FREQUENCY DOMAIN CONTROL SYNTHESIS FOR TIME-CRITICAL PLANNING

Jorge E. Tierno* **and Asif Khalak**†

* Alphatech, 6 New England Executive Park, Burlington MA 01803, USA, jtierno@alphatech.com
† Alphatech, 6 New England Executive Park, Burlington MA 01803, USA, akhalak@alphatech.com

Keywords: Model Predictive Control, Planning, Robust Control, Discrete-Event Systems, Markov Decision Processes

Abstract

We establish a connection between two well developed research areas, robust control and planning systems. The discount factors used in Markov Decision Process (MDP) models of time-critical planning are derivable from frequency domain concepts of performance and uncertainty used in robust control. The controls framework makes it possible to move beyond merely justifying the discount factor towards synthesis approaches for designing the right discount factor for a given problem.

1 Introduction and Background

Increasing interest in the development of real-time planning systems has led to a desire to merge ideas from the fields of planning, operations research, and controls engineering. Planning systems are typically modeled as Markov Decision Processes (MDPs), with the plan objective to maximize the expected discounted reward. Such an objective usually leads to computationally intractable dynamic programs, due to the high complexity of the cost-to-go computation (i.e., it is NP-complete).

One approach to address the complexity problem is to emulate *Model Predictive Control* (MPC) as applied in the area of continuous dynamical systems. In the MPC framework, a plan is generated which projects into a significant future time horizon, but only the first components of it are executed. After this, the plan is regenerated using the same technique with the most current available information. Control engineers have developed tools within this MPC framework [2, 7] that allow them to synthesize controllers that achieve the performance objective robustly [10, 4].

Within the literature of Discrete Event Systems[3], techniques have been proposed to address the complexity issues of solving MDP-type planning problems. It has been found that the use of "discount factors" —that is, the practice of exponentially discounting future value— are often beneficial to efficiently finding a good solution. However, applying such discount factors has been justified ad-hoc, based on the intuitive nature of the practice. It has been subsequently shown that the discount factor aids in the convergence of certain fixed-point algorithms for solving the underlying DP[1, 6].

In this paper, we cast a canonical problem in dynamical planning systems into the framework of "classical" dynamics control problems. In this framework, we adapt the tools of robust performance to the canonical dynamical planning problem. This allows for an interpretation of events in the future horizon from a frequency bandwidth perspective. Specifically, we show that the discount factors normally used in MDPs can be interpreted as the performance and robustness bandwidths of the control system. This makes it possible to move beyond justifying the discount factor toward synthesis approaches for designing the right discount factor for a given problem.

The paper is organized as follows. Section 2 formulates a time-critical, dynamic planning system, and discuss how such problems can be cast as an optimal control problem. Section 3 reviews some robust control theory pertaining to a formal definition of robust performance. We analyze the robustness of planning systems for ideal plant dynamics in Section 4, and suggest a strategy for robust planning in Section 5. This control strategy is applied to a routing problem in Section 6. Finally, we make some observations and conclusions in Section 7.

2 Planning System Formulation

Many plans can be characterized by accomplishment of partial objectives which lead to the main objective. If the main objective maximizes "value," then such a characterization accrues value for every partial objective achieved. Often, we would like to capture the *timeliness* of achieving partial objectives—typically, the sooner, the better.

The framework of Markov decision processes (MDPs) [1, 8, 5] is useful for formalizing this discussion of planning systems. To capture timeliness, a *discount factor*, $\gamma \leq 1$, has been used to quantify the rate at which value, $\mathcal{V}$, decays. This suggests a value function of the form

$$\mathcal{V} = \left\langle \sum_{t=0}^{\infty} \gamma^t \; v(t) \right\rangle \tag{1}$$

where $v(t)$ is the value accrued in time step t, and $<>$ represents expected value.

As discussed in the introduction, such time-criticality is found in the MDP literature. Also, on-line planning techniques [6] are designed to replan periodically to exploit new information. Such approaches are analogous to those of Model Predictive Control (MPC), which involves periodic updates to the plan. We propose that a control theoretic view of planning provides unique and useful insights by focusing attention on the *structure* of the uncertainty arising from measurement errors and unmodeled dynamics.

