International Planning Competition 2018 Probabilistic Tracks

Thomas Keller

University of Basel, Switzerland

June 29, 2018

New at IPC 2018

New at IPC 2018

new RDDL features

- action preconditions
- enum-valued variables (compilation provided)
- intermediate variables (compilation provided)

New at IPC 2018

new RDDL features

- action preconditions
- enum-valued variables (compilation provided)
- intermediate variables (compilation provided)

new participation procedure

- planner submission as in classical tracks
- planner abstracts mandatory

New at IPC 2018

new tracks

- discrete MDP track
- discrete data-based track
- discrete SSP track
- continuous tracks

New at IPC 2018

new tracks

- discrete MDP track
- discrete data-based track
  canceled due to lack of interest
- discrete SSP track
  raised a lot of interest, but canceled due to lack of participants
- continuous tracks
  postponed to 2019 (organized by Scott Sanner)

Participants

Participants

Statistics

- 5 planners (+2 withdrawn in the last week)
- 4 teams (+1 withdrawn in the last week)
- two variants of each planner allowed

Participants

A2C-Plan

- Anurag Koul, Murugeswari Issakkimuthu, Alan Fern and Prasad Tadepalli
- Oregon State University, USA
- reinforcement learning trains neural network with A2C algorithm

Participants

Imitation-Net

- Murugeswari Issakkimuthu, Alan Fern and Prasad Tadepalli
- Oregon State University, USA
- supervised learning trains neural network with one-step lookahead policy

Participants

PROST-DD-1 and PROST-DD-2

- Florian Geißer and David Speck
- University of Freiburg, Germany
- THTS planner with heuristic based on decision diagrams

Participants

Random-Bandit

- Alan Fern, Murugeswari Issakkimuthu and Prasad Tadepalli
- Oregon State University, USA
- MCTS with ϵ-greedy exploration in the root and greedy behaviour elsewhere

Participants

Conformant-SOGBOFA-F-IPC18 and
Conformant-SOGBOFA-B-IPC18

- Hao Cui and Roni Khardon
- Tufts University, USA
- gradient-based search on symbolic task representation, optimized for large-scale problems in a fractional (Conformant-SOGBOFA-F-IPC18) and binary (Conformant-SOGBOFA-B-IPC18) version

Domains

New domains

Competition Domains

- 7 new domains
- 1 domain from previous IPC
- 20 new instances per domain

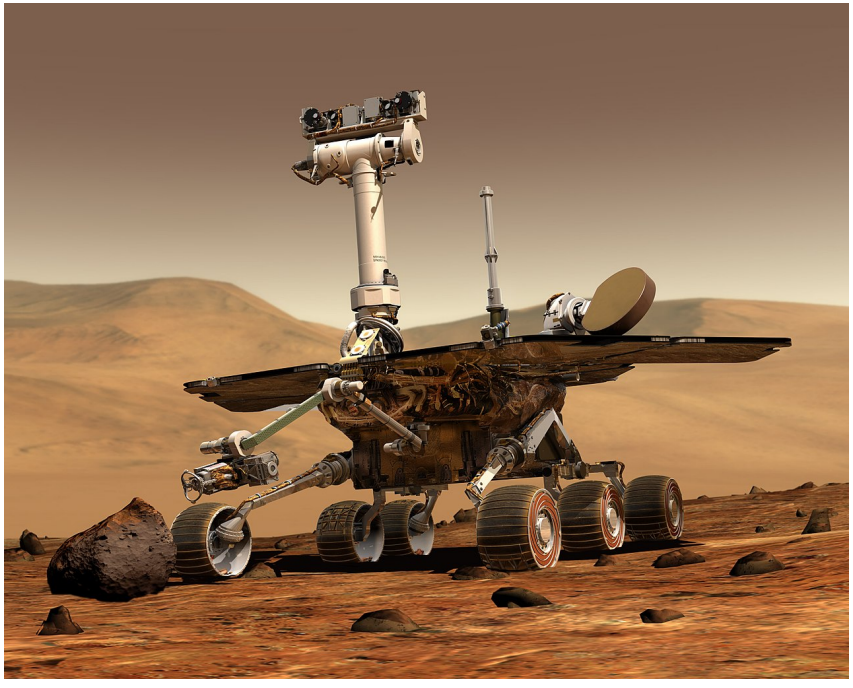
Academic Advising

- student aims to graduate
- probability of passing course higher if prerequisites passed
- requires to plan far ahead

Chromatic Dice

- version of dice game Yahtzee
- dice roll gives value and color
- highest immediate reward often different from optimal policy

Cooperative Recon



Picture by NASA/JPL/Cornell University

- rover searches for life on Mars
- collaboration required for higher probability of success
- challenging form of concurrency

Earth Observation

- satellite takes images of patches on Earth
- use weather forecast to optimize probability of high-quality images
- weather forecast analysis crucial for success

