Towards Readability Aspects of Probabilistic Mode Automata

Heinz Schmidt and Maria Spichkova

School of Science, RMIT University, Melbourne, Australia

Keywords: Software Engineering, Formal Methods, Petri Nets.

Abstract: This paper presents a new approach and design model targeting hybrid designer- and operator-defined performance budgets for timing and energy consumption. The approach is based on Petri Nets formalism. As the cognitive load is typically high while using formal methods, this increases the chances of mistakes. Our approach is focused on the readability aspects and aims to decrease the cognitive load of developers. We illustrate the proposed approach on example of a sample embedded multi-media system, a modern digital camera.

1 INTRODUCTION

In the domain of embedded systems, the trend to enhance more and more system functionalities through software solution is constantly increasing. This makes the design of these systems and the corresponding quality assurance more and more challenging Sangiovanni-Vincentelli and Martin (2001). Real-time and dependability constrains provide additional challenges, which also lead to necessity of probabilistic analysis within the phases of design and verification of these systems. Also, some constraints within embedded systems are mutually dependent, for example, timing and energy consumption constraints cannot be analysed independently of each other, see Mudge (2001); Saxe (2010); Wolf et al. (2008).

One of the successfully applied paradigms is Component-based software development, which was initially introduced many decades ago (CBSD, see Adler (1995); Clements (1995)). However, CBSD cannot solve directly issues related to the constraints on safety, timing, energy consumption, etc. Henzinger and Sifakis (2006); Spichkova et al. (2012), but can provide a solid basis for extended approaches.

In recent work Peake and Schmidt (2011) we have extended our rich architecture definition language (RADL, see Schmidt (2003)) and underlying theory Schmidt et al. (2003) to meet such industrial requirements, aiming at a scalable and compositional (component-based) approach to soft dependability guarantees: with probability, guarantee risk, execution time, cost etc. Industrial practice requires the capability to compose a variety of heterogeneous models and components, specified and designed using different methods and frameworks. Many real-world engineering environments are not locked into a single model, single framework or single-language environment. While we abstract from the programming languages underlying such an heterogeneous software engineering approach, we hope to show that, and how, our design-oriented model-based approach links with concrete programming by means of elementary modelling blocks providing abstractions directly for code blocks. This is natural and perhaps more appropriate in design of embedded systems than in other fields, as component models in this context often use architectural elements to abstract from software and hardware blocks at the same time. However we expect that this approach carries across to other domains.

In our current work, we targeting hybrid designer- and operator-defined performance budgets for timing and energy consumption. We propose an approach that is on Petri Nets formalism. Our approach is focused on the readability aspects and aims to decrease the cognitive load of developers, as having high cognitive load increases the chances of mistakes in system design and quality assurance process. We also aim to keep the method lightweight, following the classification presented in Zamansky et al. (2018).

To illustrate the proposed approach, we use an example of a sample embedded multi-media system, a modern digital camera. This allows us to demonstrate how the time (and the ensuing synchronisation) and energy constraints can be analysed taking into account their mutual dependencies. We propose that extra-functional properties have to be considered from early performance requirement specification through to model-based testing and run-time
verification. Beside the compositional approach to reasoning about and testing such properties in a hybrid modelling environment, our contribution is in the separation of concern of different aspects of modelling and in context-dependent methods of reasoning about such properties. Notably we have developed methods which allow automated contextual resource allocation strategies, under dynamically varying, and suitably parameterised, architectural configurations.

2 EXAMPLE: DIGITAL CAMERA

Consider the design of a modern digital camera from the perspective of different types of use:

Scenario 1: A busy professional sports photographer requires the ability to capture many hundreds or thousands of high quality images rapidly, with minimal shutter lag, in rapid bursts of up to 100 photos. Within the given price point afforded by budget, she is prepared to sacrifice “convenience” features, accepting shorter battery life and fewer shots per memory card while carrying extra battery packs, memory cards or even a laptop for frequent uploads, as well as extra lenses, and manage reconfiguration as needed.

