ONLINE PLANNING FOR F1 RACE STRATEGY IDENTIFICATION

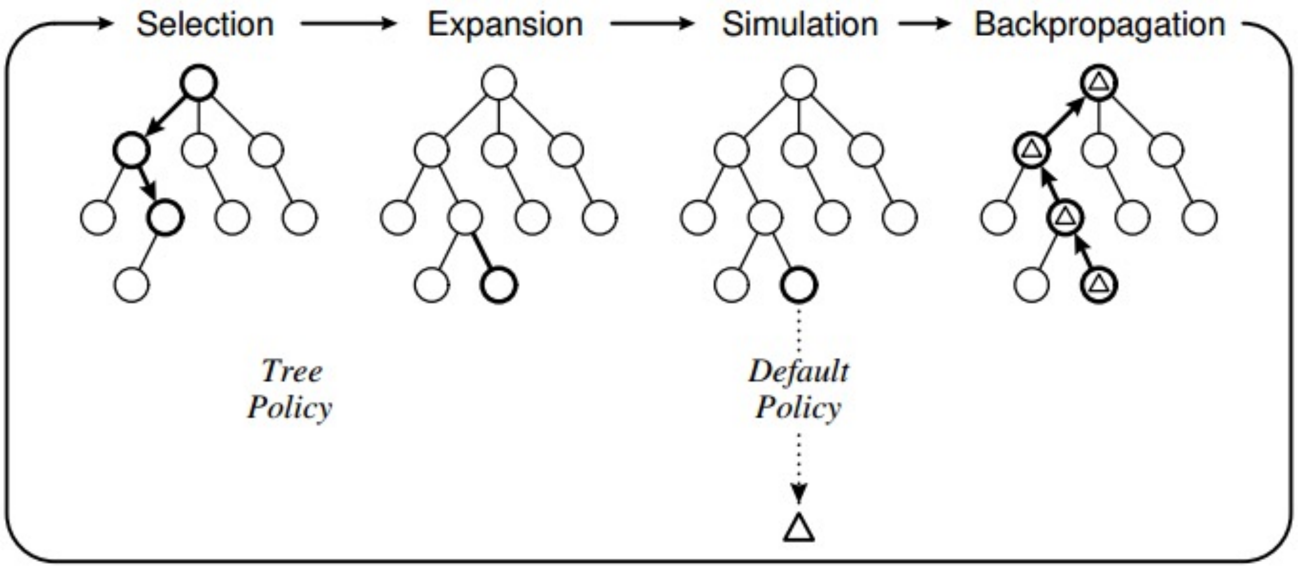
DIEGO PICCINOTTI
AMARILDO LIKMETA, NICOLÒ BRUNELLO, MARCELLO RESTELLI

diego.piccinotti@(mlcube.com, mail.polimi.it)

POLITECNICO
MILANO 1863

THEORETICAL BACKGROUND: MCTS ALGORITHMS

POLITECNICO MILANO 1863



STATE OF THE ART: PLANNING ALGORITHMS

POLITECNICO MILANO 1863

DPW-UCT[1]

Limit branching factor by using an upper bound to child states number.

OLOP[2]

Search over sequences of actions instead of sequences of states.

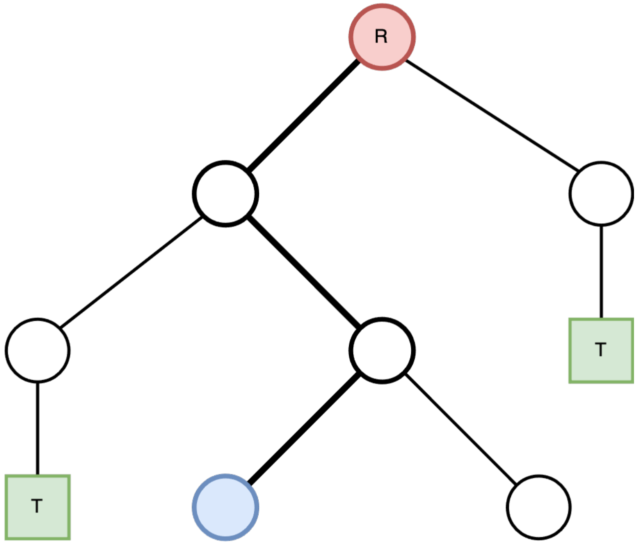
OL-UCT[3]

Identify nodes in the search tree by sequences of actions instead of environment state.

