ (\text{alt}\geq 600ft)}$ = ().

Until within Future Interval ($\varphi\, \mathcal{U}_J\, \psi$) The observer for $\varphi\, \mathcal{U}_J\, \psi$, as stated in Alg. 5, reads inputs (T_φ, T_ψ) from two synchronization queues and makes use of a Boolean flag p and three registers $m_{\uparrow\varphi}$, $m_{\downarrow\varphi}$, and m_{pre} with domain $\mathbb{N}_0 \cup \{-\infty\}$: $m_{\uparrow\varphi}$ ($m_{\downarrow\varphi}$) holds the time stamp of the latest ⌐¬ transition (¬⌐ transition) of $\langle T_\varphi \rangle$ and m_{pre} holds the latest time stamp where the observer detected $\varphi\, \mathcal{U}_J\, \psi$ to hold. Input tuples (T_φ, T_ψ) for the observer are read from synchronization queues in a *lockstep* mode: (T_φ, T_ψ) is split into (T'_φ, T'_ψ), where $T'_\varphi.\tau_e = T'_\psi.\tau_e$ and the time stamp $T''_\varphi.\tau_e$ of the next tuple (T''_φ, T''_ψ) is $T'_\varphi.\tau_e + 1$. This ensures that the observer outputs only a single tuple at each run and avoids output buffers, which would account for additional hardware resources (see correctness proof in the Appendix for a discussion). Intuitively, if T_φ does not hold (lines 22-26) the observer is synchronous to its input and immediately outputs (**false**, $T_\varphi.\tau_e$). If T_φ holds (lines 11-20) the time stamp n' of the output tuple is not necessarily *synchronous* to the time stamp $T_\varphi.\tau_e$ of the input anymore, however, bounded by $(T_\varphi.\tau_e - \max(J)) \leq n' \leq T_\varphi.\tau_e$ (see Lemma "*unrolling*" in the Appendix). To illustrate the algorithm, consider an observer for (pitch $\geq 5°$) $\mathcal{U}_{[5,10]}$ (alt $\geq 600ft$) over the execution in Fig. 1. At time $n = 0$, we have $m_{pre} = 0$, $m_{\uparrow\varphi} = 0$, and $m_{\downarrow\varphi} = -\infty$ and since $\langle T^0_{\text{pitch}\geq 5°} \rangle$ does not hold, the observer outputs (**false**, 0) in line 26. The outputs for $n \in [1,2]$ are (**false**, 1) and (**false**, 2). At time $n = 3$, a ⌐¬ transition of $\langle T_{\text{pitch}\geq 5°} \rangle$ occurs, thus we assign $m_{\uparrow\varphi} = 2$ and $m_{pre} = -\infty$ (lines 3 and 4). Since $\langle T^3_{\text{pitch}\geq 5°} \rangle$ holds and $\langle T^3_{\text{alt}\geq 600ft} \rangle$ does not hold, the predicate in line 18 is evaluated, which holds and the algorithm returns ⟨**false**, $\max(2, 3-10)$⟩ = (**false**, 2). Thus, the observer does not yield a new output in this case, which repeats for times $n \in [4,9]$. At time $n = 10$, a ⌐¬ transition of $\langle T_{\text{alt}\geq 600ft} \rangle$ occurs and the predicate in line 12 is evaluated. Since $(2 + 5) < 10$ holds, the algorithm returns (**true**, 5), revealing that $e^n \vDash$ (pitch $\geq 5°$) $\mathcal{U}_{[5,10]}$ (alt $\geq 600ft$) for $n \in [3,5]$. At time $n = 11$, a ¬⌐ transition of $\langle T_{\text{pitch}\geq 5°} \rangle$ occurs and since $\langle T^{11}_{\text{alt}\geq 600ft} \rangle$ holds, p and the truth value of the current input $\langle T^{11}_{\text{pitch}\geq 5°} \rangle.v$ are set **true** and $m_{\downarrow\varphi} = 11$. Again, line 12 is evaluated and the algorithm returns (**true**, 6). At time $n = 12$, since $\langle T^{12}_{\text{pitch}\geq 5°} \rangle$ does not hold, we clear p in line 22 and the algorithm returns (**false**, 12) in line 26, i.e., $e^n \nvDash$ (pitch $\geq 5°$) $\mathcal{U}_{[5,10]}$ (alt $\geq 600ft$) for $n \in [7,12]$. At time $n = 13$, a ⌐¬ transition of $\langle T_{\text{pitch}\geq 5°} \rangle$ occurs, thus $m_{\uparrow\varphi} = 12$ and $m_{pre} = -\infty$. The predicates in line 12 and 15 do not hold, the algorithm returns no new output in line 28. At time $n = 14$, a ¬⌐ transition of $\langle T_{\text{pitch}\geq 5°} \rangle$ occurs, thus p and $\langle T^{14}_{\text{pitch}\geq 5°} \rangle.v$ are set **true** and