Scenario 2: One weekend a family member is getting married, and as the de facto camera expert she has agreed to act as a semi-official or backup photographer for the wedding. In this capacity she aims for simplicity and convenience, so she can still enjoy the day and mingle without being conspicuous or weighed down by equipment. The couple insist they prefer photos in a standard compressed consumer format (JPEG), which at least eliminates extra effort later at her workstation, and maximises memory card capacity. She selects what she can carry easily—a single camera body and lens and perhaps a single additional memory card, but no extra battery pack. She is unwilling to spend anything like her usual time and effort on camera configuration, instead often (perhaps not always) relying on camera to automatically select exposure, focus and aperture. Occasionally, for particularly important shots she takes full control again. In this second case, battery life is paramount.

The specific challenge is to design a camera which is capable of flexible reconfiguration to suit multiple contexts, including for example these. The generic challenge is to:

(i) Characterise context in terms of user configuration choices, usage (e.g. selected modes/operations/functions) and user-visible desired properties.

(ii) Reason in a context-sensitive way about system properties and manage internal configuration to ensure consistency between configuration/profiles and desired properties. For example, to make the camera battery last longer, the camera must somehow sacrifice quality and/or performance in an acceptable way.

However the true usage context is often hard to predict. What exactly are the user’s requirements and intentions? Even the user may not know exactly what he intends beyond the immediate moment. Contextual uncertainty extends to environmental conditions, which may have a non-trivial impact on performance. For example ambient temperature may affect performance (e.g., energy consumption) of key camera components significantly, including batteries Rao et al. (2003), sensors and actuators such as lenses. This has implications for the design not only of embedded systems, but also at a macroscopic level. Thus, large-scale computing centres have significant inter-dependency on their local environment; such facilities are already planned with environmental conditions such as temperature in mind to be able to maximise performance and performance per cost while minimising cooling and energy consumption.

We extend the camera design presented by Lee Lee (2006). In our example, the camera has the following logical components:

- a general purpose processor (GPP),
- a digital signal processor (DSP),
- actuators to control, e.g., mirror and shutter curtain, lens focus and aperture,
- sensors, e.g., for auto-focus,
- a buffer to store images temporarily, and
- a flash memory as a long-term storage media.

To keep the example small enough for a conference paper, we abstract from other typical functions such as USB driver for photo download, LCD user interface, camera flashbulb, and various advanced settings.

In high performance scenarios a dedicated GPP-flash memory link is possible. We focus on the interplay of functionality relevant for taking a range of different shots involving real-time physical control, as well as selecting tradeoffs between timing and energy consumption.

As presented in Figure 1, the system has three modes, each with different resource requirements:

- IDLE mode covers waiting for shutter half/full press and pre-focusing.
- In single frame (SF) mode, the camera returns to the idle mode after shooting is completed, while
- in multi frame (MF) mode, shooting is continued as long as the shutter release button is kept pressed.
MF contains two sub-modes, high-speed (HS) and low-speed (LS). MF starts with HS and switches to LS if/when the image buffer gets full, where shooting of the consequent frame is delayed until enough space is freed in the buffer by writing to the flash memory. With these mode abstractions in mind, from a design perspective it is expected that refinements to components used in these modes may enable new features (for example smart/continuous save in HS at a performance penalty).

Figure 1: Digital Camera: Modes.

Furthermore, in each mode the user can select lens focusing and exposure metering to be performed automatically or manually, i.e. each mode has four sub-modes. More precisely, in the case of multi frame shooting, each of the MF submodes, HS and LS, has four further submodes:

- FE: automatic operations are fully enabled: both autofocus AF and automatic exposure AE are enabled;
- F: only the autofocus AF operation is enabled;
- E: only the automatic exposure AE operation is enabled;
- 0: neither autofocus AF nor automatic exposure AE are enabled.

In the IDLE mode the user may perform AF, AE or both, while composing a picture. During this time DSP cannot be activated and AF and AE operations are performed on GPP to reduce energy consumption. When the user presses the shutter release button, first, AF and AE operations that are being executed are completed, then the idle mode is terminated and the system switches to SF or MF depending on the user selection.

