PhD in Information Technology and Electrical Engineering
Università degli Studi di Napoli Federico II

PhD Student: Francesco De Lellis

Cycle: XXXV

Training and Research Activities Report

Year: First

Francesco De Lellis

Tutor: prof. Mario di Bernardo

Co-Tutor: Giovanni Russo (University of Salerno)

Date: October 21, 2020

1. Information:

- PhD student: Francesco De Lellis
- DR number: DR993887
- Date of birth: 24/04/1993
- Master Science degree: Ingegneria dell'Automazione University: Università degli studi di Napoli Federico II
- Doctoral Cycle: XXXV
- Scholarship type: *UNINA*
- Tutor: Mario di Bernardo
- Co-tutor: Giovanni Russo (University of Salerno)

2. Study and training activities:

Activity	Type[1]	Hours	Credits	Dates	Organizer	Certificate[2]
Deep Learning Onramp	Seminar	2	0.2	21/11/2019	Carlo Sansone	Y
Lo spazio cibernetico come dominio bellico	Seminar	2	0.2	15/11/2019	Guglielmo Tamburrini	Y
Additive Manifacturing: Modeling and Challenges	Seminar	2	0.2	19/02/2020	Ferdinando Auricchio	Y
Computational Biology: Large scale data analysis to understand the molecular bases of human diseases	Seminar	2	0.2	09/04/2020	Michele Ceccarelli	Y
SINCRO Research Seminar Serie	Seminar	16	1.6	04/03/2020 11/03/2020 18/03/2020 25/03/2020 08/04/2020 15/04/2020 22/04/2020 29/04/2020	Mario di Bernardo	Y
Model Predictive Control	Course	26	2	03-11/06/2020	Alberto Bemprad	Y
Innovation management, entrepreneurship and intellectual	Course	18	5	23/04/2020 – 19/06/2020	Pierluigi Rippa	Y

property						
Joint Design of Optics and Post-Processing Algorithms Based on Deep Learning for Generating Advanced Imaging Features	**Seminar**	**2**	**0.2**	**19/05/2020**	**Raya Giryes**	**Y**
Deep Reinforcement Learning per la risoluzione di problemi di Controllo	**Seminar**	**2**	**0.2**	**14/05/2020**	**Gianfranco Fiore**	**Y**
Large Scale Training of Deep Neural Networks	**Seminar**	**2**	**0.2**	**06/05/2020**	**Giuseppe Fiameni**	**Y**
Exploring autonomy in robotic colonoscopy	**Seminar**	**2**	**0.2**	**12/06/2020**	**Pietro Valdastri**	**Y**
Bias from the wild	**Seminar**	**2**	**0.2**	**26/05/2020**	**Nello Cristianini**	**Y**
SINCRO Research Seminar Serie	**Seminar**	**14**	**1.4**	**06/05/2020 13/05/2020 20/05/2020 27/05/2020 10/06/2020 17/06/2020 24/06/2020**	**Mario di Bernardo**	**Y**
Statistical Learning	**Course**	**44**	**6**	**17/03/2020-09/06/2020**	**Roberta Siciliano**	**N**
Machine Learning	**Course**	**38**	**3.6**	**06/06/2020-17/07/2020**	**Carlo Sansone**	**Y**
SINCRO Research Seminar Serie	**Seminar**	**10**	**1**	**01/07/2020 08/07/2020 15/07/2020 22/07/2020 29/07/2020**	**Mari di Bernardo**	**Y**

1) Courses, Seminar, Doctoral School, Research, Tutorship
2) Choose: Y or N

- **Study and training activities - credits earned**

	Courses	Seminars	Research	Tutorship	Total
Bimonth 1	**0**	**0.4**	**9.6**	**0**	**10**
Bimonth 2	**0**	**0.2**	**9.8**	**0**	**10**
Bimonth 3	**0**	**1.6**	**8.2**	**0.5**	**10.3**
Bimonth 4	**7**	**2.4**	**5**	**0.5**	**14.9**
Bimonth 5	**3.6**	**1**	**5**	**0**	**9.6**
Bimonth 6	**0**	**1.6**	**4**	**0.6**	**6.2**
Total	**10.6**	**6.8**	**41.6**	**1.6**	**61**
Expected	**30 - 70**	**10 - 30**	**80 - 140**	**0 – 4.8**	