$m_{\downarrow\varphi} = 14$. The predicate in line 15 holds, and the algorithm outputs $(\textbf{false}, 14)$, revealing that $e^n \not\models (\text{pitch} \geq 5^\circ)\ \mathcal{U}_{[5,10]}\ (\text{alt} \geq 600ft)$ for $n \in [13, 14]$.

3.2 Synchronous Observers

The main characteristic of our synchronous observers is that they are evaluated at every tick of the RTC and that their output tuples T are guaranteed to be synchronous to the current time stamp n. Thus, for each time n, a synchronous observer outputs a tuple T with $T.\tau_e = n$. This eliminates the need for synchronization queues. Inputs and outputs of these observers are execution sequences with three-valued verdicts. The underlying abstraction is given by $\widehat{\textbf{eval}} : \circledast \to \{\textbf{true}, \textbf{false}, \textbf{maybe}\}$, where $\circledast \in \{\neg\varphi, \varphi \wedge \psi, \blacksquare_\tau\ \varphi, \square_J\ \varphi, \varphi\ \mathcal{U}_J\ \psi\}$. The implementation of $\widehat{\textbf{eval}}\ (\neg\varphi)$ and $\widehat{\textbf{eval}}\ (\varphi \wedge \psi)$ follows the rules for Kleene logic [18]. For the remaining operators we define the verdict $T_\xi.v$ of the output tuple $(T_\xi.v, n)$, generated for inputs $(T_\varphi.v, n)$ (respectively $(T_\psi.v, n)$ for $\varphi\ \mathcal{U}_J\ \psi$), as:

$$\widehat{\textbf{eval}}\ (\blacksquare_\tau\ \varphi) = \begin{cases} \textbf{true} & \text{if } T_\varphi.v \text{ holds and } \tau = 0, \\ \textbf{false} & \text{if } T_\varphi.v \text{ does not hold,} \\ \textbf{maybe} & \text{otherwise.} \end{cases}$$

$$\widehat{\textbf{eval}}\ (\square_J\ \varphi) = \textbf{maybe}.$$

$$\widehat{\textbf{eval}}\ (\varphi\ \mathcal{U}_J\ \psi) = \begin{cases} \textbf{true} & \text{if } \begin{array}{l} T_\varphi.v \text{ and } T_\psi.v \text{ holds} \\ \text{and } \min(J) = 0, \end{array} \\ \textbf{false} & \text{if } T_\varphi.v \text{ does not hold,} \\ \textbf{maybe} & \text{otherwise.} \end{cases}$$

To illustrate our synchronous observer algorithms, consider the previously discussed formula $\blacksquare_5\ (\text{alt} \geq 600ft) \wedge (\text{pitch} \geq 5^\circ)$, which we want to evaluate using the synchronous observer:

$$\xi = \widehat{\textbf{eval}}\ (\widehat{\textbf{eval}}\ (\blacksquare_5\ (\text{alt} \geq 600ft)) \wedge (\text{pitch} \geq 5^\circ))$$

