A Metareasoning Framework for Planning and Execution in Autonomous Systems

Justin Svegliato[1]

Abstract. *Metareasoning*, a particularly effective computational approach to bounded rationality, is the process by which an autonomous system optimizes its own planning and execution in order to act effectively in its environment. The need for metareasoning has become critical to autonomous systems due to the uncertainty about the range of their potential circumstances and the limitations of their reasoning capabilities. My thesis develops a novel metareasoning framework for monitoring and controlling both the *planning* processes and *execution* processes of autonomous systems. The planning module employs efficient metareasoning techniques that relax unrealistic assumptions of earlier work while the execution module employs robust metareasoning techniques that resolve a range of exceptions and maintain a level of safety. The result is a modular and nonmyopic metareasoning framework that monitors and controls the planning and execution of autonomous systems.

1 INTRODUCTION

It has long been recognized that autonomous systems cannot be capable of perfect rationality due to the intractability of optimal decision making in complex domains [5]. As a result, there have been substantial efforts to develop computational approaches to bounded rationality [4]. *Metareasoning*, a particularly effective computational paradigm for bounded rationality, enables an autonomous system to optimize its own planning and execution processes in order to act effectively in its environment. This enables the autonomous system to handle any uncertainty about the range of its potential circumstances and the limitations of its reasoning capabilities. Consequently, due to the growth in the complexity of autonomous systems in recent years, metareasoning has become critical to automated decision making.

There has been considerable progress in developing metareasoning techniques for monitoring and controlling the planning processes of autonomous systems. For example, a recent method identifies the best algorithm to solve a problem among a portfolio of algorithms by compiling a model with a limited number of features to predict the efficiency and accuracy of the each algorithm [3]. Another recent method selects the next computation, specifically the next simulation, to be performed by Monte Carlo search techniques by representing the decision as a Bayesian selection problem that maximizes the value of information [2]. There are also many methods that determine when to interrupt an anytime algorithm and act on the current solution by using a profile that represents the performance of the anytime algorithm [1]. However, despite these advances, developing metareasoning techniques for monitoring and controlling the execution processes of autonomous systems has not seen much attention.

[1] College of Information and Computer Sciences, University of Massachusetts Amherst, USA, email: jsvegliato@cs.umass.edu

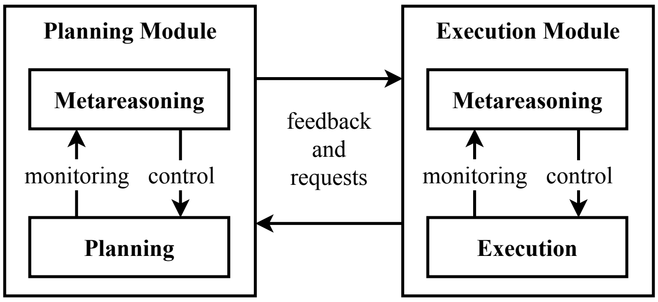

Figure 1. A metareasoning framework for monitoring and controlling the planning processes and execution processes of an autonomous system.

In my thesis, I develop a metareasoning framework for autonomous systems that improves the efficiency of meta-level control for planning processes and expands the scope of meta-level control to execution processes. The framework is composed of a pair of modules shown in Figure 1. The *planning module* enables the system to monitor and control its planning processes to generate policies for acting effectively in its environment within an acceptable amount of time. My thesis proposes methods used by the planning module that relaxes unrealistic assumptions made by earlier work. The *execution module* enables the system to monitor and control its execution processes to act appropriately on the policies generated by the planning module. My thesis introduces methods used by the execution module that resolve a range of exceptions and maintain a level of safety.

The general aim of my thesis is therefore to build a metareasoning framework, which is composed of a planning module and an execution module, that makes the following contributions.

1. **Online Performance Prediction.** Determine the optimal stopping point of an anytime algorithm by predicting the performance of the algorithm online to avoid relying on substantial offline preprocessing needed to compile and maintain a performance profile.
2. **Model-Free Meta-Level Control.** Employ reinforcement learning methods to learn and adapt a policy that indicates when to interrupt an anytime algorithm and act on the current solution to handle changes in the parameters of meta-level control.
3. **Adjustable Anytime Algorithms.** Adjust the hyperparameters of an anytime algorithm at runtime to attain the best solution in the shortest amount of time using deep reinforcement learning.
4. **Exception Recovery.** Resolve a range of exceptions that prevent an autonomous system from completing its task by recovering from unanticipated scenarios during execution.
5. **Safe Operation.** Maintain a level of safety as an autonomous system completes a task by periodically monitoring and proactively avoiding potentially unsafe situations during execution.

2 COMPLETED WORK

We describe the contributions that have been made toward the planning and execution modules of the metareasoning framework below.

Contributions 1 and 2. Autonomous systems rarely have enough time to determine the optimal solution to real world decision-making problems. To generate an acceptable solution under strict time constraints, an autonomous system often uses an anytime algorithm that gradually improves the quality of a solution as it runs and returns the current solution if it is interrupted. However, to exploit the trade-off between solution quality and computation time, it must decide when to interrupt the anytime algorithm and act on the current solution.

Existing metareasoning techniques that monitor and control anytime algorithms rely on planning with a model, called a *performance profile*, that describes the performance of a given anytime algorithm solving a specific problem on a particular system [1]. This model is compiled offline before the activation of meta-level control by using the anytime algorithm to solve thousands of instances of the problem on the system. Planning with a model, however, imposes many assumptions often violated by autonomous systems in the real world. First, there must be enough time for offline compilation of the performance profile of the algorithm. Second, the settings of the algorithm across every problem instance must be the same. Third, the distribution of problem instances solved by the algorithm must be known and fixed. Fourth, the CPU and memory conditions of the system executing the algorithm must be static.

