21st International Symposium on
Mathematical Theory of Networks and Systems
July 7-11, 2014. Groningen, The Netherlands

# Recent Advances in the Numerical Analysis of Optimal Hybrid Control Problems*

Roberto Ferretti[1], Achille Sassi and Hasnaa Zidani[2]

***Abstract*— The application of Dynamic Programming techniques to hybrid systems leads to a Hamilton–Jacobi–Bellman equation in the form of a quasi-variational inequality, whose approximation may be carried out by an adaptation of classical monotone schemes. We present here some further improvements which aim at applying this numerical framework to problems of applicative interest, and in particular the use of policy iteration to speed up the iterative solution, and a practical construction of the (approximate) optimal control in feedback form.**

## I. INTRODUCTION

In recent times, the notion of hybrid system [1], [3], [4], [6], [7] has provided a unified framework for treating a large class of control systems of heterogeneous structure. In the hybrid framework, control may be applied to a dynamical system either in the conventional form, or by choosing among a certain number of different dynamics, or even by performing discrete jumps from one point to another in the state space.

We define therefore a set of dynamics in $\mathbb{R}^d$, indexed by the integer $Q(t) \in \mathcal{I}$, with $\mathcal{I}$ a finite set of integers, so that

$$\begin{cases} \dot{X}(t) = f(X(t), Q(t), u(t)) \\ X(0) = x, \quad Q(0) = q, \end{cases} \tag{1}$$

and two sets $A$ and $C$ such that:

- On hitting $A$ at a time $\tau_i$, the state switches from $(X(\tau_i^-), Q(\tau_i^-))$ to $(X(\tau_i^+), Q(\tau_i^+)) = g(X(\tau_i^-), Q(\tau_i^-), v)$ on the basis of the value of a discrete control $v$, and the cost $C_a(X(\tau_i^-), Q(\tau_i^-), v)$ is associated to this transition.
- When the trajectory evolves in the set $C$, the controller can choose either to switch or not. In the former case, given a switching time $\xi_i$, the state is moved from $(X(\xi_i^-), Q(\xi_i^-))$ to $(X(\xi_i^+), Q(\xi_i^+))$, with a cost $C_c(X(\xi_i^-), Q(\xi_i^-), X(\xi_i^+), Q(\xi_i^+))$.

For such a system, a control strategy is a quadruple

$$\theta := (u(\cdot), v(\cdot), \{\xi_i\}, \{\tau_k\}),$$

composed of:

- A measurable control $u : \mathbb{R}_+ \to U$, mapping the positive time axis into a bounded set of admissible control values $U$;
- A discrete control $v : \mathbb{R}_+ \to \mathcal{V}$ mapping the set of autonomous transition times $\{\tau_k\}$ into a finite discrete set $\mathcal{V}$;
- Two (possibly empty) sets $\{\xi_i\}$ and $\{\tau_k\}$ of positive times, associated respectively to controlled and to autonomous transitions.

In the infinite horizon case, given a control strategy $\theta$, a running cost $\ell$ and a discount factor $\lambda > 0$, the cost functional is then defined as

$$\begin{aligned} J(x,q,\theta) \;&:=\; \int_0^{+\infty} \ell(X(t), Q(t), u(t)) e^{-\lambda t} dt \qquad (2) \\ &+ \sum_{k=0}^{\infty} C_a(X(\tau_k^-), Q(\tau_k^-), v) e^{-\lambda \tau_k} \\ &+ \sum_{i=0}^{\infty} C_c(X(\xi_i^-), Q(\xi_i^-), X(\xi_i^+), Q(\xi_i^+)) e^{-\lambda \xi_i}. \end{aligned}$$

While optimality conditions in the form of a Maximum Principle have also been proposed for hybrid systems (see, e.g., [9]), we will work here in a Dynamic Programming (DP) setting. As customary in DP techniques, we define the value function of the problem,

$$V(x,q) := \inf_\theta J(x,q,\theta), \tag{3}$$

where the infimum is taken over all $\theta$ belonging to the set of strategies defined above. Under suitable, technical assumptions ensuring continuity of the value function, it has been proved [4] that $V$ solves in the viscosity sense a Bellman equation in the form of a quasi-variational inequality on the product space $\mathbb{R}^d \times \mathcal{I}$. We will refer to this framework of hypotheses as “basic assumptions” in what follows.

On the numerical side, two of the authors have shown in a previous work [5] that classical monotone schemes for Hamilton–Jacobi–Bellman equations (e.g., upwind, Lax–Friedrichs, Semi-Lagrangian) can be recast for the case under consideration to provide a convergent approximation of the value function. A possible form of the final scheme is the value iteration

$$\begin{aligned} V_{k+1}^h(x,q) \;&=\; T^h(x,q,V_k^h) \qquad (4) \\ &= \begin{cases} M^h V_k^h(x,q) & x \in A \\ \min\left\{ N^h V_k^h(x,q), \Sigma^h(x,q,V_k^h) \right\} & x \in C \\ \Sigma^h(x,q,V_k^h) & \text{else} \end{cases} \end{aligned}$$

*This work was partially supported by the EU under the 7th Framework Programme Marie Curie Initial Training Network “FP7-PEOPLE-2010-ITN”, GA number 264735-SADCO.