For $n \in [0, 9]$, as in the case of the *asynchronous* observer, we can immediately output $(\textbf{false}, n)$. At $n = 10$, $\widehat{\textbf{eval}}\ (\blacksquare_5\ (\text{alt} \geq 600ft))$ yields $(\textbf{maybe}, n)$, thus, the observer is inconclusive about the truth value of $e^{10} \models \xi$. At $n \in [11, 12]$ since $(\text{pitch} \geq 5^\circ)$ does not hold, the outputs are $(\textbf{false}, n)$. For analogous arguments, the output at $n = 13$ is $(\textbf{maybe}, 13)$, at $n = 14$ $(\textbf{false}, 14)$, and at $n = 15$ $(\textbf{maybe}, 15)$. In this way, at times $n \in \{11, 12, 14\}$ the synchronous observer completes early evaluation of ξ, producing output that would, without the abstraction, be guaranteed by the exact asynchronous observer with a delay of 5 time units, i.e., at times $n \in \{16, 17, 19\}$.

4 Mapping Observers into Efficient Hardware

We introduce a mapping of the observer pairs into efficient hardware blocks and a synthesis procedure to generate a configuration for these blocks from an arbitrary MTL specification. This configuration is loaded into the control unit of our rt-R2U2, where it changes the interconnections between a pool of (static) hardware observer blocks and assigns memory regions for synchronization queues. This approach enables us to quickly change the monitored specification (within resource limitations) without re-compiling the rt-R2U2's hardware, supporting our REALIZABILITY requirement.

Asynchronous observers require arithmetic operations on time stamps. Registers and flags as required by the observer algorithm are mapped to circuits that can store

information, such as flip-flops. For the synchronization queues we turn to block RAMs (abundant on FPGAs), organized as ring buffers. Time stamps are internally stored in registers of width $w = \lceil log_2(n) \rceil + 2$, to indicate $-\infty$ and to allow overflows when performing arithmetical operations on time stamps. Subtraction and relational operators as required by the observer for $\blacksquare_\tau \varphi$ (Fig. 2) can be built around adders. For example, the check in line 8 of Alg. 2 is implemented using two w-bit wide adders: one for $q = T_\varphi.\tau_e - \tau$ and one to decide whether $m_{\uparrow\varphi} \geq q$. A third adder runs in parallel and assigns a new value to m_{τ_s} (line 6 of Alg. 2). Detecting a $\_\!\!\sqcap$ transition on $\langle T_\varphi \rangle$ maps to an XOR gate and an AND gate, implementing the circuit $(T_\varphi^{i-1}.v \oplus T_\varphi^i.v) \wedge T_\varphi^i.v$, where $T_\varphi^{i-1}.v$ is the truth value of the previous input, stored in a flip-flop. The multiplexer either writes a new output or sets a flag to indicate $(\lrcorner, \lrcorner)$.

Synchronous observers do not require calculations on time stamps and directly map to basic digital logic gates. Fig. 2 shows a circuit representing an $\widehat{\textbf{eval}}\,(\blacksquare_\tau \varphi)$ observer that accounts for one two-input AND gate, one two-input OR gate, and two Inverter gates. Inputs (i_1, i_2) and outputs (y_1, y_2) are encoded (to project the three-valued logic into Boolean logic) such as: **true** $(0, 0)$, **false** $(0, 1)$, and **maybe** $(1, 0)$. Input j is set if $\tau_e = 0$ and cleared otherwise.

4.1 Synthesizing a Configuration for the rt-R2U2

The synthesis procedure to translate an MTL specification ξ into a configuration such that the rt-R2U2 instantiates observers for both ξ and $\widehat{\textbf{eval}}\,(\xi)$, works as follows:

- Preprocessing. By the equivalences given in Sect. 2 rewrite ξ to ξ', such that operators in ξ' are from $\{\neg\varphi, \varphi \wedge \psi, \blacksquare_\tau \varphi, \square_J \varphi, \varphi\, \mathcal{U}_J\, \psi\}$ (**SA1**).
- Parsing. Parse ξ' to obtain an Abstract Syntax Tree (AST), denoted by AST(ξ'). The leaves of this tree are the atomic propositions Σ of ξ' (**SA2**).
- Allocating observers. For all nodes q in AST(ξ') allocate both the corresponding synchronous and the asynchronous hardware observer block (**SA3**).
- Adding synchronization queues. $\forall q \in$ AST(ξ'): If q is of type $\varphi \wedge \psi$ or $\varphi\, \mathcal{U}_J\, \psi$ add queues q_φ and q_ψ to the inputs of the respective asynchronous observer (**MA1**).
- Interconnect and dimensioning. Connect observers and queues according to AST(ξ'). Execute Alg. 6 (**MA2**).

