Robot planning and control subject to stochastic temporal tasks

Advisor: Christos Verginis, Department of Electrical Engineering, Division of Signals and Systems, UU

Co-Advisor: Dimitris Boskos, Delft Center for Systems and Control, Delft University of Technology

Co-Advisor: Subhrakanti Dey, Department of Electrical Engineering, Division of Signals and Systems, UU

1 Background and project goals

Recent technological advances have enabled the increasing deployment of autonomous robots across all major societal domains, including healthcare, warehouse logistics, agriculture, and search-and-rescue missions. Such robots are engaging in increasingly complex missions that are often safety-critical and need to be performed under fixed deadlines. In order to ensure safety and autonomy, the high-level planning of the robots' tasks needs to be performed in an automated manner and accompanied by guarantees that its execution is correct by design. To this end, computer science techniques, such as formal methods, have emerged as a successful paradigm to express complex temporal tasks in a high-level human-friendly manner. For instance, "robot 1 should be in area A within the next 20 and 30 time units to collect a package from robot 2" or "robot 1 should visit periodically every 5 minutes in areas A and B to report to a human", which are common in warehouse robotics applications. Still, this elementary example may lack absolute guarantees that robot 2 or the human will indeed be in the pre-specified region A during the designated time interval, simply because of random events that may delay their arrival. Namely, there might be uncertainty about the deadlines of the temporal task since they are defined through potential interactions with the environment, such as the uncertain arrival times of the human and robot 2. At the same time, the probability distributions that describe such uncertainty in the deadlines are typically not known and can only be inferred from limited data.

This project focuses on robot planning and control subject to temporal tasks with stochastic deadlines that are modeled via uncertain probability distributions. We will develop algorithms that aim to maximize the probability of robustly satisfying the given task. The key concept of the proposed approach lies in the integration of probabilistic model checking ([2, 11]) and distributionally robust optimization towards the synthesis of robot control policies that hedge against distributional variations of the deadlines in the tasks.

There is considerable work on planning and control with temporal tasks that have prescribed deadlines [10, 3, 15, 17]. However, unlike the tasks that we consider in this project, the time constraints in these works are deterministic and do not encode any uncertainty. Uncertainty in robot planning is typically considered only in the system dynamics, either through non-determinism [17], or through stochasticity [9, 14, 12], but not in the underlying tasks. Closer related to this project is the recent work [13], which considers temporal tasks where specific uncertain time constraints are modeled by random variables. Nevertheless, the authors adopt the limiting assumption that the probability distributions of the random variables are *known*. Since in practice the exact probabilistic model of the random variable is unknown, and may only be approximated from data, this project focuses on stochastic time constraints with *uncertain* probability distributions. To hedge against such uncertainty, distributionally robust optimization (DRO) introduced ambiguity sets of probability distributions that robustify the decisions of optimization problems. The distributions of the ambiguity sets are typically grouped using statistical divergences [5] and optimal transport metrics [16], such as the Wasserstein distance [18]. The latter are widely used for data-driven problems because they are accompanied by statistical guarantees and they facilitate tractability of the DRO problems [6].

2 Methodology

To address the challenges of this proposal, we will combine tools from probabilistic model checking, distributionally robust optimization, and the discretization of stochastic systems. The derivation of control policies for the robot requires a formal description of its underlying task through a mathematical model, which is typically an infinite-state timed automaton [11]. Similarly, the robot's motion is captured through an infinite-state dynamical system with controls that is usually described by a differential equation. Since we seek algorithmically computable solutions, we will first abstract the task automaton and the robot's dynamics into finite-state models, such as Markov Decision Processes (MDP). Subsequently, we will equip these abstractions with suitable ambiguity sets to derive robust control policies. The proposed approach builds on three main objectives: (i) abstraction of the task into a robust MDP, (ii) abstraction of the robot motion model into an interval MDP, and (iii) synthesis of distributionally robust robot policies that maximize the probability of satisfying the given task.

(i) Task abstraction into a distributionally robust MDP: The purpose of this objective is to derive a finite-state MDP representation of the temporal task. We will develop an algorithmic solution that derives such a representation in several steps: First, it converts the task to a stochastic timed automaton (STA), whose accepting runs encode the satisfaction of the task [4]. Second, it extends the STA to a distributionally robust STA (DR-STA), which captures the task's uncertain deadlines via ambiguity sets. Third, it abstracts the DR-STA into a finite-state robust MDP.

