Challenges in Applying Control Theory to Software Performance Engineering for Adaptive Systems

Davide Arcelli
University of L’Aquila, Italy
davide.arcelli@univaq.it

Vittorio Cortellessa
University of L’Aquila, Italy
vittorio.cortellessa@univaq.it

ABSTRACT
Control theory has recently started to be applied to software engineering domain, mostly for managing the behavior of adaptive software systems under external disturbances. In general terms, the main advantage of control theory is that it can be formally proven that controllers achieve their goals (with certain characteristics), whereas the price to pay is that controllers and system-to-be-controlled have to be modeled by equations. The investigation of how suited are control theory techniques to address performance problems is, however, still at the beginning. In this paper we devise the main challenges behind the adoption of control theory in the context of Software Performance Engineering applied to adaptive software systems.

CCS Concepts
• Software and its engineering → Extra-functional properties: Software performance; • Control methods → Computational control theory; • Software system structures → Software system models; Model-driven software engineering;

Keywords
Adaptive Software; Control Theory; Software Performance engineering;

1. INTRODUCTION
In the last few years, the massive introduction of software to different sectors of human society has led to complex ecosystems of systems. On the one hand, the large number of software systems and services interconnected through Internet has introduced a high degree of dependency and complexity among systems that takes the maintenance task beyond human intervention. On the other hand, customers expect particular qualities from software, e.g. the fulfillment of specific performance requirements, but since it is plunged in an unpredictable environment the requirements can also change at runtime. To tackle these issues, modern software systems are designed to be highly self-adaptive, i.e. capable of observing changes in their environment and modify their behavior accordingly to reach their goals.

Control theory tackles the challenge of adapting (physical) plants since decades in many real world domains (e.g., robotics, Cloud computing) [12, 13], and in the last few years it has started to be applied to adaptable software [29, 14, 15, 31, 16]. While adaptation of an application’s functional aspects (i.e., semantic correctness) often requires human intervention, its non-functional aspects (such as reliability, performance, energy consumption, and cost) represent an important and challenging opportunity for applying self-adaptive techniques [6]. Adaptation actions and policies are triggered (in a proactive or reactive way) to allow a software system to offer acceptable levels of Quality of Service (QoS), while preserving semantic correctness with respect to functional requirements. For example, customers may require continuous assurance of agreed performance indices (such as response time) that can be used to trigger adaptations guaranteeing requirements even in the face of unforeseen environmental fluctuations [9].

Despite the undisputed benefit of formally guaranteed control, the application of control theory to software is fairly limited [31]. This paper points out the most challenging research issues for adopting control engineering techniques in the context of software performance engineering for adaptive systems.

The paper is structured as follows. Section 2 gives an overview on adaptable software systems with embedded controllers, both from the control and software engineering viewpoint. Section 3 describes a set of common steps for designing controllers for adaptive software systems. Based on such steps, Section 4 points out the most important challenges in applying control theory to software performance engineering. Section 5 sketches some related work, and finally Section 6 concludes the paper.

2. CONTROLLING ADAPTIVE SOFTWARE
Figure 1 shows a general scheme of an adaptive system driven by a controller. Basically, it consists of a system-to-be-controlled, which has to adapt its behavior in order to meet predefined goal(s), despite an uncontrollable environment produces disturbances, that are unpredictable events that affect the system behavior. The adaptation is performed by a Controller that manipulates available knobs of the system, for sake of adaptation, on the basis of system properties monitoring.
2.1 Control Engineering perspective

Control systems are backed by a control theory that undergoes the mathematical tools supporting the definition of controllers with formally proven quality properties [1,29].

The design of a controller is based on a mathematical model of the controlled system behavior. Such model is usually a dynamic model defined by differential or difference equations. The system is usually identified with the software application under control, but it can also contain some information about the execution environment and platform. The dynamic model formalizes the relationships between the time, the system states (through a transfer function that maps the input of a system to its output in the frequency domain [18]), the control variables (i.e., the knobs), and the controlled variable (i.e., the monitored system output we want to effect). The controller input is the error between the goal(s) and the value(s) monitored from the plant, whose multiplicities should coincide. The controller output is an enforced knob configuration, i.e., such that it is closer to the goals than the previous one.

A broad variety of disturbance types have been studied to characterize physical phenomena as well as to represent errors or uncertainties about the system [26]. The main purpose of control theory is to minimize the effect of such disturbances on the controlled output variables.

A proper dynamic model of the controlled system and the relevant environmental phenomena allows the application of a broad variety of (more or less automatic) techniques for the design of controllers engaging several important qualities [26,4].