Let $\{\sigma_1, \sigma_2, \sigma_3\} \in \Sigma$ and $\xi = \sigma_1 \rightarrow (\blacklozenge_{10}\ (\sigma_2) \vee \blacklozenge_{100}(\sigma_3))$ be an MTL formula we want to synthesize a configuration for. SA1 yields $\xi' = \neg(\sigma_1 \wedge \neg(\neg\blacksquare_{10}\ (\neg\sigma_2)) \wedge \neg(\neg\blacksquare_{100}\ (\neg\sigma_3)))$ which simplifies to $\xi = \neg(\sigma_1 \wedge \blacksquare_{10}\ (\neg\sigma_2) \wedge \blacksquare_{100}\ (\neg\sigma_3))$. SA2 yields AST($\xi'$). SA3 instantiates two $\varphi \wedge \psi$, three $\neg\varphi$, one $\blacksquare_{10}\ T_\varphi$ and one $\blacksquare_{100}\ T_\varphi$ observers, both synchronous and asynchronous. MA1, introduces queues $q_{\sigma_1}, q_{\xi_2}, q_{\xi_3}, q_{\xi_4}$ and MA2 interconnects observers and queues and assigns $|q_{\sigma_1}| = 100$, $|q_{\xi_2}| = 100$, $|q_{\xi_3}| = 10$, and $|q_{\xi_4}| = 0$, see Fig. 2.

4.2 Circuit Size and Depth Complexity Results

Having discussed how to determine the size of the synchronization queues for our asynchronous MTL observers, we are now in the position to prove space and time complexity bounds.

Algorithm 6 Assigning synchronization queue sizes for AST(ξ'). Let S be a set of nodes; Initially: $w = 0$, add all Σ nodes of AST(ξ') to S; The function wcd : $\circledast \to \mathbb{N}_0$ calculates the *worst-case-delay* an asynchronous observer may introduce by: $\text{wcd}(\neg\,\varphi) = \text{wcd}(\varphi \wedge \psi) = 0$, $\text{wcd}(\blacksquare_\tau\,\varphi) = \tau$, $\text{wcd}(\square_J\,\varphi) = \text{wcd}(\varphi\,\mathcal{U}_J\,\psi) = \max(J)$.

while S is not empty **do**
  $s, w \leftarrow$ get next node from S, 0
  if s is type $\varphi\,\mathcal{U}_J\,\psi$ or $\varphi \wedge \psi$ **then**
    $w \leftarrow \max(|q_\varphi|, |q_\psi|) + \text{wcd}(s)$
  end if
  while s is not a synchronization queue **do**
    $s, w \leftarrow$ get predecessor of s in AST(ξ'), $w + \text{wcd}(s)$
  end while
  Set $|q| = w$; (q is opposite synchronization queue of s)
  Add all $\varphi\,\mathcal{U}_J\,\psi$ and $\varphi \wedge \psi$ nodes that have unassigned synchronization queue sizes to S
end while

Theorem 1 (Space Complexity of Asynchronous Observers). *The respective asynchronous observer for a given MTL specification φ has a space complexity, in terms of memory bits, bounded by $(2 + \lceil \log_2(n) \rceil) \cdot (2 \cdot m \cdot p)$, where m is the number of binary observers (i.e., $\varphi \wedge \psi$ or $\varphi\mathcal{U}_J\,\psi$) in φ, p is the worst-case delay of a single predecessor chain in* AST(φ)*, and $n \in \mathbb{N}_0$ is the time stamp it is executed.*

Theorem 2 (Time Complexity of Asynchronous Observers). *The respective asynchronous observer for a given MTL specification φ has an asymptotic time complexity of $\mathcal{O}\Big(\log_2 \log_2 \max(p, n) \cdot d\Big)$, where p is the maximum worst-case-delay of any observer in* AST(φ)*, d the depth of AST(φ), and $n \in \mathbb{N}_0$ the time stamp it is executed.*

For our synchronous observers, we prove upper bounds in terms of two-input gates on the size of resulting circuits. Actual implementations may yield significant better results on circuit size, depending on the performance of the logic synthesis tool.

