Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

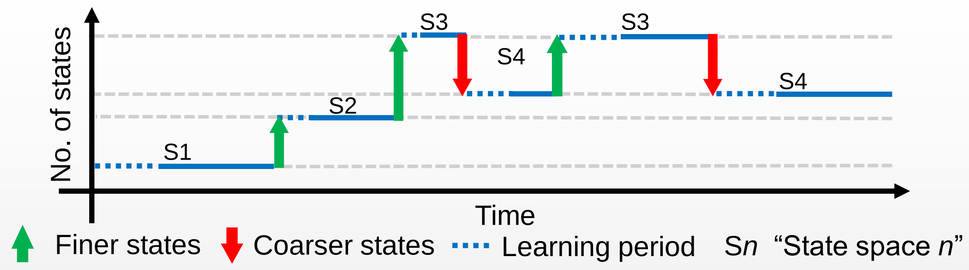
maxime.gueriau@scss.tcd.ie @maximegueriau
n.cardozo@uniandes.edu.co @ncardoz
ivana.dusparic@scss.tcd.ie @ivanadusparic

Constructivist Approach to State Space Adaptation in Reinforcement Learning

Maxime GUÉRIAU[1], Nicolás CARDOZO[2] and Ivana DUSPARIC[1]

[1]*School of Computer Science and Statistics, Trinity College Dublin, The University of Dublin, Ireland*
[2]*Systems and Computing Engineering Department, Universidad de los Andes, Colombia*

1. Context: State space adaptation in Reinforcement Learning

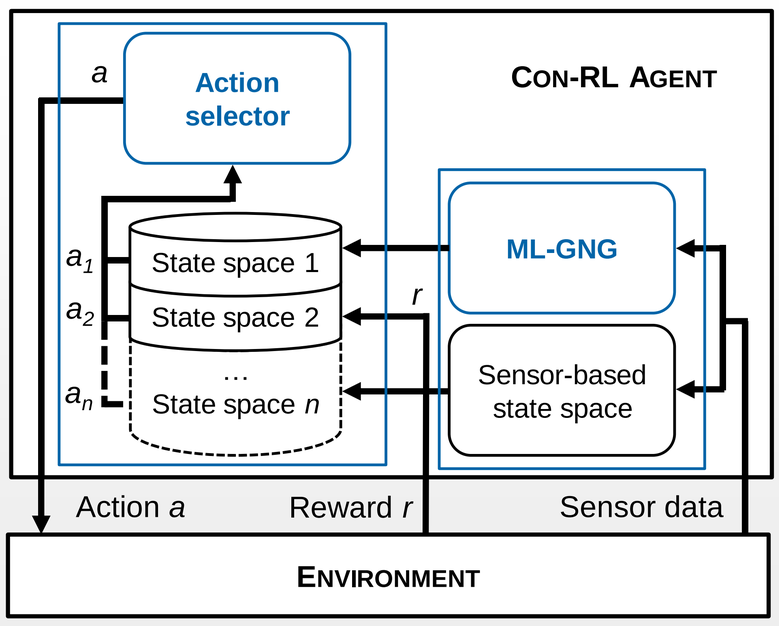


- When does an agent need to adapt its state space?
  - when its original state space is too big/small
  - when sensors are added or removed dynamically
  - when sensors input granularity changes over time

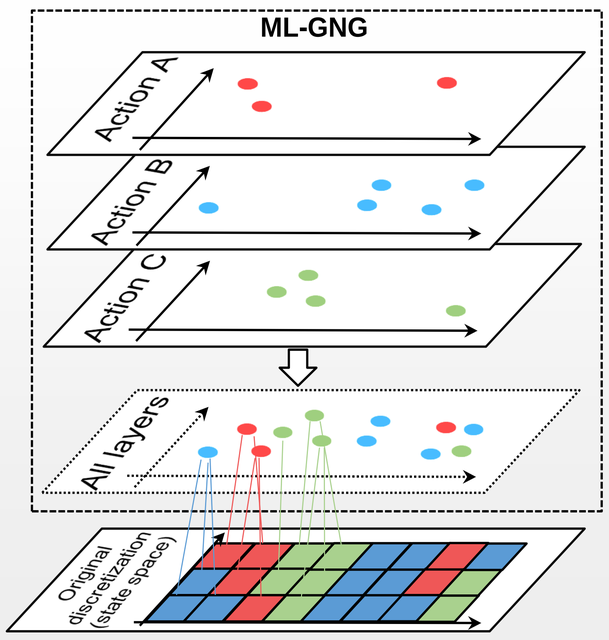
2. Constructivist approaches

- Inspired from a theory [1] that models human mind construction process
- Models the continuous construction and adaptation of knowledge through accommodation and assimilation
- A framework with RL has been proposed, but at a conceptual level [2]