3. Research activity:

Title of the whole research activity: *Development and applications of nonlinear and distributed control strategies based on reinforcement learning for complex and multi-agent systems*

Title: *Control-Tutored Q-learning*
Description: Reinforcement Learning (RL) provides a set of methods that can be used to solve control problems in the presence of uncertain or unknown dynamics [1]-[3]. In the existing literature, there are several methods aimed at finding an optimal control policy. In general, these methods tend to solve an infinite horizon optimal control problem under some specific constraints that make the problem solvable. To do so, one can think of gathering data to approximate the value of the optimal cost function starting from all possible state values. Such methods can be tabular and often refer to discrete spaces but the recent developments of nonlinear approximators, such as Neural Networks (NNs), opened the possibility to extend those methods to continuous spaces [4],[5]. On the other hand, the employment of NNs makes any rigorous proof of stability and convergence hard if not impossible to obtain. Trying to learn the optimal cost function is not the only method to solve such problems, some other strategies aim at defining a policy and then tune their parameters according to the experience gathered so far. These methods go under the name of Direct Policy Search and some of them still use cost learning in parallel, such as Actor-Critic methods [5],[6]. As these solutions require large amounts of data and long learning times, recent research has been focused on the use of control laws in combination with RL algorithms to solve these problems and to improve data efficiency [7]-[9]. Many efforts have been made to establish a unified framework to develop strategies that make use of Control Laws and Reinforcement Learning [10],[11]. During the first year, I developed a strategy named Control-Tutored Q-learning [12]. The aim of this research is to find out if nonlinear control strategies can be used in combination with RL algorithms to improve data efficiency and increase the control performance and robustness. The CTQL algorithm makes use of a

control policy to improve the learning phase of the Q-learning algorithm [13] by using a control law to recommended possible control action using partial knowledge of the environment dynamics encoded in a mathematical model. The key idea is that, during the learning phase, the value approximation made by Q-function can bring information about how good the learned policy is. Based on that, the CTQL decides if it is the case to use the action supplied by a feedback control law or that encoded in the Q-function to drive the artificial agent. This mix of Control Laws and Reinforcement Learning has been tested on two representative problems: the multi-agent herding problem, where one or more agents have to collect and contain a group of target agents, and the classical benchmark of stabilizing an inverted pendulum. In both cases, the CTQL showed its capability of reducing learning times, improving data efficiency, and achieving satisfactory results with the use of only partial model knowledge. As next steps, I plan to:

- determine the set of properties that the control policy must fulfill in the CTQL approach
- determine the hypothesis that the partial mathematical model must fulfill in the CTQL approach
- search for combinations that involve different kinds of Reinforcement Learning algorithms like the Direct Policy Search methods.

Collaboration: Fabrizia Auletta, Giovanni Russo, Pietro De Lellis and Mario di Bernardo

In addition to the research above, I also worked on deriving a model of Italy as a network of regions to capture the COVID-19 epidemic. This work showed how decisions taken at the regional level can make a huge impact on the containment of the epidemic and economic costs. We acted on this model with a simple bang-bang control strategy based on the current number of intensive care units in use at the regional level. The results showed that a localized intermittent lockdown strategy brings substantial advantages compared to an aggregate national lockdown measure. Despite being accurate as a model, several discrepancies exist with the actual epidemic dynamics. So current research aims at checking if the trivial bang-bang control strategy can be assisted by a reinforcement learning algorithm to find better control solutions that could potentially bring robustness to unmodeled dynamics.

