INTERNSHIP proposal

Subject Control with parameter estimation for Hidden Semi Markov Models with application to medical treatment optimization

Location IMAG Montpellier and Inrae Toulouse (France)

Supervision Alice CLEYNEN, CNRS researcher, IMAG, Benoîte de SAPORTA, Professor, Univ. Montpellier, IMAG, Régis SABBADIN Inrae researcher, MIAT

Starting date spring 2022 for 6 month internship

Keywords Bayesian Reinforcement Learning, Hidden observations, Numerical approximation, Optimization, Piecewise Deterministic Markov Processes, Sequential decision making, Simulation

SCIENTIFIC CONTEXT

Context motivation In long-term diseases such as cancer, patients alternate between remission and relapse phases and are monitored along time through non-invasive check-ups such as blood samples. Based on these noisy indirect disease measurements, practitioners must decide on treatment allocation, sometimes with little knowledge on the process dynamics (e.g. level of cancer cells) which may differ between patients.

Mathematical framework In the statistical study of dynamical processes, prediction, diagnosis or decision are made more difficult when the process of interest cannot be directly observed. Typically, the only information available is a noisy and incomplete version of the process. When the latter is governed by latent states, it is classically modeled in the framework of Hidden Markov Models (HMM). In sequential decision problems, the standard question is not to estimate the model parameters or to restore the hidden states, but to compute the best sequence of control actions (a strategy), according to a given objective, in order to manage the system along time. The framework of Partially Observed Markov Decision Process (POMDP) enables to solve such problems when the process itself is ill-observed. The main limitation of the HMM framework and its POMDPs control counterpart is the assumption that the sojourn times of the hidden chain in a particular state follow a geometric/exponential distribution. This assumption is known to be inappropriate in several domains. The frameworks of Hidden Semi Markov Models (HSMM [BL08]), and Partially Observed Semi MDPs (POSMDP [SP14]) have been proposed to relax these assumptions and to enable arbitrarily distributed sojourn times (e.g. Poisson, Gamma). However new advances are required to tackle more challenging real-life problems with complex interleaved dynamics, complex observations (censored, hybrid) and complex management constraints, especially in the context of a learning-while-managing approach where knowledge acquisition and control actions are taken jointly. There exist no efficient tools for solving optimal sequential decision problems when the process to control is both hidden and semi-Markov. Operational algorithms are rare and are often only applicable to small or toy problems, even for the case of a unique hidden chain [SP14] or for the case of Markov models with continuous state space [ZFM10, dSDN16].

INTERNSHIP OBJECTIVES

One of the key challenges in numerically solving a PO(S)MD is to discretize the belief space. The belief or filter is the conditional distribution of the hidden part of the process given the observations. It is itself a stochastic process taking values in a probability measure space, and it strongly depends on the sequence of actions chosen by the controller. The aim of the internship will be to explore and compare different ways of discretizing this space, for fixed control policies. Two directions will be investigated:

- approximations of the semi-Markov kernel in suitably chosen parametric kernel spaces, for example beta or Dirichlet distribution, or Gaussian processes [GMPT15], optimization using machine learning tools,
- sequential discretization of the controlled process state space based on simulations that focus on the exploration of the belief space on its part which is actually reachable by the considered policy through e.g. optimal quantization [CdS18] or Point-based POMDP solvers [SPK13].

Depending on the candidate skills, the internship might also investigate the framework of Model-free Bayesian Reinforcement learning where the underlying PO(S)MD is not inferred but the optimal policy is directly estimated through simulation-based optimization.
All methods will be tested and compared on a cohort data of multiple myeloma patients followed for several years. Discussions with practitioners will be organized.

REQUIRED EXPERTISE OF THE CANDIDATE

Master in probability, statistics or artificial intelligence with a good knowledge of Markov processes and simulation and programming skills. Prior knowledge in stochastic control is not required but would be a plus.

FUNDING AND PHD OPPORTUNITY

Funding for this internship has been acquired through the ANR Hidden Semi Markov Models: Inference, Control and Applications (HSMM-INCA). Grant: 630 €/month
If the candidate is suitable, she or he may have the opportunity to pursue with a PhD (details on the proposed program: https://imag.umontpellier.fr/~saporta/PDF/JOBS/These_en_HSMM.pdf).

REFRENCES

[BL08] V.S. Barbu and N. Limnios. Semi-Markov Chains and Hidden Semi-Markov Models. Toward Applications. Their use in Reliability and DNA Analysis. Vol. 191. Lecture Notes in Statistics. Springer, New York, 2008.
[CdS18] A. Cleynen and B. de Saporta. “Change-point detection for piecewise deterministic Markov processes”. In: Automatica 97 (2018), pp. 234–247.
[dSDN16] B. de Saporta, F. Dufour, and C. Nivot. “Partially observed optimal stopping problem for discrete time Markov processes”. In: A quarterly journal of Operations Research 15 (2016), pp. 277–302.
[GMPT15] M. Ghavamzadeh, S. Mannor, J. Pineau and A. Tamar. “Bayesian reinforcement Learning: a survey”. In: Foundations and Trends in Machine Learning 8.5-6 (2015), pp. 359-492.
[SPK13] G. Shani, J.Pineauand R. Kaplow. A survey of point-based POMDP solvers. *Auton Agent Multi-Agent Syst* **27,** 1–51 (2013). https://doi.org/10.1007/s10458-012-9200-2
[SP14] R. Srinivasan and A. Parlikad. “Semi-Markov decision process with partial information for maintenance decisions”. In: IEEE Transactions on Reliability 63.4 (2014), pp. 891–898.
[ZFM10] E. Zhou, M.C. Fu, and S.I. Marcus. “Solving Continuous-State POMDPs via Density Projection”. In: Automatic Control, IEEE Transactions on 55.5 (2010), pp. 1101–1116.

CANDIDACY

Applications should include a short CV and cover letter and sent before December 15 2021 to
alice.cleynen@umontpellier.fr
benoite.de-saporta@umontpellier.fr
regis.sabbadin@inrae.fr