[1] Couetoux and Doghmen, *Adding Double Progressive Widening to Upper Confidence Trees to Cope with Uncertainty in Planning Problems*
[2] Bubeck and Munos, *Open-Loop Optimistic Planning*
[3] Lecarpentier et al., *Open Loop Execution of Tree-Search Algorithms*

STATE OF THE ART: PLANNING ALGORITHMS

POLITECNICO MILANO 1863

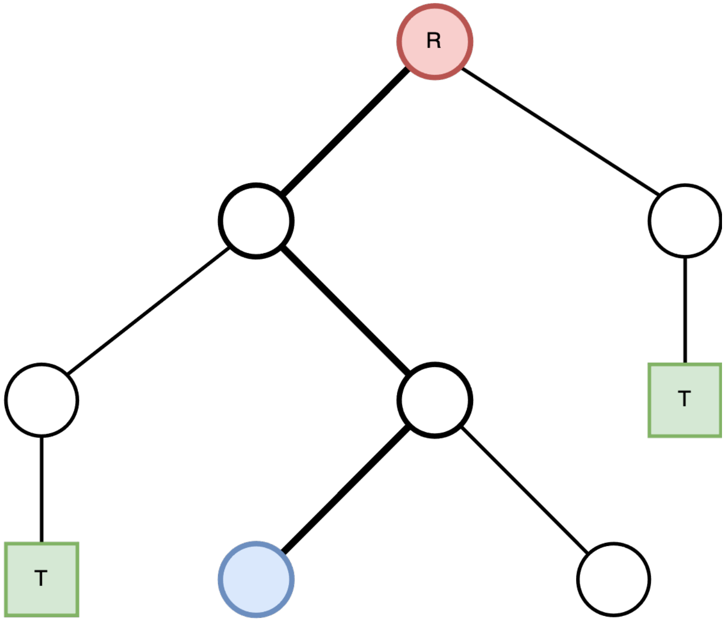


OL-UCT

Identify nodes in the search tree by sequences of actions instead of environment state.

$$V_{OL}(s,\tau) = \mathbb{E}\left[\sum_{t=0}^{m} \gamma^t r_t \mid s_0 = s, a_t \in \tau\right]$$

THEORETICAL BACKGROUND: OPEN-LOOP



$$V_{OL}(d|s_0,\tau) = \mathbb{E}_{(s_d \sim P(\cdot|s_0,\tau[1:d]))}\left[\sum_{t=0}^{m-d} \gamma^t r_{t+d} \,\middle|\, s_d, \tau[d:m]\right]$$

$$V_{OL}(s,\tau) = \mathbb{E}\left[\sum_{t=0}^{m} \gamma^t r_t \,\middle|\, s_0 = s, a_t \in \tau\right]$$

1

FORMULA 1 RACE STRATEGY

Introduction to basic concepts

POLITECNICO
MILANO 1863

USE CASE GOAL

Identify **pit-stop strategies** for Formula 1 competitions, using a **planning-based autonomous agent**

FORMULA 1: BASIC CONCEPTS

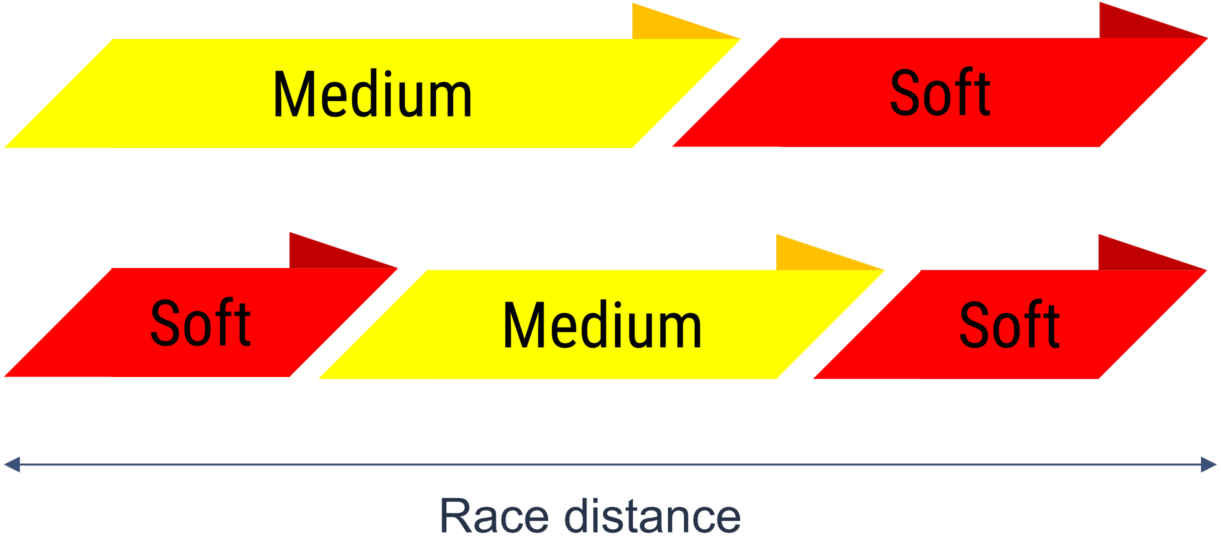
Formula 1 is the top category in motorsport, in which single-seater high-performance vehicles compete around a closed circuit.