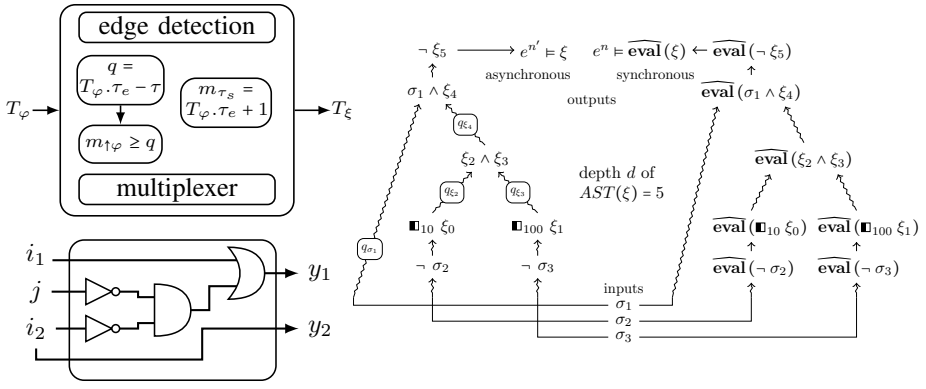


Fig. 2. Left: hardware implementations for $\blacksquare_\tau\,\varphi$ (top) and $\widehat{\textbf{eval}}\,(\blacksquare_\tau\,\varphi)$ (bottom). Right: subformulas of AST(ξ), observers, and queues synthesized for ξ. Mapping the observers to hardware yields two levels of parallelism: (i) asynchronous (left) and the synchronous observers (right) run in parallel and (ii) observers for subformulas run in parallel, e.g., $\blacksquare_{10}\,\xi_0$ and $\blacksquare_{100}\,\xi_1$.

Theorem 3 (Circuit-Size Complexity of Synchronous Observers). *For a given MTL formula φ, the circuit to monitor $\widehat{\textbf{eval}}\,(\varphi)$ has a circuit-size complexity bounded by $11 \cdot m$, where m is the number of observers in AST(φ).*

Theorem 4 (Circuit-Depth Complexity of Synchronous Observers). *For a given MTL formula φ, the circuit to monitor $\widehat{\textbf{eval}}\,(\varphi)$ has a circuit-depth complexity of $4 \cdot d$.*

5 Applying the rt-R2U2 to NASA's Swift UAS

We implemented our rt-R2U2 as a register-transfer-level VHDL hardware design, which we simulated in MENTOR GRAPHICS MODELSIM and synthesized for different FPGAs using the industrial logic synthesis tool ALTERA QUARTUS II.[3] With our rt-R2U2, we analyzed raw flight data from NASA's Swift UAS collected during test flights. The higher-level reasoning is performed by a *health model*, modeled as a Bayesian network (BN) where the nodes correspond to discrete random variables. Fig. 3 shows the relevant excerpt for reasoning about altitude. Directed edges encode conditional dependencies between variables, e.g., the sensor reading S_L depends on the health of the laser altimeter sensor H_L. Conditional probability tables at each node define the local dependencies. During health estimation, verdicts computed by our observer algorithms are provided as virtual sensor values to the observable nodes S_L, S_B, S_S; e.g., the laser altimeter measuring an altitude increase would result in setting S_L to state *inc*. Then, the posteriors of the multivariate probability distribution encoded in the BN are calculated [10]; for details of modeling and reasoning see [25].

