ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
UNIVERSITY OF PIRAEUS
ΤΜΗΜΑ ΨΗΦΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

Δ.Π.Μ.Σ. ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ

DEMOKRITOS
ΙΝΣΤΙΤΟΥΤΟ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

Προτεινόμενες Θεματικές Διπλωματικών Εργασιών[1]

Από

Καθηγητή Γεώργιο Βούρο

Απρίλιος 2026

A. Safe Reinforcement Learning (Ασφαλής Ενισχυτική Μάθηση)

Description: Theses in this thematic area aim to study state of the art safe reinforcement learning algorithms and develop novel ones with performance guarantees, demonstrating their performance in benchmarking settings.

The goal is challenging given that most safe reinforcement learning algorithms, training models through exploration, make unsafe decisions during training and inference and do not have performance guarantees. In many applications, unsafe decisions are undesirable even at the training phase.

This area shares a lot with control theory and has many applications in foundational models and robotics.

The focus is in reinforcement learning, although methods from control theory may also be investigated in combination with reinforcement learning methods.

Goals and Deliverables:

- A safe reinforcement learning algorithm with state of the art results on benchmarks
- Theoretical results with respect to performance bounds and safety guarantees of reinforcement learning algorithms

Specific learning outcomes:

- Specialization in practical and theoretical aspects of deep reinforcement learning
- Specialization in safe decision making in high-stakes domains

References

[1] Το έγγραφο αναφέρεται σε θεματικές και όχι σε συγκεκριμένα θέματα. Τα θέματα θα εξειδικευτούν αναλόγως των ενδιαφερόντων των φοιτητριών /φοιτητών και του διδάσκοντα.

1. Lukas Brunke et al., Safe Learning in Robotics: From Learning-Based Control to Safe Reinforcement Learning, 2021, https://arxiv.org/abs/2108.06266
2. W.Zhao et al., State-wise safe reinforcement learning: A survey, https://www.ijcai.org/proceedings/2023/0763.pdf
3. Luca Marzari, Ferdinando Cicalese, Alessandro Farinelli, Christopher Amato, and Enrico Marchesini. 2025. Verifying Online Safety Properties for Safe Deep Reinforcement Learning. ACM Trans. Intell. Syst. Technol. 17, 1, Article 3 (February 2026), 27 pages. https://doi.org/10.1145/3770068

B. Ontology Alignment through Linguistic Reinforcement Learning with Foundational models

(Αντιστοίχιση Οντολογιών με αξιοποίηση θεμελιωδών μοντέλων και χρήση ενισχυτικής μάθησης βασιζόμενης στη γλώσσα)

Description: Ontologies are artifacts for describing data in semantically valid ways, formalizing domain and data conceptualizations. Their necessity to combine distinct and/or fragmented, disparate and multi-modal data sources is of importance to allow agents to reason jointly, providing context to their reactions and reasoning tasks. However, having multiple ontologies describing distinct data sources is a major problem, as these may describe common terms and relations in different ways, at different levels of abstraction etc. By aligning ontologies, we aim to find semantic correspondences between ontology elements, with the objective agents to share knowledge by mapping entities from their own ontology to others, obtaining a common context.

The use of LLMs, providing rich linguistic and domain knowledge, aim to enhance the discovery of semantic relations among ontology entities. Foundational models for different data modalities (e.g. VLMs) can help in aligning ontologies with respect to multi-modal data. Theses in this thematic area aim to address the ontology alignment problem with the use of foundational models, also investigating the use of reinforcement learning techniques in this context.

The focus regarding novelty can be either in reinforcement learning, or in ontology alignment or in both.

Goals and Deliverables:

- Linguistic reinforcement learning method under semantic constraints
- A method for ontology alignment using foundational models and reinforcement learning techniques
- Results from evaluating the proposed method in Ontology Alignment Evaluation Initiative benchmarks.

