Approximate Policies for Time dependent MDPs

Emmanuel Rachelson [1]

Frédérick Garcia [2]
Patrick Fabiani [1]

[1]ONERA-DCSD — Toulouse, France

[2]INRA-BIA —Toulouse, France

September 22nd, 2007

Illustration

- n_i, passengers in station i
- x_j, current station of train j
- y_j, is train j working properly ?
- z_j, passengers on-board train j
- t_j, train j's starting time

$\longrightarrow$ time-dependent dynamics

$\longrightarrow$ non-controlable events

Optimize network exploitation cost

Probabilistic Temporal Planning

Discrete Time:

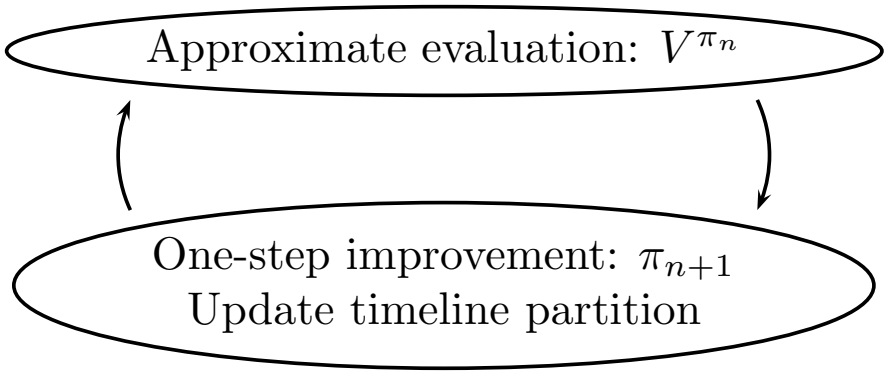
- CoMDP (concurrent actions)

Continuous Time:

- CTMDPs and SMDPs (stationary problems)
- TMDP (Boyan & Littman, 01)
- GSMDP (Younes & Simmons, 04)
- Continuous resources (SSP algorithms, HAO*, ALP, CPH, . . .)
- “Classical” planning approaches (Prottle, IxTeT, . . .)

Our research focus

We investigate *Approximate Temporal Policy Iteration*



$\longrightarrow$ family of algorithms for temporal policy search

Algorithms

- ATPI with TMDPpoly approximation
  - $\longrightarrow$ Idea:
    - piecewise polynomial approx. for SSP-like problems
  - $\longrightarrow$ Poster
- Simulation-based ATPI
  - $\longrightarrow$ Idea:
    - Heuristic search for large state spaces, non-SSP problems using policy iteration
  - $\longrightarrow$ Issues:
    - representing π (timeline partition estimator + BDD)
    - Convergence of API
    - Simulation framework (DEVS)

Extensions

- UAV - robot coordination
- Satellite operations planning

Extensions

- UAV - robot coordination
- Satellite operations planning