Another way to represent system modes (which can be related to the same submodes hierarchy as introduced in Figure 2) is to work parallel with on mode variables, because the choice to activate AF and AE operations is highly independent of whether the camera is in the IDLE, SF or one of the multi frame modes. Let call them CameraMode and AutoMode defined over enumeration types

\{IDLE, SF, HS, LS\}

and

\{FE, F, E, 0\}

respectively. We can also see this as a feature composition/interaction, see e.g., Calder and Magill (2000); Apel et al. (2010); Broy (2010). Thus, one feature is responsible for the choice of the current value of CameraMode and for the processes in the corresponding mode, where the second feature solely deals with the AF and AE operations.

Figure 2: Digital Camera: Submodes Hierarchy.

Figure 3: Digital Camera: Parallel Model for the Submodes Hierarchy.

Table 1 lists some of the relevant software components, their descriptions and their implementation platform (GPP/DSP). Some components are implemented in both processors to allow dynamic reconfiguration of the system in order to provide optimal resource usage. Within these constraints, a key challenge is allocating computing resources for the software elements to best suit partly predictable usage conditions. The DSP is especially suited to image processing operations, yet the DSP has significant energy overheads. We characterise the main design problems for the camera as follows. (i) Given an overall objective (e.g. minimise time consumption), satisfy that objective at run time. (ii) Given a usage profile, minimise energy and time consumption at run time.

3 PROPOSED VISUALISATION APPROACH

One of the problems using formal representation is that often only two factors are considered as important: the method must be sound and give such a representation, which is concise and beautiful from the mathematical point of view, without taking into account any question of readability, usability, etc.,
but even small syntactical changes of a method can make it more understandable and usable for engineers Constantine (2003); Dhillon (2004); Klare (2000); Spichkova (2013). Figure 4 presents an example of Petri net specifying HS mode details for the digital camera, which provides a typical representation of a coloured Petri net. Within our approach, we propose the following enhancements: To make representation more readable, first of all we should take into account the human factor. Thus, if a path (in this case a colour marked path, green or red) starts on the left/right of the net, we should proceed to draw it on the same side if possible and avoid cross moving the paths without any important reason.

Thus, on Figure 4 two paths are switched after the operation do AS, which can confuse some readers. Then, we can try to find a solution to avoid a lot of crossing arrows having different meanings: the blue and maroon arrows indicate synchronisation of the counters, and we can replace them by visual grouping of operations on the same counter. As result we obtain an optimised coloured Petri net presented in Figure 5, which is semantically equivalent to the representation in Figure 4. This optimisation increases ease of use by human readers (designers, testers etc.) without decreasing simplicity for machine readability and semi-automated support or expressiveness/power (for the domain or domains of choice).

Usability derives from the following aspects:

<table>
<thead>
<tr>
<th>Description</th>
<th>GPP</th>
<th>DSP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Operations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AF</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>AUTOFocus: Automatic lens focusing</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>AE</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Automatic Exposure metering</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>IS</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>Image Storage:</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>(e.g. aperture adjust)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IB</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Image Buffering:</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>Transfer image from sensor to local buffer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AS</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Activate Shutter etc.</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>BC</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Buffer check:</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Check if buffer is full</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Modes</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>IDLE</td>
<td>Idle mode</td>
<td>-</td>
</tr>
<tr>
<td>SF</td>
<td>Single-Frame shutter</td>
<td>✓</td>
</tr>
<tr>
<td>MF</td>
<td>Multi-Frame shutter</td>
<td>✓</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Submodes</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>FE</td>
<td>AF and AE enabled</td>
<td>✓</td>
</tr>
<tr>
<td>F</td>
<td>AF only enabled</td>
<td>✓</td>
</tr>
<tr>
<td>E</td>
<td>AE only enabled</td>
<td>✓</td>
</tr>
<tr>
<td>0</td>
<td>AF &amp; AE disabled</td>
<td>✓</td>
</tr>
</tbody>
</table>

Figure 4: HS mode details presented as a coloured Petri net.

Figure 5: HS mode details: Another Visualisation.
• Lowering the barrier between the simplified and expressive language for the machine support and that of the domain languages of the user(s) and associated with the purpose, e.g., by using controlled natural languages that try to avoid disadvantages of both natural and formal languages and being a subset of a natural language with a well-defined syntax and semantics, see Kuhn (2010); Fuchs and Schwitter (1995); Macias and Pulman (1993).