Points are awarded to the drivers completing the prescribed amount of laps in the top 10 positions. Points obtained in race events cumulate through the season, determining both the driver and teams ranking.

The goal of a driver is therefore to consistently finish races in the best placement achievable.

EXAMPLE OF POSSIBLE TIRE STRATEGIES

POLITECNICO MILANO 1863



- Soft tyre compound
- Medium tyre compound

RACE STRATEGY

Allowing to gain time against competitors, good race strategy is critical in close fight races!

POLITECNICO
MILANO 1863

2

PROBLEM MODELING

Markov Decision Process model

POLITECNICO
MILANO 1863

MARKOV DECISION PROBLEM MODEL

We model a single-agent MDP

- **State**: features for each driver + flags
- **Actions**: stay on track, pit for compound
- **Reward**: negative (normalized) lap time
- **Transition model**: lap time simulator
- **Discount factor**: 1

MAIN DIFFICULTIES

POLITECNICO
MILANO 1863

- Continuous state space
- High stochasticity
- Reward difference between actions: pit-stops cost around 30s more than staying on track
- Good policies are unbalanced (typically 2-3 pit-stops per race)

3

SOLUTION

Race simulator and proposed algorithm

POLITECNICO
MILANO 1863

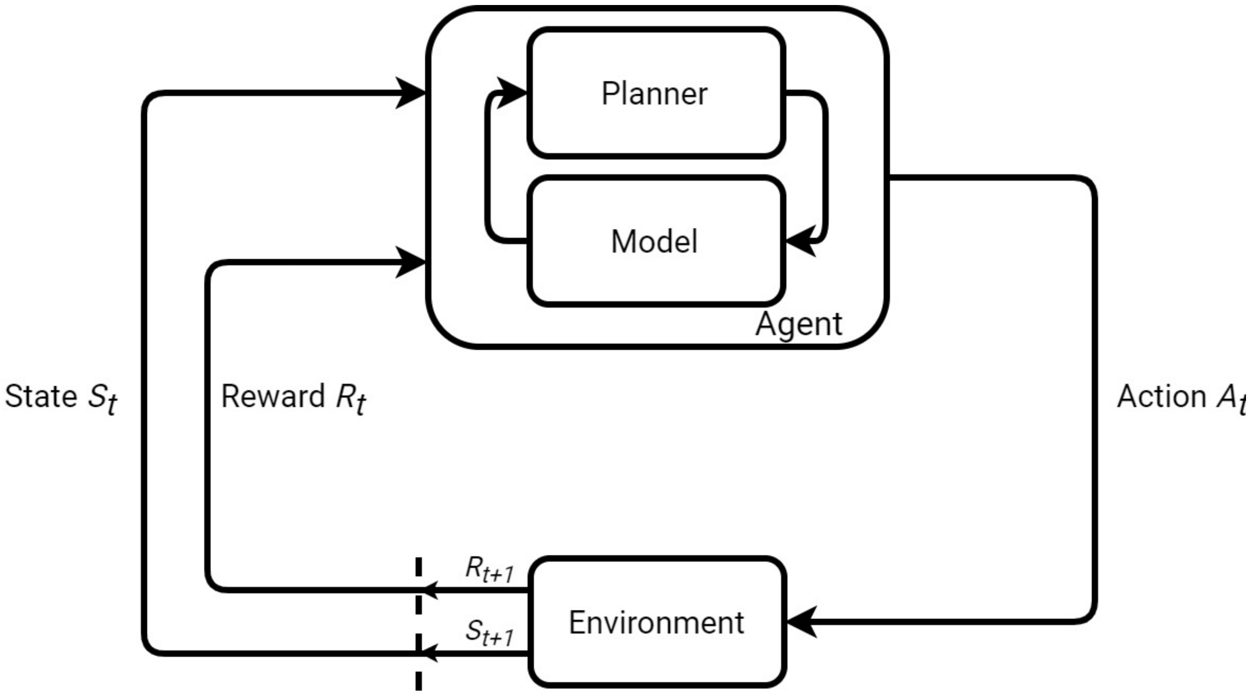
ELEMENTS ADDRESSED BY SOLUTION

- Limit state space
- Tackle stochasticity

→ Open-Loop approach

- Reduce variance in returns → TD update
- Use domain knowledge → ESPN rollout

PLANNING GOAL



PLANNING GOAL

$$Q^*(s,a) = \max_{\pi} \left[R(s,a) + \gamma \sum_{s' \in S} P(s'|s,a) \sum_{a' \in A} \pi(a'|s') Q^{\pi}(s',a') \right]$$

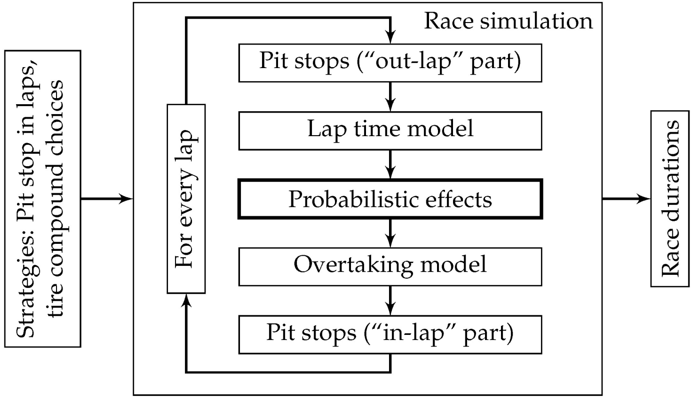
Bellman Expectation Equation

for State Action Value Function

RACE SIMULATOR - PROBABILISTIC

We use a probabilistic lap-time simulator[1] as transition model

- Lap time as sum of contributes
- Probabilistic approach
- **Extension**: specify actions for drivers lap-by-lap
- **Extension**: dynamically add Safety Car during simulation



[1] A. Heilmeier et al., *Application of Monte Carlo Methods to Consider Probabilistic Effects in a Race Simulation for Circuit Motorsport*