A TA is defined over a set $\mathtt{AP}$ of *atomic propositions*, describing properties of interest of the environment, a set $X = \{x_1, \ldots, x_m\}$ of *clocks* taking values on $\mathbb{R}^n_{\geq 0}$, and a set of *guards* $\mathcal{G}(X)$ over X, which consists of clock constraints $x \sim c$, with $x \in X$, $c \in \mathbb{Q}$, and $\sim \in \{<, \leq, \geq, >\}$. *A timed automaton is a tuple* $\mathcal{TA} = (L, X, E, \mathcal{I}, \mathcal{L})$ *where* L *is a finite set of locations,* X *is a finite set of clocks,* $E \subset L \times \mathcal{G}(X)^X \times L$ *is a finite set of edges,* $\mathcal{I} : L \to \mathcal{G}(X)$ *assigns an invariant to each location, and* $\mathcal{L} : L \to 2^{\mathtt{AP}}$ *is a labeling function.* An accepting run of a TA is determined through a sequence of transitions $s_0 \xrightarrow{\tau_1, e_1} s_2 \cdots \xrightarrow{\tau_{n-1}, e_{n-1}} s_n$, where each state $s_k = (\ell_k, v_k)$ comprises a location $\ell_k \in L$ and a valuation $v_k \in \mathbb{R}^m_{\geq 0}$ of the clocks, s_n is an accepting state, and the transitions respect the guard conditions of the state invariants and the edges. The accepting runs of the automaton are in one-to-one correspondence with timed sequences that satisfy the task. An example of a TA is given in Fig. 1. Given a TA and the uncertain task deadlines, the algorithm will construct the respective STA by assigning probability distributions $\mu = (\mu_s)_{s \in L \times \mathbb{R}^X_{\geq 0}}$ over time delays in the locations and a set of weights $q = (q_e)_{e \in E}$ that assign probabilities to each edge.

The second step of the approach is to generalize the notion of an STA to a distributionally robust STA (DR-STA) by allowing the distributions of the time delays, which encode the uncertainty of the task deadlines, to belong to a Wasserstein ambiguity set. This ambiguity set will essentially be a ball in the Wasserstein distance that is centered around a nominal distribution for the time delays.

Finally, we will leverage tools for the abstraction of stochastic systems with uncertain distributions [8] to abstract the DR-STA of the task automaton into a robust MDP (RMDP). To this end, we will generalize the abstraction algorithm of timed automata into finite-state region automata [1]. We will equip the region automaton in our case with a family of transition kernels, which assign to each region an ambiguity set of transition probability distributions. Using the nominal distribution of the time delays, we will construct the center of the ambiguity as the corresponding transition probability distribution of the STA with the nominal distribution. The ambiguity set will then contain all distributions up to an appropriate optimal transport distance from the nominal transition probability distribution.

(ii) Abstraction of robot motion model: The robot's motion is modeled via a controlled stochastic differential equation of the form

$$dX_t = f(t, X_t, u_t) + \sigma(t, X_t, u_t) dB_t,$$

where $X_t \in \mathbb{R}^n$ is the robot's continuous state-space, $u_t \in U \subset \mathbb{R}^m$ is its control, and B_t denotes the r-dimensional Brownian motion. Again, our objective is to derive a finite-state system from the continuous dynamics of the robot.

To obtain the abstraction of the system, we will perform an appropriate discretization Q of the robot's continuous state space, which respects the regions of interest for the underlying temporal task. The abstraction algorithm will be correct-by-design, in the sense that it associates appropriate feedback controllers of the form $u_t(X_t)$ with values in the continuous control space U to actions from a discrete set $\mathcal{A}$ of primitives that achieve safe transitions between regions of the discretized workspace. We will quantify upper and lower probability bounds for these transitions using the continuous-time stochastic model. The gap between these upper and lower bounds is a consequence of the space discretization. As a result, the transition probabilities belong to intervals and the abstracted robot model can be conveniently represented by an Interval Markov Decision Process (IMDP) [7]. An IMDP $\mathcal{M} = (Q, \mathcal{A}, \underline{P}, \bar{P})$ is a special case of a RMDP, where Q is a finite set of states, $\mathcal{A}$ is a finite set of actions, and $\underline{P} : Q \times \mathcal{A} \times Q \to [0, 1]$ and $\bar{P} : Q \times \mathcal{A} \times Q \to [0, 1]$ are upper and lower bounds for the transition probabilities between the states in Q.