• Applying an appropriate automatisation of a number of steps within the modelling and verification process: this not only saves human time and allows to get results much faster then humans can produce manually, but also (partially) excludes the human element as the most “unreliable” in failure, see Redmill and Rajan (1996); Spichkova and Zamansky (2016). For example, a formal specification can be generated from the corresponding CASE tool representation which can be edited in a more readable way also using predefined templates, see Spichkova et al. (2013); Vo and Spichkova (2016).

• Supporting directly common and standard abstractions that are well-established (and hence part of the software engineering training), e.g. Message Sequence Charts Genest and Muscholl (2008); Harel and Thiagarajan (2003), or defined in standards (such as UML, IEC-61131, etc):

• Unification of the representation of any information we are dealing with (see, e.g., Spichkova (2013));

• Easing the use of novel compositional principles and high-level tools, that are opening novel and powerful methods to users of formal specification or specification-based/model-driven methods.

Having a representation like presented on Figure 5, we can easily transform a Petri net to a hierarchical MSC. In the case a component-based specification of the system is need in addition to the process representation, an MSC can be schematically translated to the corresponding formal specification as shown in Spichkova et al. (2012); Spichkova (2010). Let us also shortly discuss translation/representation of the following modelling artefacts: (global) parameters, local, time and counter variables.

Local Variable: use can be translated into state and transition label expansions for NF purposes Katoen (2005), but can also be intuitively understood in data types and data structures that capture state. However for Petri net normal form used in our approach and compositionality considerations restrictions need to be designed along with such capabilities to limit the scopes of these variables appropriately, viz. to FSM components of nets, in terms of their use in guards and assignments associated with transitions and states.

Global Parameter: use are of a similar nature with respect to normalisation but needs to be limited to achieve compositionally. For example, we could say the global parameters may occur locally in guards (i.e., they are read-only) as well as in initialisation expressions (for the initial states when FSM objects are created) or with re-assignments limited to higher-level FSMs (such as mode automata) when submodes are entered and before these branch out into rational parallel processes. Another common example is the use of iterator and bounded loop process constructs that have a very structured use of local counter variables which never serve synchronisation but are providing a reasoning tool for local termination and performance approximation, based on an interplay of local (loop) invariants and loop control variables, which implies strict monotonicity and boundedness.

Time Variables (Clocks): are a further example and in some sense a special case of counter control above – in the sense that all practical approaches to timed automata and synchronous time models discretise an infinite number of real-time points into a real-time intervals with integer bounds and then solve a linear convex hull problem to determine feasibility and/or optimal schedules that meet time constraints. There is also a significant difference here, that needs to be considered, relative to counters. In general, counter processes can be explained as a macro structure based on sequence and choice, and hence are lower-level automata (or process expressions) themselves, and hence they do not add ‘new’ semantics but can be explained in terms of existing semantics. For example if we are in rational parallel processes, they are just a syntactic sugar extension that does not take us out of this class. Likewise with other classes of processes (such as pushdown automata). In contrast, timed extension are true semantic extensions, in that they define a different class of behaviours and automata, because the define what the legitimate processes (occurrence nets) are that are traces of the given language (net system or process expression).

4 RELATED WORK

Component-based software engineering utilises a well-defined composition theory to enable the prediction of such properties, as performance and reliability. This is one of the largest fields of software and system engineering, there are many approaches on component-based design (CBD) cover-
ing different aspects and focusing on requirements, quality, timing properties etc. (see e.g., CBS (2003); Broy (2010, 1999); Broy et al. (2008)). Several component-based prediction approaches, e.g. Palladino Kapova et al. (2010); Martens et al. (2008); Becker et al. (2007), CB-SPE Bertolino and Miranda (2004), ROBOCOP Bondarev et al. (2004) (see also a survey in Becker et al. (2006)) derive the benefits of reusing well-documented component specifications. In our approach we focus on the questions of resource-awareness and adaptivity of systems as well as on the readability aspects of the formalism.