The objective of robust, optimal control design is to find a control policy, c, to minimize an objective function, J, in the presence of disturbance, u. The system diagram which captures the principal elements of the formulation is shown in Figure 1.

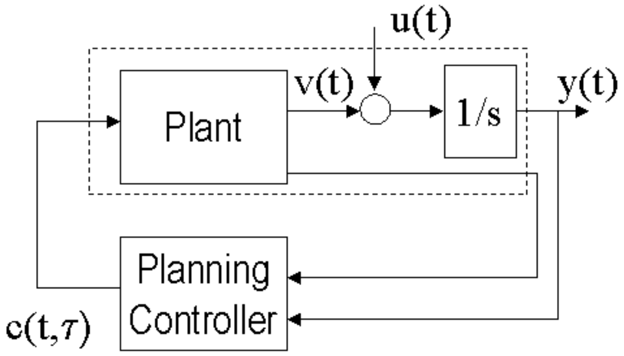

Figure 1: System diagram for planning system

To make the connection from the MDP value maximization problem to an objective function minimization problem we work with potential value, y, rather than captured value, v. The value captured at a given time, (i.e. $v(t)$ of equation (1)), leads to a reduction in the potential value, $y(t)$. Thus, a good plan is one which drives the potential value to zero (i.e. by acquiring it). The change in potential value at a given time, $\dot{y}(t)$, is the reduction in potential value achieved by executing the plan,

$$\dot{y}(t) = u(t) - v(t) \leq 0 \tag{2}$$

where v is planned value capture, and u is the execution error in the value capture. In this context, execution errors, u, model the failure to capture planned value and therefore are bounded by the planned value, v. Thus, performance can be expressed in terms of how the controller drives the output signal to zero remaining value, which is analogous to the standard disturbance rejection formulation of classical control. The associated optimal control problem seeks the optimal control policy, c^* such that

$$c^* = \arg \inf_{c=f(x)} \sup_{\|u\|<1} \int_{-\infty}^{\infty} J(x, c, u)dt \tag{3}$$

where x is the system state.

Solving this dynamic program is typically not feasible, often due to the sheer complexity of x. A common approximation is to solve the problem using *Model Predictive Control* (MPC), which makes some assumptions about disturbances into the future, defining a predictor model. Based on the predictor model, we may define a policy, c, in terms of time, t, rather than state, x. Only the beginning of this solution is used, and then a new optimization problem is solved, taking advantage of new information which has meanwhile become available. Although a closed-loop predictor is theoretically preferable, algorithmic inefficiencies often make an open-loop predictor more practical. This is particularly true when the system contains nonlinearities [9].

Note that the control policy, c, shown in Figure 1 has two time indices, t, and τ. This is a somewhat unique aspect of planning systems, since the control does not simply involve the actions executed time, t, but also those anticipated for future times, τ (i.e. planning time).

The MPC approximation modifies (3) into the following problem

$$\begin{aligned} c(t, \tau) &= c_t^*(\tau) \\ c_t^*(\tau) &= \arg \inf_{c^*(\tau)} \sup_{u \in U} \int_t^{\infty} J(x, c^*, u)d\tau \end{aligned} \tag{4}$$

where U represents the set of allowable disturbances, and c_t^* is a single plan generated at time t. Thus, the problem of control synthesis becomes one of selecting J appropriately.

2.1 Issues in Controller Synthesis

Perhaps the most important benefit to a control theoretic view of dynamic planning is one of perspective. Control theory provides powerful tools for engineers to explicitly account for system performance and model uncertainty. In particular, frequency domain techniques allow us to specify "where" system performance is important and "where" uncertainties are likely to lie (in the frequency domain). Often, performance is important at "low" frequencies (below some threshold of interest), and the uncertainties are present at "high" frequencies. Such an analysis is practical because it can often be done with only a rough understanding system components which ignores the dynamical details.

These considerations generically hold true for time-critical planning systems as well. We will show how the Discounted Markov Decision Process objective (1) can be derived from control theoretic performance notions. To proceed with this, we must conduct a more formal performance and robustness analysis in the planning domain. First, however, we review some of the tools of analysis.