2.2 Software Engineering perspective

From a software engineering perspective, the distinction between the environment and the adaptive system is made based on the extent of control [21,36]. Typically, the environment is monitored for changes by an adaptive software system. The system represents the domain specific application functionality. The controller adapts the system according to some logic that deals with the system goal(s). To keep the goal(s) from being violated, the controller continuously monitors the environment and the system, starting the adaptation process when needed 1.

Core components of an adaptive software system can be totally separated or highly intersected in terms of implementation, and they can also be centralized or decentralized. A known approach to build a control system is the Monitor-Analyze-Plan-Execute (MAPE) feedback loop [21]. The loop relies on a Knowledge component that stores runtime models maintained by the control system to support the MAPE functions. The control system can consist of one or more feedback loops, while MAPE functions can be implemented by different components or integrated in a single component.

2.3 Glossary

Differences in the terminologies used in control and software engineering have practically limited the potential application of control engineering techniques to software engineering problems [7]. In this section we introduce a short glossary of the main terms used in this paper, and we emphasize their meanings in control and software engineering.

- A goal is a reference value (or range) for a certain property of the system. From the software engineering perspective, it coincides with a QoS requirement, e.g., on the response time of the system. From the control engineering perspective, it is a reference value for a controlled variable, namely setpoint.

- A knob is an instrument for modifying some system properties. From the control engineering perspective, using a knob means to manipulate a control variable. From the software engineering perspective, it means performing an adaptation action, e.g., allocating a certain number of virtual machines.

- A model in the software engineering domain is generally defined as “an abstraction of reality” [6]; different models may be used in different development phases (requirements, architecture design, deployment) and for different purposes (communication, analysis, code generation). In the control engineering domain, instead, a model usually takes an analytical form as a set of differential equations [18,33]. Such difference is not only syntactical, but also semantic, and it may limit the applicability of control engineering techniques to software engineering. The purpose of modeling in this domain is always to devise a controller that adapts the system to keep a setpoint, i.e. goal.

- A controlled system in the control engineering domain is the aggregate of the plant and the controller [17], whereas in the software engineering domain the same term usually only refers to the system-to-be-controlled.

3. DESIGNING ADAPTIVE SOFTWARE

Most of current control theory approaches target running software, with the aim of leading minimal changes to the latter. While controlling existing software is a most wanted capability to empower it with self-adaptive functionalities, just as much important is to define development processes accounting for the controllability of the software system. Controllability can be informally defined as the capability of a software system to meet target goals under external disturbances. This is because controllability has an impact on the underlying control system or to control the cooperation of multiple control systems.

1 Additional controller can be added on top of an adaptive system to, for example, manage the adaptation logic of an...


on many stages of software life-cycle: (i) in requirements analysis, where suitable stakeholder needs can be matched to control goals, (ii) in design, where controllers may become first class elements exactly like dynamic behavior of the software is, (iii) in quality assurance, where the active role of the controllers introduces new challenges that will be discussed in the remainder of this paper.

On the basis of current practices, a set of steps can be identified for designing controllers for adaptive software systems [18], and they are summarized in the following.

1. **Define the goals.** Quantifiable and measurable system goals (i.e., requirements) are defined.

2. **Identify the disturbances.** Non-controllable system variables coming from the external environment are identified.

3. **Build the software model.** Once completed the previous steps, a software behavioral model has to be built, usually represented in terms of equations.

4. **Define the knobs.** What can be modified on the system and the domain of the involved changes have to be defined.

5. **Design the controller.** Given the model of the system, a controller has to be synthesized, and this step can be accomplished with several approaches. An example is to choose a controller structure, for example a Proportional Integral and Derivative (PID) controller [3], and to select the parameters for the chosen structure.

6. **Prove properties of the closed-loop system.** Once the model and the controller are both defined, the next step is to analyze the closed-loop system and to prove its intrinsic properties, namely SASO properties [18].
   - **Stability:** A controlled system is asymptotically stable if, under reasonable assumptions on the initial state, the system will tend to an equilibrium point (i.e., for any given input, the output converges to a specific value) within a convenient accuracy.
   - **Accuracy:** The magnitude of the control error at the system steady-state has to be kept conveniently short.
   - **Settling time:** The time to converge to the setpoint (or the closest feasible equilibrium) has to be kept conveniently short.
   - **Overshoot:** The system is required to converge to the setpoint despite both the effect of disturbances and possible inaccuracies in the dynamic model.

If the desired SASO properties are not satisfied, then either a different controller has to be synthesized or new knobs have to be devised and added into the system. This iterative process continues until the desired properties are met.

