

# A tractable class of Partially Observed Markovian Decision Process: det-POMDP

- **Cyrille Vessaire** (CERMICS, Ecole des Ponts, Marne-la-Vallée, France, [cyrille.vessaire@enpc.fr](mailto:cyrille.vessaire@enpc.fr))
- Jean-Philippe Chancelier (CERMICS, Ecole des Ponts, Marne-la-Vallée, France, [jean-philippe.chancelier@enpc.fr](mailto:jean-philippe.chancelier@enpc.fr))
- Michel De Lara (CERMICS, Ecole des Ponts, Marne-la-Vallée, France, [michel.delara@enpc.fr](mailto:michel.delara@enpc.fr))
- Pierre Carpentier (UMA, ENSTA Paris, Palaiseau, France, [pierre.carpentier@ensta-paristech.fr](mailto:pierre.carpentier@ensta-paristech.fr))
- Alejandro Rodríguez-Martínez (IAM, TotalEnergies SE, Pau, France, [alejandro.rodriguez-martinez@totalenergies.com](mailto:alejandro.rodriguez-martinez@totalenergies.com))
- Anna Robert (OneTech/R&D, TotalEnergies SE, Saclay, France, [anna.robert@totalenergies.com](mailto:anna.robert@totalenergies.com))

**Keywords:** POMDP, Dynamic programming

**Abstract:** Partially Observed Markovian Decision Processes constitute a class of problems which can be theoretically solved by reformulating them as a much bigger fully observed MDP (Markovian Decision Process). Such reformulation cannot be exactly solved through dynamics programming in the general case. However, when considering deterministic dynamics and observations functions, we obtain a class of POMDP first studied by Littman [1] and called det-POMDP. Det-POMDP is of interest as the size of the reachable state sets used for dynamic programming algorithms are bounded, be it in the finite or infinite horizon cases. The focus of this talk is on finite horizon problems, whereas most of the literature on POMDPs focuses on problems with infinite horizon.

We first present improvements of the bounds presented by Littman. Then, by adding further conditions on the dynamics and observations, we get a sub-class of det-POMDP whose bounds are again improved. Through this subclass, we are able to further push back the dimensionality wall of dynamic programming. We hence obtain partially observed problems that are tractable by dynamic programming algorithms, whereas they are usually solved by approximating the Bellman value functions when considering general POMDPs algorithms.

We illustrate this class of problems on a small problem.

## References:

- [1] M.L. Littman, Algorithms for Sequential Decision Making, *Ph.D. thesis, Brown University*, 1996.