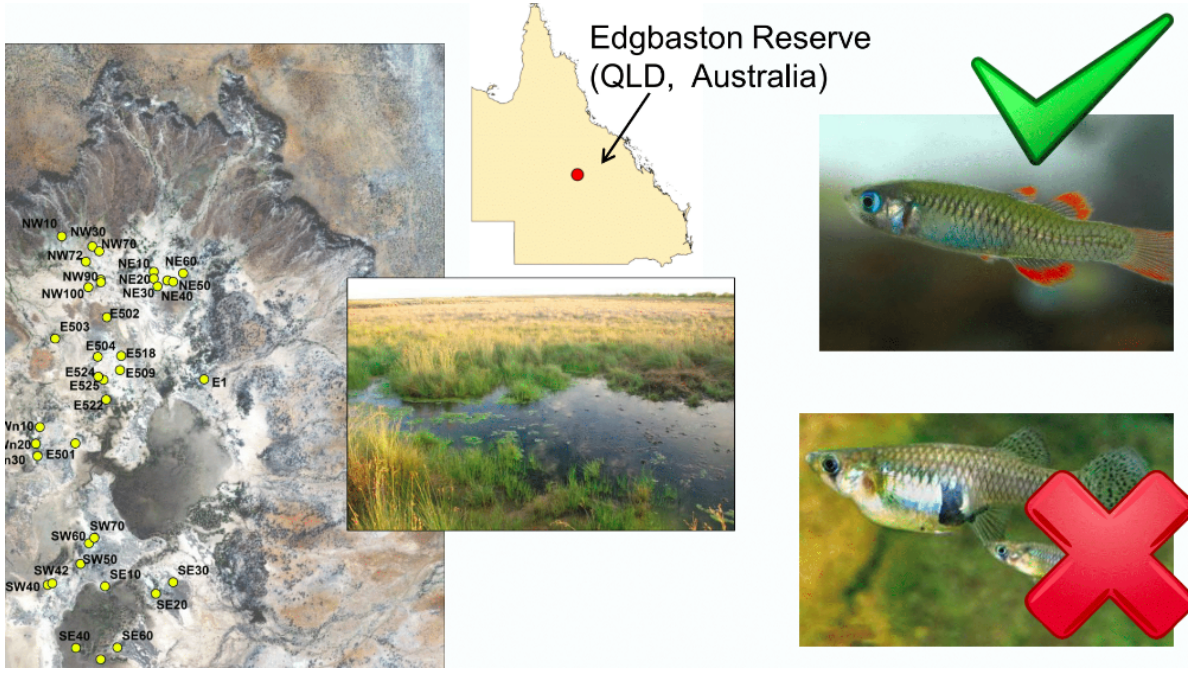
Manufacturer

- buy, produce and sell goods
- influence prices and automate production by hiring staff
- accept negative short-term reward for increased long-term reward

Push Your Luck

- roll dice until choose to cash out or value shows up again
- reward is product of all rolled values
- punishes determinization

Red-finned Blue-eye



Picture by Iadine Chadès

- Red-finned Blue-eye population threatened by Gambusia
- springs connected probabilistically during rain season
- find strategy to save Red-finned Blue-eye from extinction
- challenging due to highly entwined probabilities

Wildlife Preserve

- protect wildlife from poaching
- respond to dynamic attacker strategy
- attacker strategy complex, large action space

Results

Experimental setup

changes in experimental setup

- more instances per domain
- variable horizon
- 75 runs
- significantly more time per instance
  (between $\approx$ 1 hour and $\approx$ 4.5 hours)
- central evaluation on grid
- baseline planners: PROST 2011 and PROST 2014

Experimental setup

- execute policy by interacting with rddlsim
- obtain average reward over 75 runs
- compute instance score by comparing to artificial min policy and highest score among the participants
- domain score is sum over all instance scores

Awards

Winner

Florian Geißer and David Speck with
PROST-DD

Runner-Ups

Murugeswari Issakkimuthu, Alan Fern and Prasad Tadepalli with
Random-Bandit and

Hao Cui and Roni Khardon with
Conformant-SOGBOFA

Results

Score	Academic Advising	Chromatic Dice	Cooperative Recon	Earth Observation	Manufacturer	Push your Luck	Red-finned Blue-eye	Wildlife Preserve	SUM
PROST 2014	3.3	10.1	10.7	**19.9**	2.7	14.2	6.0	7.9	74.7
PROST-DD-2	5.8	7.6	10.3	6.5	3.3	**15.0**	5.9	**14.3**	68.8
PROST 2011	3.2	12.8	9.0	18.7	**7.1**	6.3	6.9	3.9	67.9
PROST-DD-1	**6.6**	7.5	**12.0**	5.3	2.8	12.7	5.4	**14.3**	66.5
Random-Bandit	0.7	17.1	1.5	12.8	4.1	13.1	5.6	10.8	65.6
Conf.-SOGBOFA-B	4.1	**19.4**	6.9	7.4	0.0	1.4	18.3	4.8	62.3
Conf.-SOGBOFA-F	4.9	18.9	6.4	7.1	0.0	1.3	**18.7**	4.8	62.1
Imitation-Net	0.0	3.8	0.0	0.6	0.3	8.8	5.0	10.1	28.6
A2CPlan	1.4	0.6	4.8	1.6	2.7	6.9	4.8	3.8	26.6

Thanks

- Iadine Chadès and Guillem Francès for their help with Red-finned Blue-eye
- Fei Fang and Thanh Nguyen for their help with Wildlife Preserve
- Augusto Blaas Corrêa
- Florian Pommerening

Thank You!