[1]Roberto Ferretti is with Dipartimento di Matematica e Fisica, Università Roma Tre, Roma, Itlay ferretti@mat.uniroma3.it

[2]Achille Sassi and Hasnaa Zidani are with the Unité de Mathématiques Appliquées, ENSTA–ParisTech, Saclay, France ach.sassi@gmail.com, hasnaa.zidani@ensta-paristech.fr

ISBN: 978-90-367-6321-9

where $h$ denotes the discretization step, $\Sigma^h$ is a monotone approximation scheme for the continuous Hamiltonian and $M^h$ and $N^h$ are suitable discretizations of the jump operators (see [5] for details). Denoting by $I[\phi](x, q)$ a monotone space reconstruction of $\phi$ computed at the point $(x, q)$, the approximate jump operators (defined respectively for $(x, q) \in A$ and for $(x, q) \in C$) are in the form

$$M^h\phi(x, q) := \inf_{w\in\mathcal{V}}\{I[\phi](g(x, q, w)) + C_a(x, q, w)\},$$
$$N^h\phi(x, q) := \inf_{(x',q')\in D}\{I[\phi](x', q') + C_c(x, q, x', q')\}.$$

Via the Barles–Souganidis theory [2], it is possible to prove the following result [5]:

*Theorem 1:* Under the basic assumptions, the value iteration (4) converges to a unique fixed point $V^h$ as $k \to \infty$. Moreover, $V^h$ converges to $V$ for $h \to 0$.

However, convergence of value iteration can be very slow, and a first improvement towards an efficient scheme would be to accelerate it, for example by replacing value iteration with *policy iteration*. A second important point for control applications is the synthesis of the optimal feedback.

## II. POLICY ITERATION

In a more abstract form, the value iteration (4) can be written as

$$V^h_{k+1} = \min_\alpha \left(B(\alpha)V^h_k - c(\alpha)\right), \tag{5}$$

where $\alpha$ represents the control strategy as a function of the state (i.e., a *policy*), and $B$ is a linear operator depending on $\alpha$ (note that, depending on the set to which the state $(x, q)$ belongs, $\alpha$ may represent a switching and/or a continuous control).

Slowness of the value iteration (4) is mostly related to the need of performing a large number of minimizations. In fact, at each node of the grid, a minimization is required to compute the numerical Hamiltonian $\Sigma^h$, as well as the switching operators $M^h$ and $N^h$. The policy iteration algorithm reduces the required number of minimizations, thus achieving a speed-up in the computation of $V^h$, by means of the following general idea.

First, an initial policy $\alpha_0$ is chosen. Then, for $k \ge 0$ and until a suitable stopping criterion is satisfied, the algorithm performs alternately the steps of *policy evaluation*, in which $V^h_k$ is defined as the solution of

$$B(\alpha_k)V^h_k = c(\alpha_k),$$

and *policy improvement*, in which a new policy is defined as

$$\alpha_{k+1} = \arg\min_\alpha \left(B(\alpha)V^h_k - c(\alpha)\right).$$

While in conventional control systems the step of policy evaluation requires solving nothing but a linear advection equation, in hybrid systems a more complex definition of the related linear problem is called for (see [8]).

First numerical experiments show that a considerable increase in convergence speed is achieved by this strategy. At the theoretical level, in the hybrid case the algorithm still produces a monotone decreasing sequence of solutions (as in the standard case) and is therefore convergent, although no superlinear convergence result is still available.

## III. FEEDBACK SYNTHESIS

While the use of DP techniques is usually motivated with the possibility of constructing optimal controls in feedback form, very few theoretical results exist on the topic.

A provably convergent strategy to construct feedback controls (see [5]) might be based on a discrete DP process in which the controller works by sampling the state at discrete times $t_k = k\delta$, with a time step $\delta$.