3 Concepts from Robust Control

The theory of robust control provides us with a solid framework to discuss performance of uncertain systems, and to design feedback strategies which achieve this performance robustly. We review some elements of the theory pertaining to *robust performance*. For more comprehensive reviews on the subject, see [10, 4].

3.1 Performance Measures

Many control problems can be cast mathematically as disturbance rejection problems (i.e. keeping some norm of a set of relevant signals small). Given the size of a disturbance signal, the relevant problems are (a) to compute the associated influence on the controlled output (the analysis problem), and (b) to ensure that the influence is kept small enough so that the design requirements are met (the synthesis problem). For linear, time invariant systems the influence is represented by the system induced gain defined as:

$$\sup_{\|u\|<1} \|y\|$$

Where u is the disturbance signal to be "rejected", and y is the output signal we want to isolate from the disturbance.

In general, a control system cannot reject disturbances at all frequencies. Thus, control engineers design a performance filter, h_p, to preferentially reject disturbances in the feedback signal (i.e. the output) within the dynamic range of the plant. Typically, h_p takes the form of a low-pass filter, passing all output below some cutoff frequency.

3.2 Robustness and the Small Gain Theorem

To account for the inevitable incompleteness and uncertainty in system modeling, control engineers often define a range of possibilities for the unmodeled dynamics (called an uncertainty model). Achieving the required performance for all of these possibilities yields a "robust controller". There are several approaches to defining the range of possible models. When the uncertainty is dominated by sensor errors and neglected high frequency dynamics, a multiplicative uncertainty model, as shown in Figure 2, is appropriate for synthesis and analysis. Specifically, the uncertainty is modeled as $W_u\Delta$.

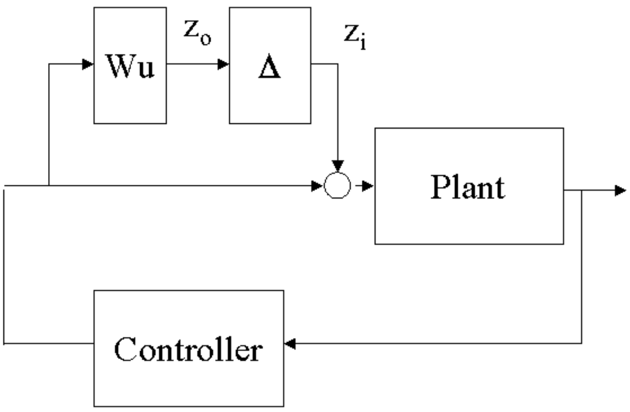

Figure 2: Typical model of a system with multiplicative uncertainty and feedback control.

The W_u block in Figure 2 is a filter on the command signal which defines the frequency content of the uncertainty, Δ. Typically, W_u takes the form of a high-pass filter which is commensurate with the uncertainty associated with sensor errors and high-frequency unmodeled dynamics. The uncertainty, Δ, is bounded, $\|\Delta\| < 1$, but otherwise arbitrary. A well known result in control theory states that the closed loop system will be stable for all (bounded) Δ's if it is stable for $\Delta = 0$ and the following condition on closed loop induced gain:

$$\sup_{\|z_i\|<1} \|z_o\| < 1.$$

This result is usually referred to as the "small-gain" theorem.

3.3 Robust Performance

The control objectives are achieved reliably when the system is stable and performance requirements are met for all values of Δ. This condition is referred to as *robust performance.* A sufficient condition for robust performance is that the system is stable for $\Delta = 0$ and the following induced norm condition is verified

$$\|y\|_\infty < 1 \text{ and } \|z_o\|_\infty < 1 \tag{5}$$

for any bounded inputs, $\|u\| < 1$ and $\|z_i\| < 1$. We will use this test of robust performance in the development that follows.

4 Planning System Robust Performance Analysis

Figure 3 shows the planning system of Figure 1 with filters attached to the output signal, y, and plan signal, c. To satisfy the performance and robustness conditions of equation (5), we minimize the system "gain" from the disturbance signal u to the filtered output signal, y_f, and the filtered plan signal, c_f. For convenience, we choose the ∞-norm to measure the system gain.