7. **Implement and integrate the controller.** Once the controller is designed and proven, it has to be implemented and integrated in the system.

8. **Controller testing and validation.** Finally, the controlled system has to be tested and validated towards the goal(s).

### 4. CHALLENGES

One major advantage of control theory comes from the analytical guarantees it can provide on the system behavior, due to its mathematical grounding [18] [15]. In order to exploit such an advantage in the software performance context, performance-specific aspects have to be treated in the design process of adaptive software. In this section we devise the main challenges under the application of the steps in Section 3 in a software performance context.

#### 4.1 Define the goals

Performance requirements are typically defined on usual indices: response times, throughputs, and utilization. The goals of a controlled adaptive software system obviously come from such requirements. The use of control theory to achieve performance goals heavily depends on the granularity to which the latter are considered. A performance goal defined on a single component/device is more controllable than one defined on a subsystem or on the whole system. This is due to the fact that, in the former case, a controller can perform local adaptation (without considering the rest of the system), whereas in the latter case a controller has to achieve the goal despite different (and possibly conflicting) behaviors of many system entities.

Moreover, performance goals are not always simple to address, because the ideal range for performance indices could be very narrow. For example, a throughput goal with respect to a minimal threshold has to be targeted while controlling that the system throughput does not achieve too high values that may lead to system saturation.

Nowadays, an ever increasing amount of software applications run on mobile devices with limited resources. Hence, in this context energy and temperature are keen concerns to deal with. For example, a mobile software system that performs a high amount of computation requires a high amount of energy, thus leading CPU(s) towards high temperature and energy consumption. Defining and achieving energy consumption and temperature goals may help to mitigate such effects.

#### 4.2 Identify the disturbances

Disturbances are a source of uncertainty. In fact, as they can change at runtime, the dynamic model of the system has to capture such changes, evaluate the need for an adaptation, and possibly perform the latter in order to face them. In the software performance domain, two main sources of disturbances can be identified: (i) workload and (ii) operational profile. Usually, these types of parameters are represented by probability distribution functions. A significant advantage of applying control theory to software performance is that it allows to deal with such disturbances even when they assume the form of irregular curves. This is particularly useful when workload and/or operational profile come from empirical observations of a running system (see the definition in the context of software engineering in Section 2.3).

#### 4.3 Build the software model

In order to obtain mathematical models of software systems, numerous methods from system identification can be used [27]. Broadly, these methods try to obtain an accurate model from input-output data recorded on a running system. Software is particularly prone to this approach be-
cause running a piece of software is typically much cheaper and can be done more often or faster than for a physical system. This means large amounts of data for identification purposes can be available, possibly from previous releases of the software system under development.

However, unlike physical systems, software behavior is hard to model by means of dynamic systems of equations, because the algorithmic nature of a program often leads to the introduction of complex non-linearities in the models, thus reducing their suitability for control theory application.

Although most of the analytical models used to describe software performance are not straightforwardly mappable to a dynamic system of equations, they can be exploited to fill the semantic gap between software models/artifacts and equations, as firstly proposed in [14].

The benefit of using established analytical models as pivot for generating dynamic systems of equations is twofold: (i) it simplifies the construction of the dynamic model by providing a more concise and precise quantitative view on the system; (ii) it broadens the applicability of the controller design methodology to every system that can be formalized through the intermediate model. In other words, defining a general control design methodology for Queuing Networks (QNs) would allow the construction of controllers for a variety of systems whose performance concerns can be captured through a QN model.

4.4 Define the knobs

Knobs are fundamental for adaptive systems, because they allow the system to perform adaptation policies for facing disturbances. For example, a possible goal for a QN model is to keep queue length under a certain threshold. When this value increases, a knob should be available to allow a service center to increase its rate, and vice versa. However, it is worth to notice that the types of knobs depend on the analytical model (e.g., a Markov Chain rather than a Queuing Network).

Here below we identify some knobs with a specific reference to QN models, and for each knob we provide an example of corresponding adaptation action on the software system:

- Changing the rate of a service center. As an example, the same request can be served at different rates because different alternative implementations of the same service are deployed on the same node.
- Changing service center multiplicity. As an example, additional virtual machines and/or disk mirrors can be allocated/deallocated.
- Changing job routing among similar service centers that process jobs in parallel, with the aim of balancing their loads. As an example, this could correspond to increasing the hit-rate of a cache with respect to the access to main memory.
- Decreasing the demand to a set of resources and increasing the one to another set of resources. This action could represent, for example, a software component splitting and consequent re-deploying on a different site.
- Moving communication demand (e.g., network) to computation demand (e.g., CPU). As an example, zipping messages introduces additional computation for compression but reduces communication due to shorter messages.