Mode automata have a long history motivated by real-time design practices and methods used in industry in connection with statecharts. Maraninchi et al. Maraninchi and Rémond (2003) capture the notion of modes formally for a practical extension of the real-time synchronous language Lustre and include elements of the well-known I/O-automata. Mode automata define synchronous mode automata as a hybrid between data-flow and transition systems. Talpin et al. Talpin et al. (2006) extend this work to so-called polychronous mode automata to work with the multi-clock data-flow formalism SIGNAL. Both these types of automata are non-deterministic and do not deal with probabilities. The (bisimulation) equivalence and therefore compositional reasoning for mode automata is undecidable. However, Maraninchi et al. introduce a synchronised (lock-step) parallel product for modes in which shared symbols (intersection of alphabets nonempty) are synchronised while local symbols (the symmetric difference of the alphabets) are independent. While the modes of a single automaton are mutually exclusive in their approach, and the behaviour of these mode automata is fully abstract wrt. probabilistic testing, the automata product suffers from combinatorial explosion (state space explosion), due to the aim of allowing arbitrary shared variables and interference of parallel processes.

Cheung et al. Cheung et al. (2012) describe an architecture-level method, SHARP, for predicting reliability (and timing) of concurrent systems. Whereas SHARP is specifically designed for reliability and timing prediction, our method is intended to be generic thus also catering, e.g., for energy consumption. SHARP models involve scenarios which are either basic (similar to message sequence charts) or hierarchical, involving sequential, conditional or concurrent composition. SHARP supports concurrent composition of finite numbers of instances of a particular scenario, corresponding to symmetrically replicated components. SHARP derives completion time and reliability predictions from scenarios for use at higher levels of abstraction. For each basic scenario, SHARP requires transition rates for all individual actions, then calculates a single continuous-time Markov model model from which completion time and reliability are derived. For an hierarchical scenario a system level CTMC is constructed using abstraction techniques such as queuing networks and abstraction of sequential components into single global states. In contrast, our approach requires probabilities/rates at the system level only. Our approach seeks to avoid or defer calculation of monolithic models.

Our cost estimation is inspired by Valiant’s bulk synchronous-parallel model Valiant (1990) of parallel computing where global strong synchronisation conservatively approximates systems which may in reality use more fine grained synchronisation and indeed may allow for more asynchrony than the above approximation would suggest. In performance benchmarks reported in Yusuf et al. (2011), Yusuf et al. demonstrated that such conservative predictions may still be accurate enough if there is enough WCET variation and a large enough number of activities/tasks scheduled on individual processing elements. Thus adjacent modes may be assumed to be strongly separated in the global model while in fact such modes are partially interleaved with respect each other (subject to restrictions on repetition such as boundedness for message sequence graphs as described by Alur and Yannakakis (1999) and star-connectivity in trace languages). For conservative cost estimation purposes this seems reasonable. We expect that (with diminishing returns) such models can be refined selectively, to bound costs of adjacent sequences of overlapping modes, in a context-dependent way.

An interesting approach on integration of synchronous and asynchronous communication was presented by Hennicker et al. Hennicker et al. (2010); Hennicker and Knapp (2011). In this approach, I/O-transition systems were used as the formal background for modelling of system behaviour. As result, a refinement relation was defined, which is compositional w.r.t. synchronous and asynchronous connections of components and which preserves connection-safety, and next existing interface theories for modal I/O-transition systems were extended to support assemblies, (greybox) assembly refinement and assembly encapsulation, also showing that communication-safety is preserved by assembly refinement, that black-box refinement of component interfaces is compositional w.r.t. grey-box refinement of assemblies and, conversely, that assembly encapsulation maps grey-box to black-box refinement.
5 CONCLUSIONS

In this paper, we proposed a Petri-Nets-based approach targeting hybrid designer- and operator-defined performance budgets for timing and energy consumption. The core focus of this approach is on decreasing the cognitive load of the designers to decrease the chances of design mistakes. To achieve better readability, we extended the coloured Petri Nets formalism. To illustrate the proposed solution, we presented an example of a sample embedded multimedia system, a modern digital camera.

Future Work: We are going to integrate the presented approach with the results of our prior work, a probabilistic global behaviour analysis approach developed for reliability and fault-tolerance studies (including fault injection) and a parallelism/concurrency focused framework centred on partially ordered traces, Petri nets and timing/energy costs.

REFERENCES


