Chapter 7
Towards Adaptive Enterprise:
Adaptation and Learning

Harshad Khadilkar
https://orcid.org/0000-0003-3601-778X
TCS Research, Tata Consultancy Services, India

Aditya Avinash Paranjape
TCS Research, Tata Consultancy Services, India

ABSTRACT

The key to a successful adaptive enterprise lies in techniques and algorithms that enable the enterprise to learn about its environment and use the learning to make decisions that maximize its objectives. The volatile nature of the contemporary business environment means that learning needs to be continuous and reliable, and the decision-making rapid and accurate. In this chapter, the authors investigate two promising families of tools that can be used to design such algorithms: adaptive control and reinforcement learning. Both methodologies have evolved over the years into mathematically rigorous and practically reliable solutions. They review the foundations, the state-of-the-art, and the limitations of these methodologies. They discuss possible ways to bring together these techniques in a way that brings out the best of their capabilities.

INTRODUCTION

A key objective of the digital enterprise is to improve the speed and quality of response of the enterprise to external inputs or context. IT infrastructure and its security are prerequisites for working towards this objective. In this chapter, we describe methodology to use the existing infrastructure and the idea of automated decision-making to make a digital enterprise *adaptive*. This is an overarching term that covers both the flexibility of the automated decision-making process when first deployed, and also its ability to respond to changing business environments.

DOI: 10.4018/978-1-7998-0108-5.ch007

As such, any methodology that leads to automated decision-making needs two common ingredients. First, it must be able to measure and suitably quantify its own performance vis-à-vis its objectives. Second, it must be able to build an internal model (either explicit or implicit) of its environment. The decision-making process solves the following problem: given the available information (model), maximise the performance metrics of the organization.

When the environment and/or the enterprise are volatile, or time-varying, it follows that the internal model used for decision-making (even if it is implicit) must be tweaked continuously. More often than not, the cues needed for tweaking come from an observation of the system's interaction with the environment rather than through any pre-defined prescription. We refer to this process as *learning* or *adaptation*, although the contexts in which these terms are employed are slightly different and this will become apparent in the chapter.

We cover two approaches, broadly divided into a model-based technique (Model Reference Adaptive Control, or MRAC) and a model-free technique (Reinforcement Learning, or RL). By model-free, we mean that the technique relies on an implicit model of the system rather than an explicit model. These notions will become clear later in the chapter. While RL can also accommodate model-based approaches, their implementation closely resembles existing model-driven control methods. MRAC and RL can be viewed as controllers for driving the state of a plant, which represents the enterprise. Figure 1 illustrates the conceptual similarities and differences between them. MRAC aims to track a reference signal r and uses a model M to update a control policy parameterised by $\hat{\theta}$. On the other hand, the form of RL shown in Figure 1, and known as advantage actor-critic [Konda and Tsitsiklis, 2000], replaces the model by a critic C, which evaluates the goodness of the current state of the plant. This estimate is fed to an actor, which computes the decisions or actions. The effect of the actions is quantified in the form of a reward signal R, which is used for the learning process.

Figure 1. High-level block diagrams of MRAC and a form of model-free RL known as A2C.