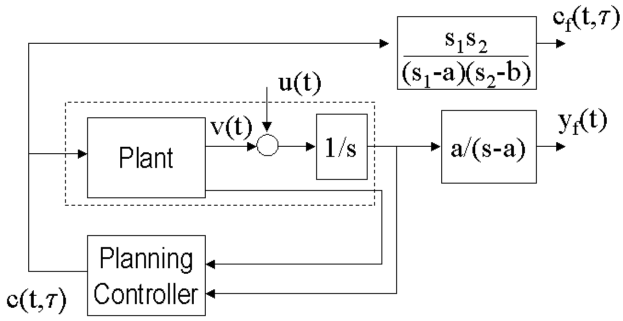

Figure 3: Planning system with filtered outputs

4.1 Performance Filter on the Output Signal

The performance requirements of the system are encoded in the performance filter. The timeliness present in equation (1) indicates that plan quality is not only a matter of how small the potential value, y gets, but how quickly it is driven to zero. From a controls perspective, the performance filter accomplishes this by emphasizing disturbances rejection at frequencies where the filter gain is large and ignoring disturbances where the filter gain is small. The filter used, h_l, is a first order, anti-causal, low pass filter of bandwidth a, which has the following transfer function,

$$h_l = \frac{s}{s-a}.$$

Such a filter is **stable, anti-causal** provided that $a > 0$. Note that we have access to the entire signal y(t), via our predictive model, so an anti-causal filter does not violate physical causality. An anti-causal low pass will prefer value acquisition for short times in the future to value acquisition in the far future (i.e. sooner is better than later). The output of the filter can then be written:

$$y_f(t) = h_l * y = a \int_t^\infty y(\tau)e^{a(t-\tau)}d\tau \tag{6}$$

This is analogous to the discounted value of (1) as found in the MDP literature. Our interpretation of this as a low-pass, anti-causal filter on the output signal is novel, however.

Integrating by parts in (6) produces

$$\begin{aligned} y_f(t) &= -y(\tau)e^{a(t-\tau)}\Big|_t^\infty + \int_t^\infty \dot{y}(\tau)e^{a(t-\tau)}d\tau \\ &= y(t) + \int_t^\infty (u(\tau) - v(\tau))e^{a(t-\tau)}d\tau \end{aligned} \tag{7}$$

If t_0 is the current time, we can derive an upper bound for the filtered signal for times t greater than t_0.

Lemma 1: A filtered signal, y_f, which arises from the action of a stable, anti-causal filter (6) on a signal obeying (2), has the property

$$y_f(t) \leq y_f(t_0) \quad \forall t > t_0. \tag{8}$$

Proof: Differentiating (7) with respect to time, and substituting the definition of $\dot{y}$ from (2) produces

$$\frac{d}{dt}y_f(t) = a\int_t^\infty (u(\tau) - v(\tau))e^{a(t-\tau)}d\tau$$

The first two terms cancel, and the integrand is negative from (2). Thus, the sign of dy_f/dt is negative if $a > 0$, which comes from the stability of filter. ■

4.2 Robustness Filter on the Plan Signal

Recall that the plan signal, c, is a function of two variables, the time t at which the plan is generated, and the time index along the plan, τ. Unmodeled dynamics in the system are excited by two high frequency phenomena: (a) Plan variability, or how fast the plan changes in t for any given τ, (b) Plan complexity, or how much the plan changes in t for any given τ. Robustness requires penalties to both high variability and high complexity, which is accomplished using a two-dimensional, anti-causal, high-pass filter, h_h. This has the following form in the Laplace domain:

$$h_h = \frac{s_1 s_2}{(s_1 - a)(s_2 - b)}$$

where s_1 and s_2 are the Laplace variables associated with the time variables, t, and τ, respectively.