POSSIBLE ROLLOUT STRATEGIES

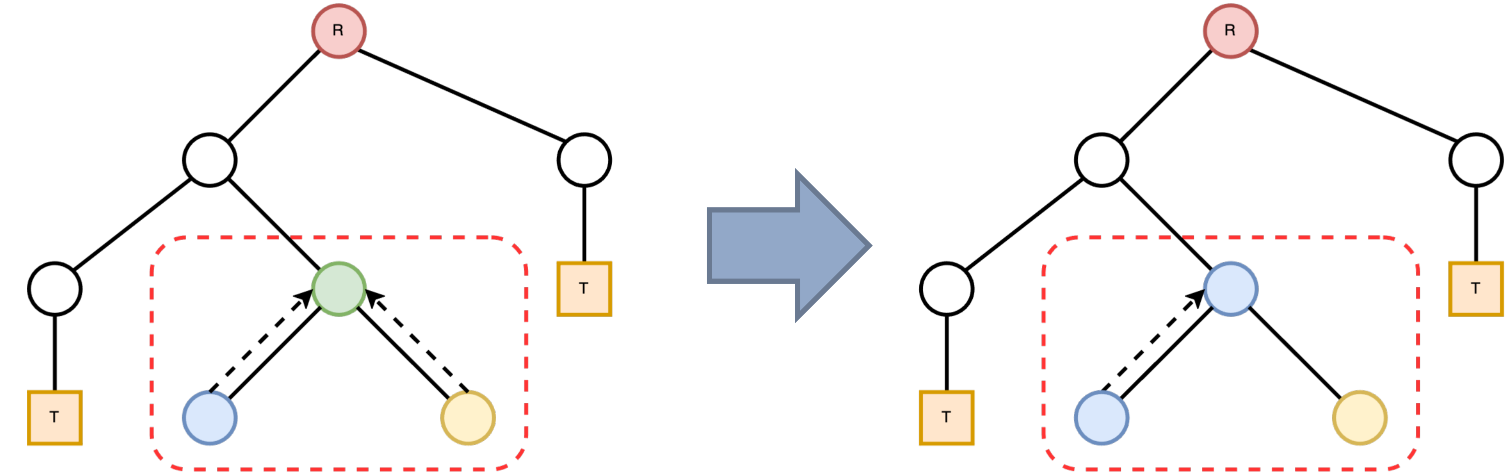
POLITECNICO
MILANO 1863

- Uniform (default in UCT)
- Mostly stay on track
- Rollout according to ESPN[1] predicted strategies

We deploy the **ESPN** rollout in our algorithm.

[1] See, for instance, https://www.espn.com/f1/story/_/id/27087391/austrian-grand-prix-strategy-guide

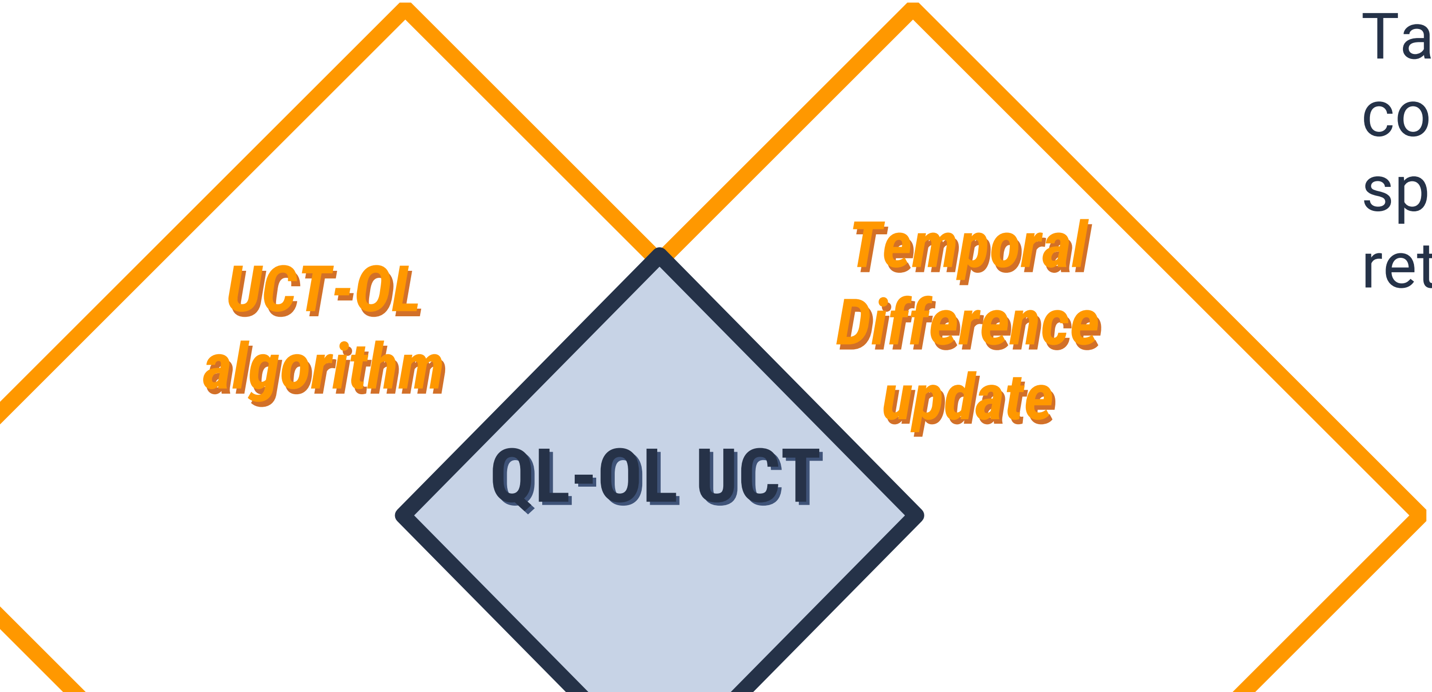
MONTE-CARLO vs Q-LEARNING



$$\mathcal{Q}(\mathcal{N},a) \leftarrow \mathcal{Q}(\mathcal{N},a) + \frac{\sum_{t=0}^{m-d} \gamma^t r_t - Q(\mathcal{N},a)}{\mathcal{N}.n}$$

$$\mathcal{Q}(\mathcal{N},a) \leftarrow (1-\alpha)\mathcal{Q}(\mathcal{N},a) + \alpha(r_d + \gamma \max_{C(\mathcal{N})} \mathcal{Q}(\mathcal{N}',a))$$

QL-OL UCT



Tackling continuous state space and high return variance

4

EXPERIMENTAL EVALUATION

Let’s race!