Our temporal specifications are evaluated by our runtime observers and describe flight rules (φ_1, φ_2) and virtual sensors:

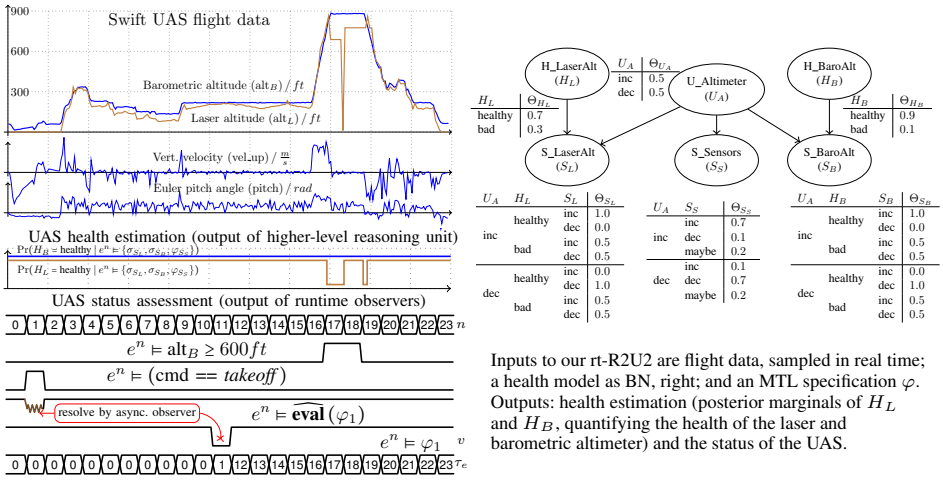
$$\varphi_1 = (\text{cmd} == \textit{takeoff}) \rightarrow \blacklozenge_{10}\ (\text{alt}_B \geq 600ft)$$
$$\varphi_2 = (\text{cmd} == \textit{takeoff}) \rightarrow \blacklozenge_{*}\ (\text{cmd} == \textit{land})$$

φ_1 encodes our running example flight rule; φ_2 is a mission-bounded LTL property requiring that the command *land* is received after *takeoff*, within the projected mission time, indicated by $*$. Fig. 3 shows the execution sequences produced by both the asynchronous ($e^n \vDash \varphi_1$) and the synchronous ($e^n \vDash \widehat{\textbf{eval}}\,(\varphi_1)$) observers for flight rule φ_1. To keep the presentation accessible we scaled the timeline to just 24 time stamps; the actual implementation uses a resolution of 2^{32} time stamps. The synchronous observer is able to prove the validity of φ_1 immediately at all time stamps but one ($n = 1$), where the output is $(\textbf{maybe}, 1)$, indicated by ≈. The asynchronous observer will resolve this inconclusive output at time $n = 11$, by generating the tuple $(\textbf{false}, 1)$, revealing a violation of φ at time $n = 1$. The verdicts of $\sigma_{S_{L\uparrow}}$,$\sigma_{S_{L\downarrow}}$, $\sigma_{S_{B\uparrow}}$, $\sigma_{S_{B\downarrow}}$, $\varphi_{S_{S\uparrow}}$, and $\varphi_{S_{S\downarrow}}$ are mapped to inputs S_L, S_B, S_S of the health model:

$$\sigma_{S_{L\uparrow}} = (\text{alt}_L - \text{alt}'_L) > 0 \qquad \sigma_{S_{L\downarrow}} = (\text{alt}_L - \text{alt}'_L) < 0$$
$$\sigma_{S_{B\uparrow}} = (\text{alt}_B - \text{alt}'_B) > 0 \qquad \sigma_{S_{B\downarrow}} = (\text{alt}_B - \text{alt}'_B) < 0$$

$\sigma_{S_{B\uparrow}}$ observes if the first derivation of the barometric altimeter reading is positive, thus, holds if the sensors values indicate that the UAS is ascending. We set S_B to *inc* if $\sigma_{S_{B\uparrow}}$ holds and to *dec* if $\sigma_{S_{B\downarrow}}$ holds. The specifications $\varphi_{S_{S\uparrow}}$ and $\varphi_{S_{S\downarrow}}$ subsume the pitch and the velocity readings to an additional, indirect altitude sensor. Due to sensor noise,

[3] Simulation traces are available in the Appendix; tools can be downloaded at `http://www.mentor.com` and `http://www.altera.com`.