In the MPC framework, the recomputation at time t_0 can introduce discontinuities, which leads to some extra terms in the filtered plan, $c_f = h_h * c$ of the following form:

$$\begin{aligned} c_f(t,\tau) = \; & \int_t^\infty \int_\tau^\infty \Big[\tfrac{\partial}{\partial t}\tfrac{\partial}{\partial \tau}c(\sigma,\epsilon) + \tfrac{\partial}{\partial \tau}\Delta_{t_0}(\tau)\delta(\sigma - t_0) + \\ & \tfrac{\partial}{\partial t}\Delta_{\tau_0}(t)\delta(\epsilon - t_0) + \Delta_{t_0,\tau_0}\delta(\sigma - t_0, \epsilon - t_0)\Big] \\ & e^{a(t-\sigma)}e^{b(\tau-\epsilon)}d\epsilon d\sigma \end{aligned} \tag{9}$$

where the Δ discontinuity factors[1] are,

$$\begin{aligned} \Delta_{t_0}(\tau) &= \begin{cases} c(t_0^+,\tau) - c(t_0^-,\tau), & \tau > t_0 \\ 0, & \text{otherwise} \end{cases} \\ \Delta_{\tau_0}(t) &= \begin{cases} c(t,\tau_0^+) - c(t,\tau_0^-), & t > \tau_0 \\ 0, & \text{otherwise} \end{cases} \\ \Delta_{t_0,\tau_0} &= c(t_0^+,\tau_0^+) - c(t_0^-,\tau_0^-) \end{aligned}$$

[1]Not to be confused with the Δ uncertainty block of Figure 2

For all times τ before t, the plan has already executed. Therefore, $\partial c/\partial \tau = 0$ for $\tau < t$. Also, for all $t > t_0$ our assumption is that the plan is executed as planned at t_0; thus,

$$\frac{\partial}{\partial t}\frac{\partial}{\partial \tau}c(t,\tau) = 0 \qquad \text{for } \{t > t_0, \forall\tau\} \text{ or, } \{\tau < t\}. \tag{10}$$

Lemma 2: A filtered signal, c_f, which arises from the action of a 2D, anti-causal, high-pass filter (9) on a plan signal obeying (10), can be expressed

$$\begin{aligned} c_f(t,\tau) = \; & \int_t^{t_0}\int_t^{t_0} \frac{\partial^2}{\partial t \partial \tau}c(\sigma,\epsilon)e^{a(t-\sigma)}e^{b(\tau-\epsilon)}d\epsilon d\sigma + \\ & b\int_{t_0}^\infty \Delta_{t_0}(\epsilon)e^{a(t-t_0)}e^{b(\tau-\epsilon)}d\epsilon \end{aligned} \tag{11}$$

Proof: Substituting (10) into (9) shows that $c_f = 0$ unless $t \leq t_0$ and $\tau > t$. Elsewhere, we can break up the integral in (9) into two components: for $\tau < t_0$ and for $\tau > t_0$ in the following manner:

$$\begin{aligned} c_f(t,\tau) = & \int_t^{t_0}\int_t^{t_0} \tfrac{\partial^2}{\partial t\partial \tau}c(\sigma,\epsilon)e^{a(t-\sigma)}e^{b(\tau-\epsilon)}d\epsilon d\sigma + \\ & \int_t^{t_0}\int_{t_0}^\infty \Big[\tfrac{\partial}{\partial \tau}\Delta_{t_0}(\epsilon)\delta(\sigma - t_0) + \\ & \Delta_{t_0,\tau_0}\delta(\sigma - t_0, \epsilon - t_0)\Big]e^{a(t-\sigma)}e^{b(\tau-\epsilon)}d\epsilon d\sigma \end{aligned}$$

Applying integration by parts to the second integral, and using the properties of the δ-function we obtain the desired result as follows:

$$\begin{aligned} & \int_t^{t_0}\int_{t_0}^\infty \Big[\tfrac{\partial}{\partial \tau}\Delta_{t_0}(\epsilon)\delta(\sigma - t_0) + \\ & \quad \Delta_{t_0,\tau_0}\delta(\sigma - t_0, \epsilon - t_0)\Big]e^{a(t-\sigma)}e^{b(\tau-\epsilon)}d\epsilon d\sigma \\ & = \int_t^\infty \tfrac{\partial}{\partial \tau}\Delta_{t_0}(\tau)e^{a(t-t_0)}e^{b(\tau-\epsilon)}d\epsilon + \Delta_{t_0,\tau_0}e^{a(t-t_0)}e^{b(\tau-t_0)} \\ & = b\int_{t_0}^\infty \Delta_{t_0}(\epsilon)e^{a(t-t_0)}e^{b(\tau-\epsilon)}d\epsilon \quad ■ \end{aligned}$$

Equation (11) composes the filtered plan as the sum of two terms, the first depending on past decisions and the second depending on future decisions.