POLITECNICO
MILANO 1863

EXPERIMENTAL SETTING

POLITECNICO
MILANO 1863

- 2014 – 2019 races
- Focus on Vettel at Ferrari (2015 - 2018)
- Select races with small gap from competitors (~10s)
- Test against real race strategy, ESPN, and a selection of planners from literature
- 100 experiments for each of 10 races

Best UCT planner

80% races

Best overall

40% races

Improved true strategy

70% races

RESULTS ON TEST RACES - I

POLITECNICO MILANO 1863

	ESPN	True	VSE[1]	Sarsa UCT[2]	Power UCT[3]	OL UCT[4]	QL-OL UCT
Japan 2015	+5.66	+7.17	+4.99	+5.01	+12.9	+7.64	**4570.35 ±1.0 ***
Japan 2016	+2.5	+2.19	+43.66	+3.55	+19.1	+13.68	**4505.35± 0.9 ***
Australia 2017	+10.68	+6.51	+18.12	+6.85	+14.41	+20.19	**4459.71± 2.4 ***
Spain 2017	+14.37	+21.84	+19.89	+8.33	+12.78	+23	**5188.05± 1.3 ***
Austria 2017	+95.04	**4430.84 ±1.7 ***	+60.82	+45.59	**+13.54**	+53.3	+35.01

- Best UCT planner
- Best overall planner

[1] Heilmeier et al., *Virtual Strategy Engineer: Using Artificial Neural Networks for Making Race Strategy Decisions in Circuit Motorsport*
[2] Vodopivec et al., *On Monte Carlo Tree Search and Reinforcement Learning*
[3] Dam et al., *Generalized Mean Estimation in Monte-Carlo Tree Search*
[4] Lecarpentier et al., *Open Loop Execution of Tree-Search Algorithms*

RESULTS ON TEST RACES - II

POLITECNICO MILANO 1863

	ESPN	True	VSE[1]	Sarsa UCT[2]	Power UCT[3]	OL UCT[4]	**QL-OL UCT**
Belgium 2017	+29.4	+20.44	**4236.00 ±0.6 ***	+19.98	+23.52	+24.24	**+10.09**
Russia 2017	+7.11	**4412.87 ±1.2 ***	+15.75	+12.23	+24.13	+18.06	**+8.67**
China 2018	+45.36	+38.67	**5095.34 ±2.0 ***	+3.99	+33.29	+17.97	+3.59
Italy 2018	+10.99	**3898.38 ±1.9 ***	+45.57	+8.84	+20.04	+12.86	**+5.29**
Brazil 2018	**4678.24 ±2.1 ***	+22.12	+33.37	+14.46	+21.01	+28.7	**+8.08**

- Best UCT planner
- Best overall planner

[1] Heilmeier et al., *Virtual Strategy Engineer: Using Artificial Neural Networks for Making Race Strategy Decisions in Circuit Motorsport*
[2] Vodopivec et al., *On Monte Carlo Tree Search and Reinforcement Learning*
[3] Dam et al., *Generalized Mean Estimation in Monte-Carlo Tree Search*
[4] Lecarpentier et al., *Open Loop Execution of Tree-Search Algorithms*

5

CONCLUSIONS AND FUTURE WORK

Takeaway and improvement points

POLITECNICO MILANO 1863

CONCLUSIONS

POLITECNICO MILANO 1863

- Open Loop to tackle continuous and stochastic state transition
- TD update to address high variance
- Improved performance of real strategies
- Performance dependent on rollout strategy

POSSIBLE EXTENSIONS

- AlphaZero setting
- Multi-agent modeling
- Reward considering ranking gains
- Improve lap time simulation

POLITECNICO
MILANO 1863

THANKS!

Any questions?

You can find me at

diego.piccinotti@(mlcube.com, mail.polimi.it)