(iii) Distributionally robust probabilistic planning: The purpose of this objective is the derivation of control policies that maximize the probability of robustly satisfying the given task. In particular, given the abstract IMDP model of Objective (ii), we will develop an algorithm that computes a strategy $\sigma : \mathrm{Paths_{fin}}(\mathcal{M}) \to \mathcal{A}$, which, for any finite path $s_0 s_1 \cdots s_n \in \mathrm{Paths_{fin}}(\mathcal{M})$ of the IMDP $\mathcal{M}$, picks an action $a \in \mathcal{A}$ in such a way that it maximizes the worst-case probability of satisfying the temporal task.

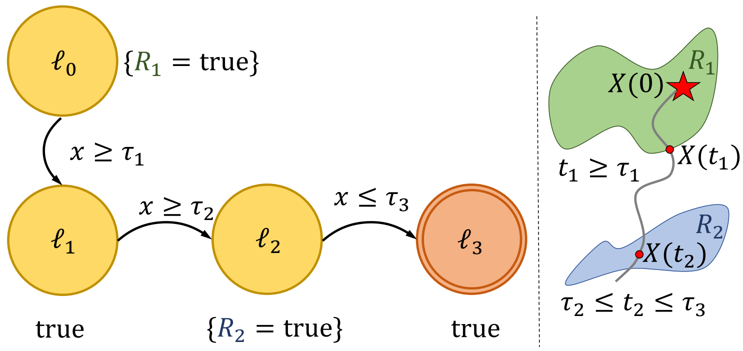

Fig. 1: The timed automaton in the figure describes the following task: "the robot should stay in region R_1 at least until time τ_1 and then it should reach region R_2 sometime between τ_2 and τ_3". Any accepting run of this timed automaton is of the form $(\ell_0, 0) \xrightarrow{t_1} (\ell_1, t_1) \xrightarrow{t_2 - t_1} (\ell_2, t_2) \xrightarrow{0} (\ell_3, t_2)$ with $\tau_1 \leq t_{\leq}\tau_2 \leq t_2 \leq \tau_3$ and clearly satisfies the task.

To synthesize such a strategy, the algorithm will build an appropriate RMDP by computing the product of the task RMDP abstraction and the IMDP abstraction of the robot dynamics. Subsequently, it will iteratively devise optimal strategies using a distributionally robust version of the value-iteration algorithm from dynamic programming. Special focus will be devoted to the computational efficacy of the algorithm, which solves a minimax problem over all the states of the product RMDP. Namely, from each state, it selects the best action to maximize the cumulative probability of satisfying the task against the worst-case ambiguous distribution that determines its successor state.

3 Interdisciplinary aspects and collaborating partners

This project innovatively combines tools from mathematics, computer science, and engineering, and will have a major impact on the deployment of autonomous systems, such as robots and self-driving cars. The developed algorithms will hinge on scientific innovations that combine state-of-the-art tools from formal methods, optimal transport, and stochastic differential equations while drawing motivation from real-world engineering problems.

The department of Electrical Engineering at UU and the Delft Center for Systems and Control offer a very active research environment with internationally recognized experts in most of the aforementioned fields, as well as stimulating course environments to provide the necessary foundations for learning.

Advisors: Christos Verginis received his Ph.D. degree in electrical engineering from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2020. Between October 2020 and January 2022, he held Postdoctoral Researcher with the University of Texas at Austin. Since February 2022, he is an Assistant Professor in intelligent systems at Uppsala University, Uppsala, Sweden. His Ph.D. thesis received the EECI Ph.D. Award on Control for Complex and Heterogeneous Systems in 2020.