5 Controller Synthesis

As stated in Section 2 we seek an approximate real-time solution to the dynamic optimal control problem using the approach of model predictive control. For robust performance, as defined in (5), we develop a plan which minimizes $\|[y_f(t), c_f(t,\tau)]\|_\infty$ where:

$$\|[y_f(t), c_f(t,\tau)]\|_\infty = \max\left(\|y_f(t)\|_\infty, \|c_f(t,\tau)\|_\infty\right)$$

In what follows, we derive formulas to compute online $\|y_f(t)\|_\infty$ and an upper bound on $\|c_f(t,\tau)\|_\infty$.

We consider the performance term, y_f, first, and then the robustness term, c_f. Based on the upper bound derived in Lemma 1, the infinity norm of $y_f(t)$ is minimized when $y_f(t_0)$ is minimized. Substituting the current time, t_0, into equation (7), we obtain

$$y_f(t_0) = y(t_0) + \int_{t_0}^\infty (u(\tau) - v(\tau))e^{a(t_0-\tau)}d\tau$$

Note that the first term, $y(t_0)$, is independent of the plan due to causality of the system dynamics, and that t_0 is a constant. Therefore, $y_f(t_0)$ is minimized when the following expression is minimized

$$\mathcal{J}_{\mathcal{P}} = \int_{t_0}^{\infty} (u(\tau) - v(\tau))e^{-a\tau} d\tau \qquad (12)$$

Next we consider the robustness term. The ∞-norm in the planning space is defined as

$$\|c_f(t,\tau)\|_\infty = \sup_t \sup_\tau |c_f(t,\tau)|.$$

At time t_0 we would like to minimize this norm as a function of the future plan. To make the computation more tractable, we will instead minimize an upper bound, which is a conservative approach. The triangle equality applied to the result in Lemma 2, (11), yields

$$\begin{aligned} |c_f(t,\tau)| < \quad & \int_t^{t_0}\int_t^{t_0} |\frac{\partial}{\partial t}\frac{\partial}{\partial \tau} c(\sigma,\epsilon)| e^{a(t-\sigma)} e^{b(\tau-\epsilon)} d\epsilon d\sigma + \\ & b\int_{t_0}^{\infty} |\Delta_{t_0}(\epsilon)| e^{a(t-t_0)} e^{b(\tau-\epsilon)} \end{aligned}$$

Since the first term only depends on the past plan (again due to causality), it becomes a constant, K, in our optimization, producing

$$\begin{aligned} \|c_f(t,\tau)\|_\infty \quad & < \quad K + \sup_{t<t_0} \sup_{\tau>t_0} \int_{t_0}^{\infty} |\Delta_{t_0}(\epsilon)| e^{a(t-t_0)} e^{b(\tau-\epsilon)} d\epsilon \\ & < \quad K + \sup_{\tau>t_0} \int_{t_0}^{\infty} |\Delta_{t_0}(\epsilon)| e^{b(\tau-\epsilon)} d\epsilon. \end{aligned}$$

where $t = t_0$ gives the greatest possible value in the second term. To implement the maximization at an arbitrary time, t_0, we determine the plan that minimizes this upper bound, i.e., we find the plan for $\tau > t_0$ such that

$$\mathcal{J}_{\mathcal{R}} = \sup_{\tau>t_0} \int_{t_0}^{\infty} |\Delta_{t_0}(\epsilon)| e^{b(\tau-\epsilon)} d\epsilon \qquad (13)$$

is minimized.

The conditions (12) and (13), coupled with the goal of robust performance as defined in (5) provide the appropriate guidance with which to design the controller. In the context of the current planning system domain, we presume that there is an algorithm available to solve the system in (4). The issue is to define the objective function, J to be minimized. We simply take the sum of the performance and robustness terms for this minimization. For an optimization performed at time t_0, this produces the following expression for J:

$$\begin{aligned} \mathcal{J} = \int J_{t_0}(x,c,u) d\tau = \Big\langle \int_{t_0}^{\infty} (u(\tau) - v(\tau)) e^{-\tau/\tau_p} d\tau + \\ \alpha \int_{t_0}^{\infty} |\Delta_{t_0}(\tau)| e^{-\tau/\tau_r} d\tau \Big\rangle \end{aligned} \qquad (14)$$

where α is a tuning parameter and we presume that plant models can provide v and Δ as a function of c and x, and $\langle\rangle$ indicates an expected value.