Inputs to our rt-R2U2 are flight data, sampled in real time; a health model as BN, right; and an MTL specification φ. Outputs: health estimation (posterior marginals of H_L and H_B, quantifying the health of the laser and barometric altimeter) and the status of the UAS.

Fig. 3. Adding SHM to the Swift UAS

simple threshold properties on the IMU signals would yield a large number of false positives. Instead $\varphi_{S_{S\uparrow}}$ and $\varphi_{S_{S\downarrow}}$ use $\blacksquare_\tau\, \varphi$ observers as filters, by requiring that the pitch and the velocity signals exceed a threshold for multiple time steps.

$$\varphi_{S_{S\uparrow}} = \blacksquare_{10}\,(\text{pitch} \geq 5°)\ \wedge\ \blacksquare_5\,(\text{vel\_up} \geq 2\tfrac{m}{s})$$
$$\varphi_{S_{S\downarrow}} = \blacksquare_{10}\,(\text{pitch} < 2°)\ \wedge\ \blacksquare_5\,(\text{vel\_up} \leq -2\tfrac{m}{s})$$

Our real-time SHM analysis matched post-flight analysis by test engineers, including successfully pinpointing a laser altimeter failure, see Fig 3: the barometric altimeter, pitch, and the velocity readings indicated an *increase* in altitude ($\sigma_{S_{B\uparrow}}$ and $\varphi_{S_{S\uparrow}}$ held) while the laser altimeter indicated a *decrease* ($\sigma_{S_{L\downarrow}}$ held). The posterior marginal $\Pr(H_L = \text{healthy} \mid e^n \vDash \{\sigma_{S_L}, \sigma_{S_B}, \varphi_{S_S}\})$ of the node H_L, inferred from the BN, dropped from 70% to 8%, indicating a low degree of trust in the laser altimeter reading during the outage; engineers attribute the failure to the UAS exceeding its operational altitude.

6 Conclusion

We presented a novel SHM technique that enables both real-time assessment of the system status of an embedded system with respect to temporal-logic-based specifications and also supports statistical reasoning to estimate its health at runtime. To ensure REALIZABILITY, we observe specifications given in two real-time projections of LTL that naturally encode future-time requirements such as flight rules. Real-time health modeling, e.g., using Bayesian networks allows mitigative reactions inferred from complex relationships between observations. To ensure RESPONSIVENESS, we run both an over-approximative, but *synchronous* to the real-time clock (RTC), and an exact, but *asynchronous* to the RTC, observer in parallel for every specification. To ensure UNOBTRUSIVENESS to flight-certified systems, we designed our observer algorithms with a light-weight, FPGA-based implementation in mind and showed how to map them into efficient, but reconfigurable circuits. Following on our success using rt-R2U2 to analyze real flight data recorded by NASA's Swift UAS, we plan to analyze future missions of the Swift or small satellites with the goal of deploying rt-R2U2 onboard.