Once the state $(x, q) = (X(t_k), Q(t_k))$ is fixed, an (approximate) optimal control strategy $\theta^\delta_*$ on the time interval $[t_k, t_{k+1})$ is defined by

$$\theta^\delta_* := \arg\min_\theta \left\{\Gamma^\delta(x, q, \theta) + e^{-\lambda\delta}V^h\left(X^\delta(\theta), Q^\delta(\theta)\right)\right\}. \tag{6}$$

In (6), the term $\Gamma^\delta(x, q, \theta)$ represents an approximation of the cost of the strategy $\theta$ on the time interval $[t_k, t_k + 1)$, while $(X^\delta(\theta), Q^\delta(\theta))$ is an approximation of the system's state at $t_{k+1}$ as a result of applying the control strategy $\theta$. Note that this notation includes the case of both a conventional control and a switch or jump in the state. More explicitly, (6) is implemented as follows:

1) If $(x, q) \in A$, then define
$$w^\delta_* := \arg\min_{w\in\mathcal{V}}\{I[V^h](g(x, q, w)) + C_a(x, q, w)\}, \tag{7}$$
and in the time interval $[t_k, t_{k+1})$ perform an autonomous transition to $(x^\delta_*, q^\delta_*) = g(x, q, w^\delta_*)$;
2) If $(x, q) \in (\mathbb{R}^d \times \mathcal{I})\backslash(A \cup C)$, then define
$$u^\delta_* := \arg\min_{u\in U}\{\delta\ell(x, q, u) + e^{-\lambda\delta}V^h(x + \delta f(x, q, u), q)\}, \tag{8}$$
and apply the control $u(t) \equiv u^\delta_*$ for $t \in [t_k, t_{k+1})$;
3) If $(x, q) \in C$, set:
$$c_1 = \min_{u\in U}\{\delta\ell(x, q, u) + e^{-\lambda\delta}V^h(x + \delta f(x, q, u), q)\}$$
$$c_2 = \min_{(x',q')\in D}\{V^h(x', q') + C_c(x, q, x', q')\}.$$
Two subcases are possible:
   a) If $c_1 \ge c_2$, then define
   $$(x^\delta_*, q^\delta_*) := \arg\min_{(x',q')\in D}\{V^h(x', q') + C_c(x, q, x', q')\}$$
   and in the time interval $[t_k, t_{k+1})$ perform a jump to the state $(x^\delta_*, q^\delta_*)$;
   b) If $c1 < c_2$, then define a constant control via (8) and apply it for $t \in [t_k, t_{k+1})$.

Note that in the time interval $[0, \delta)$ this approximate feedback either applies a constant control, or performs a jump, but does not blend the two forms of control.

In order to prove that this strategy of synthesis is quasi-optimal, it is necessary to introduce both a suitable set of assumptions for the existence of an optimal control and a compatibility condition between the two discretization steps $h$ and $\delta$. We assume therefore that

1) The set of admissible controls $U$ is convex;
2) The running cost $\ell(x, q, u)$ is convex w.r.t. the control $u$;
3) The dynamics $f(x, q, u)$ is linear w.r.t. the control $u$, that is,

$$f(x, q, u) = f_1(x, q) + f_2(x, q)u.$$

4) The value function is Hölder continuous with exponent $\gamma > 1/2$.

The following result [5] states the asymptotical optimality of the approximate feedback:

*Theorem 2:* Let the basic assumptions be satisfied, along with assumptions 1–4 above. Assume moreover that $\|V^h - V\|_\infty \to 0$ and that $\delta$ is chosen so as to have $\|V^h - V\|_\infty = o(\delta)$. Then, the control $\theta_*^\delta$ defined by (6) satisfies

$$J(x, q; \theta_*^\delta) \to V(x, q)$$

as $h, \delta \to 0$.

## REFERENCES

[1] G. Barles, S. Dharmatti and M. Ramaswamy, "Unbounded viscosity solutions of hybrid control systems", ESAIM:COCV, vol. 16, pp. 176–193, 2010.
[2] G. Barles and P. Souganidis, "Convergence of approximation schemes for fully nonlinear second order equations", Asymptotic Analysis, vol. 4, pp. 271–283, 1991.
[3] M.S. Branicky, V. Borkar and S. Mitter, "A unified framework for hybrid control problems", IEEE Trans. on Automated Control, vol. 43, pp. 31–45, 1998.
[4] S. Dharmatti and M. Ramaswamy, "Hybrid control system and viscosity solutions", SIAM J. on Control and Optimization, vol. 44, pp. 1259–1288, 2005.
[5] R. Ferretti and H. Zidani, "Monotone numerical schemes and feedback construction for hybrid systems", submitted, available online at `http://hal.archives-ouvertes.fr/`.
[6] R. Goebel, R. G. Sanfelice and A. R. Teel, "Hybrid dynamical systems", Control Systems, IEEE, vol. 29(2), pp. 28–93, 2009.
[7] D. Liberzon, Switching in systems and control, Springer, 2003.
[8] A. Sassi, "Tecniche di Programmazione Dinamica nell'Ottimizzazione di Sistemi di Controllo Ibridi", M.Sc. dissertation, Dipartimento di Matematica, Universtià Roma Tre, Roma, Italy.
[9] H. Sussman, "A maximum principle for hybrid optimal control problems", Decision and Control, 1999. Proceedings of the 38th IEEE Conference on, vol. 1, IEEE, 1999.