More complex adaptation patterns might be devised as additional knobs. Performance antipatterns can be helpful in this direction. A performance anti-pattern is in fact a common solution to a recurring performance problem.

In knowledge-based control (e.g., MAPE feedback loop), the knowledge is typically represented as adaptation policies/strategies. Such knowledge can be learned by employing an appropriate learning mechanism such as reinforcement learning [35]. However, identifying when a learned knowledge should replace an existing one, i.e., turning point detection, is a relevant challenge in this context. With this respect, introducing performance anti-patterns awareness might help in such identification. In fact, if a performance anti-pattern is detected, probably the current knob is not suitable anymore, hence an adaptation is needed to solve the detected performance problem. Once such adaptation has been performed, if the performance anti-pattern has been removed then the system might learn that such adaptation could be performed for removing the same anti-pattern in the future.

4.5 Design the controller

Control engineering generally works by applying a specific control technique (e.g., linear control, quadratic control, model-driven control, etc.) to an instance of a problem. In contrast, software engineering prefers to use methodologies or off-the-shelf components that are applicable in a wide range of cases. Therefore, in order to make control techniques useful to software engineers without requiring special skills, a solution consists in finding types of controllers that apply to classes of applications. As an analogy, this is similar to the way design patterns have emerged to aid software engineers in taking design decisions.

Currently, there are no tools that support the design of controllers as part of the application. Hence, if we stay at the “software engineering side”, then we must implement one or more components for a designed controller(s). If we move to the “control engineering side”, then we have some support to the design of controllers, but we have to redefine the model as it was described, for example, in [2], where Queuing Network constructs have been redefined in Modelica to embed controllers in QN models.

Control anti-patterns are analogous to performance anti-patterns when control theory is used in the design of software systems that fulfill their performance requirements. Coming up with a list of common control anti-patterns (and suggested remedies) would allow software engineers to more effectively use control engineering techniques. For example, having two uncoordinated controllers targeting the same goal with different knobs may lead to undesirable behavior of the software system, such as oscillations. Unless particular provisions are taken in the control scheme design (e.g., the two controllers act on very different time scales), this would lead to the adoption of a specific control pattern such as cascade control [34].

As combining software engineering and control theory is still in an early phase, extensive research would need to be performed, either top-down or bottom-up. Top-down would consist in first devising controllers and then testing what applications would most benefit from them, as done in [12]. Bottom-up would consist in first finding applications that require to be adaptive, then applying control techniques to each of them, while trying to generalize the best found
methodology, as done in [23]. However, not for all combinations of model, knobs, goals, and disturbances, it is possible to build controller with formally-proven guarantees.

4.6 Prove, implement and test/validation

Nothing is really specific in proving, implementing and testing/validating a controller in the context of software performance engineering, once the previous steps have been accurately executed.

5. RELATED WORK

Control theory [12, 18] is capturing an increasing interest from the software engineering community that looks at adaptation as a mean to meet QoS requirements despite unpredictable changes of the execution environment [31]. Examples of this trend can be seen in research on control of web servers [22, 28], data centers and clusters management [13, 24], operating systems [25], and across the system stack [20].

In the domain of model-based performance control of adaptive software, some effort has been spent to raise adaptation techniques driven by performance (or more in general QoS) requirements at the software architecture level [32], where adaptive verification techniques have also been studied [33]. Adaptation approaches for specific architectural paradigms, domains, and problems, have been introduced, such as, respectively, Service-Oriented-Architectures [10], mechatronic systems [5], and .NET thread pools [19]. An interesting work has been introduced in [37] for automatically extract adaptive performance models from running applications. An overview of techniques that apply control theory to software engineering can be found in [31, 36]. However, none of such techniques applies it for controlling performance models, like recently done in [2].

6. CONCLUSION

In this paper we have worked towards the identification of the main research challenges on the way to apply control theory techniques to software performance engineering domain. In particular, we have provided examples of specific aspects of software performance that require control theory to be tailored to the domain. We believe that this field is still very unexplored and, despite the intrinsic limits of application of control theory, it can represent a sharp solution for a range of software performance problems, especially the ones related to adaptive software systems.

7. ACKNOWLEDGMENTS

The authors are very grateful to Alberto Leva (Politecnico di Milano) and Antonio Filieri (Imperial College) for being their guides (and co-authors) in the field of control theory. The ideas in this paper stem from previous work and discussions with them.

8. REFERENCES