References

1. Alur, R., Henzinger, T.A.: Real-time Logics: Complexity and Expressiveness. In: LICS. pp. 390–401. IEEE (1990)
2. Backasch, R., Hochberger, C., Weiss, A., Leucker, M., Lasslop, R.: Runtime verification for multicore soc with high-quality trace data. ACM Trans. Des. Autom. Electron. Syst. 18(2), 18:1–18:26 (2013)
3. Barre, B., Klein, M., Soucy-Boivin, M., Ollivier, P.A., Hallé, S.: MapReduce for parallel trace validation of LTL properties. In: RV. LNCS, vol. 7687, pp. 184–198. Springer (2012)
4. Barringer, H., Falcone, Y., Finkbeiner, B., Havelund, K., Lee, I., Pace, G.J., Rosu, G., Sokolsky, O., Tillmann, N. (eds.): Runtime Verification (RV), LNCS, vol. 6418. Springer (2010)
5. Basin, D., Klaedtke, F., Müller, S., Pfitzmann, B.: Runtime monitoring of metric first-order temporal properties. In: FSTTCS. pp. 49–60 (2008)
6. Basin, D., Klaedtke, F., Zălinescu, E.: Algorithms for monitoring real-time properties. In: RV. pp. 260–275. LNCS, Springer (2011)
7. Bauer, A., Leucker, M., Schallhart, C.: Comparing LTL semantics for runtime verification. J. Log. and Comp. 20, 651–674 (2010)
8. Bauer, A., Leucker, M., Schallhart, C.: Runtime verification for LTL and TLTL. ACM Trans. Softw. Eng. M. 20, 14:1–14:64 (2011)
9. Colombo, C., Pace, G., Abela, P.: Safer asynchronous runtime monitoring using compensations. FMSD 41, 269–294 (2012)
10. Darwiche, A.: Modeling and Reasoning with Bayesian Networks. Cambridge University Press, New York, NY, USA, 1st edn. (2009)
11. Divakaran, S., D'Souza, D., Mohan, M.R.: Conflict-tolerant real-time specifications in metric temporal logic. In: TIME. pp. 35–42 (2010)
12. Finkbeiner, B., Kuhtz, L.: Monitor circuits for LTL with bounded and unbounded future. In: RV, LNCS, vol. 5779, pp. 60–75. Springer (2009)
13. Fischmeister, S., Lam, P.: Time-aware instrumentation of embedded software. IEEE Trans. Ind. Informatics 6(4), 652–663 (2010)
14. Geilen, M.: An improved on-the-fly tableau construction for a real-time temporal logic. In: CAV. pp. 394–406 (2003)
15. Havelund, K.: Runtime verification of C programs. In: TestCom/FATES. LNCS, vol. 5047, pp. 7–22. Springer (2008)
16. Ippolito, C., Espinosa, P., Weston, A.: Swift UAS: An electric UAS research platform for green aviation at NASA Ames Research Center. In: CAFE EAS IV (April 2010)
17. Johnson, S., Gormley, T., Kessler, S., Mott, C., Patterson-Hine, A., Reichard, K., Philip Scandura, J.: System Health Management: with Aerospace Applications. Wiley & Sons (2011)
18. Kleene, S.C.: Introduction to Metamathematics. North Holland (1996)
19. Lichtenstein, O., Pnueli, A., Zuck, L.: The glory of the past. In: Logics of Programs, LNCS, vol. 193, pp. 196–218. Springer (1985)
20. Lu, H., Forin, A.: The design and implementation of P2V, an architecture for zero-overhead online verification of software programs. Tech. Rep. MSR-TR-2007-99 (2007)
21. Maler, O., Nickovic, D., Pnueli, A.: On synthesizing controllers from bounded-response properties. In: CAV. LNCS, vol. 4590, pp. 95–107. Springer (2007)
22. Maler, O., Nickovic, D., Pnueli, A.: Checking temporal properties of discrete, timed and continuous behaviors. In: Pillars of Comp. Science. pp. 475–505. Springer (2008)
23. Pike, L., Niller, S., Wegmann, N.: Runtime verification for ultra-critical systems. In: RV. pp. 310–324. LNCS, Springer (2011)
24. Reinbacher, T., Függer, M., Brauer, J.: Real-time runtime verification on chip. In: RV. LNCS, vol. 7687, pp. 110–125. Springer (2012)
25. Schumann, J., Mbaya, T., Mengshoel, O., Pipatsrisawat, K., Srivastava, A., Choi, A., Darwiche, A.: Software health management with Bayesian Networks. Innovations in Systems and SW Engineering 9(4), 271–292 (2013)
26. Schumann, J., Rozier, K.Y., Reinbacher, T., Mengshoel, O.J., Mbaya, T., Ippolito, C.: Towards real-time, on-board, hardware-supported sensor and software health management for unmanned aerial systems. In: PHM (2013)
27. Tabakov, D., Rozier, K.Y., Vardi, M.Y.: Optimized temporal monitors for SystemC. Formal Methods in System Design 41(3), 236–268 (2012)
28. Thati, P., Roşu, G.: Monitoring Algorithms for Metric Temporal Logic specifications. ENTCS 113, 145–162 (2005)