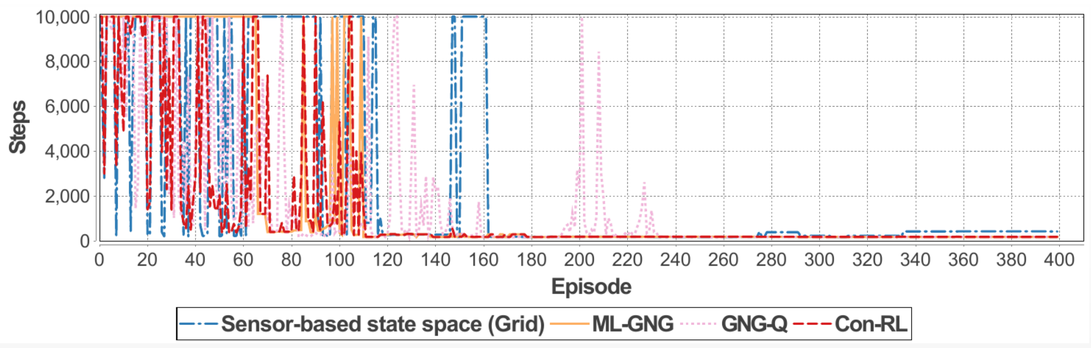
3. Con-RL: Constructivist RL for dynamic state space adaptation



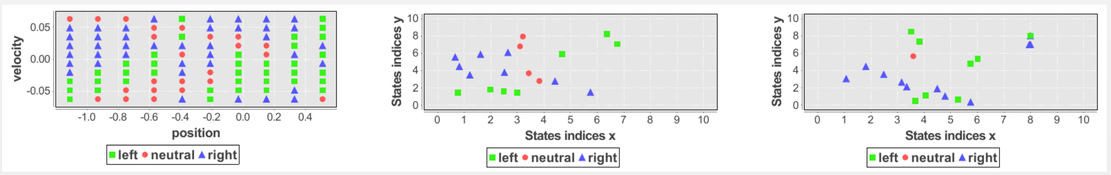
- Dynamic state space learning
  - Using a Multi-Layer Growing Neural Gas
  - Each layer is a GNG [3], a self-organizing network, specialized in one action and is triggered when an action was picked θ times for the same state
  - All layers are combined as a new learnt state space to provide a generalization of the sensor-based state space
- Dynamic state space selection
  - Action selection relies on a confidence value and configurable thresholds



4. Evaluation in Mountain Car

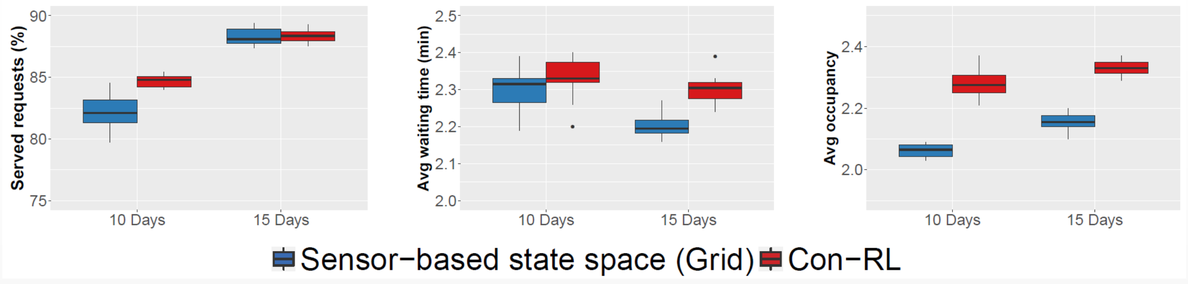


- Sensor-based discretization: 10x10 grid-based state space
- Baseline: GNG-Q [4]



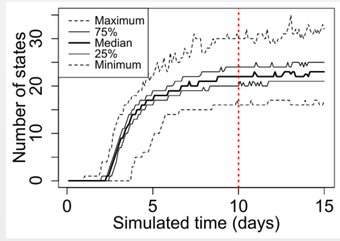
Sensor-based policy | ML-GNG nodes/actions | GNG-Q nodes/actions

5. Results in Shared Autonomous Mobility on Demand

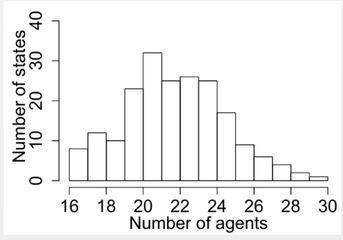


- 200 agents in SAMoD ride-sharing simulation [5]
- New York City taxi requests [6], 15 Tuesdays (7-10am)

Evolution of ML-GNG number of states during learning



Distribution of ML-GNG number of states after 10 days



Acknowledgments

This work is supported in part by a research grant from Science Foundation Ireland (SFI) under Grant Number 13/RC/2077 and 16/SP/3804 and by Irish Research Council through "Surpass: how shared autonomous cars will transform cities" New Horizons award.

References

[1] J. Piaget. The Construction of Reality in the Child; translated by Margaret Cook. Ballantine New York, 1954.
[2] R. Sutton. Presentation: Mind and time: A view of constructivist reinforcement learning. 2008. 8th European Workshop on Reinforcement Learning.
[3] B. Fritzke. A self-organizing network that can follow non-stationary distributions. In International conference on artificial neural networks, pages 613–618. Springer, 1997.
[4] M. Baumann and H. Kleine Büning. Adaptive function approximation in reinforcement learning with an interpolating growing neural gas. In 12th International Conference on Hybrid Intelligent Systems, pages 512–517, 2012.
[5] M. Guériau and I. Dusparic. SAMoD: Shared autonomous mobility-on-demand using decentralized reinforcement learning. 21st International Conference on Intelligent Transportation Systems (ITSC), pages 1558–1563, 2018.
[6] NYC Taxi and Limousine Commission. Tlc trip record data, 2018. http://www.nyc.gov