Addressing these unrealistic assumptions, we propose two metareasoning approaches in recent work [9, 10, 7]. Both approaches monitor the performance of the anytime algorithm and estimate the stopping point at runtime by expressing the state of computation in terms of solution quality and computation time. However, the first approach predicts the performance online *with* a model by using a performance predictor while the second approach learns the performance through experience *without* a model by using reinforcement learning. For each approach, we empirically show their effectiveness on a set of common benchmark domains and a mobile robot domain.

Contribution 4. Due to the complexity of the real world, autonomous systems use decision-making models that rely on simplifying assumptions to make them computationally tractable and feasible to design. However, since these limited representations cannot fully capture the domain of operation, an autonomous system may encounter unanticipated scenarios that cannot be resolved effectively.

A simple approach to ensuring the necessary conditions of normal operation is to place the entire responsibility on the operator deploying the autonomous system. However, while relying on human judgment can improve performance, it is desirable to limit human involvement when the conditions of normal operation are violated. Recent work in automated exception recovery has focused on fault diagnosis—detecting and identifying faults—during normal operation. For instance, many approaches diagnose faults by using particle filters or multiple model estimation with neural networks [11]. While these approaches *detect* and *identify* exceptions, they do not offer a way to *handle* exceptions without human assistance. Building on recent work in fault diagnosis, our goal is to develop an exception recovery framework that detects, identifies, and handles exceptions.

We offer an approach to *introspective autonomous systems* in recent work [8]. This system uses belief space metareasoning to recover from exceptions by interleaving a main decision process with exception handlers based on a belief over exceptions that can arise during normal operation. We show that an introspective autonomous vehicle is effective in simulation and on a fully operational prototype.

3 FUTURE WORK

This year, we will make the remaining contributions toward the planning and execution modules of the metareasoning framework below.

Contribution 3. Although our contributions have focused on how to determine when to interrupt an anytime algorithm and act on the current solution, adjusting its internal parameters to optimize performance has not been explored yet. We are developing a metareasoning approach that learns how to adjust its internal parameters by using deep reinforcement learning with a rich state of computation. The state of computation could include features specific to the problem, algorithm, or system. Formally, this approach learns a policy of an MDP with states representing the state of computation and actions representing the internal parameters of the anytime algorithm. We will evaluate our approach by learning how to adjust the weight of Anytime A* and the growth factor of RRT*.

Contribution 5. Ensuring safety is critical to autonomous systems that operate in the real world. We are developing an approach that monitors and controls the execution of an autonomous system to maintain and restore a degree of safety: a set of meta-level monitors address any potential safety problems as a main decision process completes a task. Formally, the main decision process is an MDP that recommends actions that complete a task while each meta-level monitor is an MDP that recommends actions that constrain the actions of the main decision process. The objective of each meta-level monitor is to maximize the probability of remaining in a safe region of the state space while minimizing any interference to the main decision process. We will evaluate our approach in a simulation with an autonomous vehicle that must navigate a route while encountering intermittent safety issues, such as a loss of traction or overheating.

Acknowledgements. This work was published alongside a conference paper [6] and supported by NSF Graduate Research Fellowship DGE-1451512 and NSF grants IIS-1724101 and IIS-1813490.

REFERENCES

[1] Eric A. Hansen and Shlomo Zilberstein, 'Monitoring and control of anytime algorithms: A dynamic programming approach', *JAIR*, (2001).

[2] Nicholas Hay, Stuart Russell, David Tolpin, and Solomon Eyal Shimony, 'Selecting computations: Theory and applications', *arXiv preprint arXiv:1408.2048*, (2014).

[3] Falk Lieder, Dillon Plunkett, Jessica B. Hamrick, Stuart J. Russell, Nicholas J. Hay, and Thomas L. Griffiths, 'Algorithm selection by rational metareasoning as a model of human strategy', in *NIPS*, (2014).

[4] Stuart J. Russell and Eric H. Wefald, *Do the Right thing: Studies in Limited Rationality*, MIT Press, Cambridge, MA, 1991.

[5] Herbert A. Simon, *Models of Bounded Rationality*, MIT Press, Cambridge, MA, 1982.

[6] Justin Svegliato, Samer Nashed, and Shlomo Zilberstein, 'An integrated approach to moral autonomous systems', in *ECAI*, Santiago de Compostela, Spain, (2020).

[7] Justin Svegliato, Prakhar Sharma, and Shlomo Zilberstein, 'A model-free approach to meta-level control of anytime algorithms', in *ICRA*, Paris, France, (2020).

[8] Justin Svegliato, Kyle Hollins Wray, Stefan J Witwicki, Joydeep Biswas, and Shlomo Zilberstein, 'Belief space metareasoning for exception recovery', in *IROS*, Macau, China, (2019).

[9] Justin Svegliato, Kyle Hollins Wray, and Shlomo Zilberstein, 'Meta-level control of anytime algorithms with online performance prediction', in *IJCAI*, Stockholm, Sweden, (2018).

[10] Justin Svegliato and Shlomo Zilberstein, 'Adaptive metareasoning for bounded rational agents', in *IJCAI/ECAI Workshop on AEGAP*, Stockholm, Sweden, (2018).

[11] Vandi Verma, Geoff Gordon, Reid Simmons, and Sebastian Thrun, 'Particle filters for fault diagnosis', *RA Magazine*, (2004).