Dimitris Boskos received his PhD degree in Applied Mathematics in 2014 from the National Technical University of Athens in Greece. Between August 2014 and July 2018, he held a post-doctoral position at the Division of Decision and Control Systems at KTH Royal Institute of Technology. Between August 2018 and August 2020, he held a post-doctoral position at the Department of Mechanical and Aerospace Engineering at the UC San Diego. Since October 2020, he is an assistant professor at the Delft Center for Systems and Control at TU Delft.

Subhrakanti Dey received the Ph.D. degree from the Australian National University, Canberra, in 1996. He is currently the Head of Division of Signals and Systems and Professor of Signal Processing at Uppsala University. Prior to this, he was with the National University of Ireland, Maynooth, Ireland, during 2018-2022. He was also a Professor with the Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Australia, during 2000-2013. During 1995-2000, he was a Postdoctoral Research Fellow with the Australian National University and University of Maryland, College Park.

4 Financial Support

The PhD student position is supported to 50% by CIM and the remaining 50% by the Department of Electrical Engineering, UU.

References

[1] R. Alur and D. Dill, "A theory of timed automata," *Theoretical Computer Science 126*, vol. 126, pp. 183–235, 1994.
[2] C. Baier and J.-P. Katoen, *Principles of model checking.* MIT press, 2008.
[3] C. Belta, A. Bicchi, M. Egerstedt, E. Frazzoli, E. Klavins, and G. J. Pappas, "Symbolic planning and control of robot motion [grand challenges of robotics]," *IEEE Robotics & Automation Magazine*, vol. 14, no. 1, pp. 61–70, 2007.
[4] N. Bertrand, P. Bouyer, T. Brihaye, Q. Menet, C. Baier, M. Größer, and M. Jurdzinski, "Stochastic timed automata," *arXiv preprint arXiv:1410.2128*, 2014.
[5] G. C. Calafiore and L. E. Ghaoui, "On distributionally robust chance-constrained linear programs," *Journal of Optimization Theory & Applications*, vol. 130, no. 1, pp. 1–22, 2006.
[6] P. M. Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations," *Mathematical Programming*, vol. 171, no. 1-2, pp. 115–166, 2018.
[7] R. Givan, S. Leach, and T. Dean, "Bounded-parameter markov decision processes," *Artificial Intelligence*, vol. 122, no. 1-2, pp. 71–109, 2000.
[8] I. Gracia, D. Boskos, L. Laurenti, and M. M. Jr, "Stochastic timed automata," *arXiv preprint arXiv:2212.14260*, 2022.
[9] R. B. Ilyes, Q. H. Ho, and M. Lahijanian, "Stochastic robustness interval for motion planning with signal temporal logic," *arXiv preprint arXiv:2210.04813*, 2022.
[10] H. Kress-Gazit, G. E. Fainekos, and G. J. Pappas, "Temporal-logic-based reactive mission and motion planning," *IEEE transactions on robotics*, vol. 25, no. 6, pp. 1370–1381, 2009.
[11] M. Kwiatkowska, G. Norman, and D. Parker, "Stochastic model checking," in *International School on Formal Methods for the Design of Computer, Communication and Software Systems*, 2007, pp. 220–270.
[12] B. Lacerda, D. Parker, and N. Hawes, "Optimal and dynamic planning for markov decision processes with co-safe ltl specifications," pp. 1511–1516, 2014.
[13] L. Li and J. Fu, "Policy synthesis for metric interval temporal logic with probabilistic distributions," *arXiv preprint arXiv:2105.04593*, 2021.
[14] R. Luna, M. Lahijanian, M. Moll, and L. E. Kavraki, "Asymptotically optimal stochastic motion planning with temporal goals," pp. 335–352, 2015.
[15] C. N. Mavridis, C. Vrohidis, J. S. Baras, and K. J. Kyriakopoulos, "Robot navigation under mitl constraints using time-dependent vector field based control," *IEEE Conference on Decision and Control (CDC)*, pp. 232–237, 2019.
[16] G. Pflug and D. Wozabal, "Ambiguity in portfolio selection," *Quantitative Finance*, vol. 7, no. 4, pp. 435–442, 2007.
[17] C. K. Verginis and D. V. Dimarogonas, "Timed abstractions for distributed cooperative manipulation," *Autonomous Robots*, vol. 42, no. 4, pp. 781–799, 2018.
[18] C. Villani, *Optimal transport: old and new.* Springer, 2008, vol. 338.