Remarks: The MPC implementation achieves performance by minimizing the filtered value (12). The low-pass filter encodes a discount factor analogous to that found in the objective of Discounted Markov Decision Processes (MDP) in equation (1). These two can be related by setting $\gamma = e^{-a}$ and $u = 0$. This derivation shows how the discount factor of MDP can be interpreted as a performance bandwidth requirement. Further, the robustness requirement (13), is achieved by minimizing the discounted *change* in plan. Such a robustness discount factor represents the "bandwidth" of the uncertain dynamics.

6 A Vehicle Routing Problem Application

In this section, we apply the concepts from the previous sections to the problem of planning the routes of multiple delivery vehicles. The output of the routing process is an ordered list of activities in time and space for multiple vehicles. There are multiple activity types with uncertainty (e.g. on-time pick-up and delivery), and various constraints handled by the optimization algorithm. The plans are executed in a high-fidelity simulation environment, which accounts for traffic constraints and vehicle dynamics. The planning is done in a dynamic, on-line fashion, analogous to MPC.

As shown in (14), the robustness criterion is highly dependent upon computing the plan difference Δ_{t_0} from the previous plan. In the context of discrete-event systems, this is subject to interpretation since the "plan" may not live in an obvious metric space. The plan difference metric used in the current study is shown in Figure 4. It essentially maps the difference from the previous plan into the net Euclidean distance between associated points.

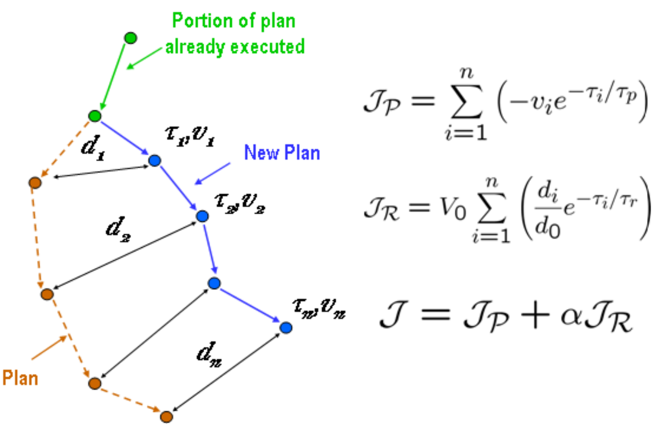

Figure 4: Definition of the planning and robustness components of objective, $\mathcal{J}$.

To assess our controller, we compare our robust plan generator to those which use a discounted value for performance only (but ignore robustness). These are (a) A performance-only "fast" controller with a relatively small timescale τ_p ($\alpha = 0$), (b) A performance-only "slow" controller with a relatively larger timescale τ_p ($\alpha = 0$), and (c) A robust controller ($\alpha \neq 0$) at the "fast" timescale τ_p. We find that there are significant benefits to using a robust planner in the presence of uncertainty.

Figure 5 shows the statistical long-time performance of the three controllers, computed from approximately 100 Monte Carlo trials. The robust controller achieves the best statistical value capture, and is followed by the "slow" controller and then by the "fast" controller. Apparently, the performance-only controller loses efficiency in the face of uncertainty. However, the "slow" controller does better in the long term than the "fast" controller.

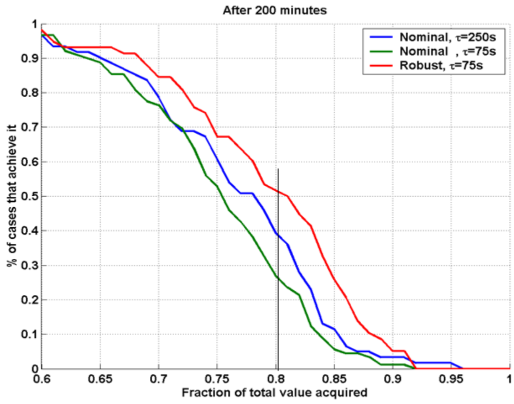

Figure 5: Performance statistics of long term value for the controllers tested: slow(blue), fast(green), and robust(red).