Collaboration: Fabio Della Rossa***, Davide Salzano*, Aanna Di Meglio*, Marco Coraggio*, Carmela Calabrese*, Agostino Guarino*, Ricardo Cardona-Rivera*, Pietro De Lellis*, Davide Liuzza****, Francesco Lo Iudice*, Giovanni Russo** and Mario di Bernardo*

*(University of Naples Federico II)

**(University of Salerno)

***(Polytechnic of Milan)

****(ENEA)

4. Research products:

- **Journal Papers:**
  - F. Della Rossa*, D. Salzano*, A. Di Meglio*, F. De Lellis*, M. Coraggio, C. Calabrese, A. Guarino, R. Cardona-Rivera, P. De Lellis, D. Liuzza, F. Lo Iudice, G. Russo, M. di Bernardo -- "*A network model of Italy shows that intermittent regional strategies can alleviate the COVID-19*" -- Nature Communications, 11, 5106, 2020.
- **Preprint**
  - F. De Lellis, F. Auletta, G. Russo, P. De Lellis, M. di Bernardo -- "*Control-Tutored Reinforcement Learning*", arXiv:1912.06085, 2019.

5. Conferences and seminars attended

- **IEEE CSS Modeling and Control the COVID-19 outbreak**
  - Dates: 24/04/2020
  - Location: Web hosted
  - Co-author of the paper presented by prof. Mario di Bernardo: "*A network model of Italy shows that intermittent regional strategies can alleviate the COVID-19"*
- **Learning for Dynamics and Control (L4DC)**
  - Dates: 11-12/06/2020
  - Location: ETH in Zürich, Switzerland (Web hosted)
- **Second Symposium on Machine Learning and Dynamical Systems**
  - Dates 21-29/09/2020
  - Location: Fields Institute, Toronto (Web hosted)

6. Activity abroad:

- **University College of Dublin (Dublin, Ireland):**
  - Carried out theoretical studies on Reinforcement Learning techniques under the supervision of prof. Giovanni Russo of the department of Electrical & Electronic Engineering
  - Dates: 12/01/2020 – 12/02/2020
- **Number of months spent abroad in 2020:** 1

7. Tutorship

- Co-supervisor of the thesis of 2 master students in Ingegneria dell'Automazione
- Weekly 2 hours tutorship ("ricevimento") for the course of Dinamica e Controllo Non Lineare in Ingegneria dell'Automazione

8. References

[1] D. P. Bertsekas, and J. N. Tsitsiklis, “Neuro-dynamic programming.” Athena Scientific, 1996.
[2] D. P. Bertsekas, “Reinforcement learning and optimal control.” Belmont, MA: Athena Scientific, 2019.
[3] R. S. Sutton and A. G. Barto, “Reinforcement learning: An introduction.” MIT press, 2018.
[4] V. Mnih, et al., "Human-level control through deep reinforcement learning." Nature, p. 529-533, 2015.
[5] T. P. Lillicrap, et al. "Continuous control with deep reinforcement learning." arXiv:1509.02971, 2015.
[6] V. R. Konda, and J. N. Tsitsiklis. "Actor-critic algorithms." Advances in neural information processing systems, p. 1008-1014, 2000.
[7] M, Rathi, P. Ferraro, and G. Russo. "Driving reinforcement learning with models." Proceedings of SAI Intelligent Systems Conference. Springer, Cham, p. 70-85, 2020.
[8] A. M. Fadel, J. Fürst, and B. Cheng. "Tutor4RL: Guiding Reinforcement Learning with External Knowledge." AAAI Spring Symposium: Combining Machine Learning with Knowledge Engineering, 2020.
[9] P. Abbeel, M. Quigley, and A. Y. Ng. "Using inaccurate models in reinforcement learning." Proceedings of the 23rd international conference on Machine learning, p. 1-8, 2006.
[10] N. Matni, et al. "From self-tuning regulators to reinforcement learning and back again." IEEE 58th Conference on Decision and Control (CDC), p. 3724-3740 ,2019.
[11] B. Recht, "A tour of reinforcement learning: The view from continuous control." Annual Review of Control, Robotics, and Autonomous Systems 2, p. 253-279, 2019.
[12] F. De Lellis, F. Auletta, G. Russo, P. De Lellis, M. di Bernardo -- "Control-Tutored Reinforcement Learning", arXiv:1912.06085, 2019.
[13] C. Watkins, P. Dayan, “Q-learning.” Machine learning, p.279-292, 1992.