Figure 6 shows the response time of the three controllers. The response time is measured in terms of the lag in seconds from the time a disturbance is detected by the control system to the time that the system executes a plan action to mitigate the disturbance. The response time of the "fast" controller and the robust controller are approximately the same, and both are considerably faster than the "slow" controller. Thus, the better performance achieved by the robust controller in the long-term does not entail a sacrifice in response time.

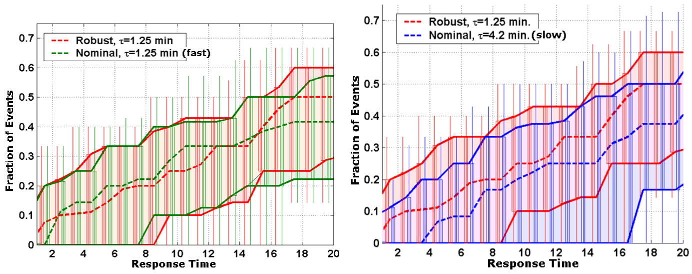

Figure 6: Peformance statistics for short term value for the controllers tested: slow (blue), fast (green), and robust (red). The shaded region shows the spread of results for various trials.

7 Conclusions

In this paper we have established a connection between two well developed research areas, robust control and planning controllers for Markov Decision Processes.

A canonical time-critical planning system is defined in the context of classical continuous-time systems. Some extensions are necessary to capture the fact that the plans are made into the future, using estimates at the current time. The system is set up such that the initial value represents "potential" value of the system, which the planning system attempts to capture by driving the value to zero. This is analogous to classical systems which reject disturbances by driving their output to zero.

Motivated by performance and robustness analysis of classical systems, we apply filters to the output and plan signals. We show that minimizing the infinity norm of the output filtered by an anti-causal, first order, low pass filter is equivalent to minimizing the aggregate discounted future value. (Note that the use of an anti-causal performance filter does not affect the causality of the plant or the control system.)

The robustness condition for classical control involves a penalty on high frequency control actions. The analogous condition for our canonical MDP planning system is a discounted cost on future variations in the plan. The cost is highest for events planned to occur close to the time of planning and lower for events farther in the future.

These concepts have been applied to a dynamic vehicle routing problem with uncertainty and time-criticality. The robust control synthesis strategy described here produces results with good performance in both long term and short term, beyond that achievable without explicitly considering robustness.

Acknowledgements

This research was supported by the Defense Advanced Research Projects Agency (DARPA/IXO) under the MICA program, Contract Number F33615-01-C-3149, Sharon Heise, Program Manager.

References

[1] Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 2. Athena Scientific, 2nd edition, 2001.

[2] E. F. Camacho, Carlos Bordons, and M. Johnson. *Model Predictive Control.* Advanced Textbooks in Control and Signal Processing. Springer Verlag, June 1999.

[3] Christos G. Cassandras and Stephane Lafortune. *Introduction to Discrete Event Systems.* The Kluwer International Series on Descrete Event Dynamic Systems. Kluwer Academic Publishers, 1999.

[4] John C. Doyle, Bruce Francis, and Allen Tannenbaum. *Feedback Control Theory.* Macmillan Publishing Company, 1992.

[5] E. A. Feinberg. Constrained discounted markov decision processes and hamiltonian cycles. *MOR: Mathematics of Operations Research*, 25, 2000.

[6] D. Koller and R. Parr. Policy iteration for factored mdps, 2000.

[7] Jan M. Maciejowski. *Predictive Control with Constraints.* Prentice Hall, 2002.

[8] P. Marbach and J. Tsitsiklis. A neuro-dynamic programming approach to call admission control in integrated service networks: The single link case, 1997.

[9] S.J. Qin and T.A. Badgwell. An overview of nonlinear model predictive control applications. In F. Allgöwer and A. Zheng, editors, *Nonlinear Predictive Control*, pages 369–392. Birkhäuser, 2000.

[10] Kemin Zhou, John C. Doyle, and Keith Glover. *Robust and Optimal Control.* Prentice Hall, 1996.