Specific learning outcomes:

- Specialization in linguistic reinforcement learning methods with foundational models
- Specialization in ontology alignment with the assistance of foundational models

References

1. Guanting Dong et al., Agentic Reinforced Policy Optimization, 2025, https://arxiv.org/abs/2507.19849
2. N.Shinn, et al., Reflexion: Language Agents with Verbal Reinforcement Learning, 2024, https://arxiv.org/abs/2303.11366
3. Y.Zhou et al., Prompted Policy Search: Reinforcement Learning through Linguistic and Numerical Reasoning in LLMs
4. Y.He et al., Exploring large language models for ontology alignment
5. R.Amini et al., Towards complex ontology alignment using large language models, 2024, https://link.springer.com/chapter/10.1007/978-3-031-81221-7_2
6. https://oaei.ontologymatching.org/

C. Language assisted Deep Reinforcement Learning (Βαθιά Ενισχυτική Μάθηση υποβοηθούμενη από πράκτορες φυσικής γλώσσας)

Description: The focus of such a thesis will be to combine deep reinforcement learning with language agents. This fusion, often termed **"language-based reinforcement learning"** allows agents to use natural language to critique their own decisions, which then guides the numerical optimization of DRL.

The benefits of this combination can be as follows:

Improved Sample Efficiency: By incorporating human-like reasoning into exploration, agents require fewer interactions to learn.

Enhanced Stability: Language agents help address the "non-stationary" problem in DRL (where the environment appears to change because the policy changes), making training more stable.

Reduced Reward Engineering: The verbal reflection acts as a natural "reward shaper," guiding the agent without requiring the design of reward functions.

Better Generalization: Agents may be able toi generalize better to unseen scenarios.

Goals and Deliverables:

- Study existing methods and design a method that deep reinforcement learning and language-assisted reinforcement learning.
- Evaluate the method in benchmarks, and prove that it has certain benefits compared to deep reinforcement learning and language-assisted agents, alone.

Specific learning outcomes:

- Specialization in language-based reinforcement learning methods with foundational models
- Deep dive into agentic AI

References

1. Shinn, N., Cassano, F., Gopinath, A., Narasimhan, K. and Yao, S., 2023. Reflexion: Language agents with verbal reinforcement learning. *Advances in neural information processing systems*, *36*, pp.8634-8652.

2. Zhang, W., Tang, K., Wu, H., Wang, M., Shen, Y., Hou, G., Tan, Z., Li, P., Zhuang, Y. and Lu, W., 2024, August. Agent-pro: Learning to evolve via policy-level reflection and optimization. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 5348-5375).
3. Hao Pang, Zhenpo Wang, Guoqiang Li, Large language model guided deep reinforcement learning for safe autonomous vehicle decision making, Transportation Research Part C: Emerging Technologies, Volume 184, 2026, 105511, ISSN 0968-090X, https://doi.org/10.1016/j.trc.2025.105511.
4. Su, Haoran, Yandong Sun, and Congjia Yu. "The End of Reward Engineering: How LLMs Are Redefining Multi-Agent Coordination." *arXiv preprint arXiv:2601.08237* (2026).

D. Modelling human gait with imitation learning methods
(Μοντελοποίηση ανθρώπινης κίνησης με μεθόδους μιμητικής μάθησης)

Description: This thesis aims to model human gait by means of imitation learning techniques, using available datasets of healthy persons at various ages.

The benefits of this modelling can be as follows:

- Diagnose abnormal gait
- Support humans to stay safe while moving
- Identifying humans using gait patterns

Goals and Deliverables:

- Study state of the art imitation learning methods.
- Formulate the problem and select the most prominent methods.
- Evaluate comparatively methods in modelling human gait.

Specific learning outcomes:

- Specialization in imitation learning
- Knowledge on applying advanced machine learning in modelling human gait

References

1. Wiles TM, et al.. NONAN GaitPrint: An IMU gait database of healthy young adults. Sci Data. 2023 Dec 5;10(1):867. doi: 10.1038/s41597-023-02704-z. PMID: 38052819; PMCID: PMC10698035. https://pubmed.ncbi.nlm.nih.gov/38052819/
2. Wiles TM, et al.,. NONAN GaitPrint: An IMU gait database of healthy older adults. Sci Data. 2025 Jan 24;12(1):143. doi: 10.1038/s41597-024-04359-w. PMID: 39856107; PMCID: PMC11759707.https://pubmed.ncbi.nlm.nih.gov/39856107/
3. Wiles TM, Kim SK, Manifrenti MK, Brink KJ, Grunkemeyer A, Stergiou N, Likens AD. NONAN GaitPrint: An IMU gait database of healthy middle-aged adults. Sci Data. 2025 Sep 30;12(1):1603. doi: 10.1038/s41597-025-05892-y. PMID: 41027945; PMCID: PMC12484833. https://pubmed.ncbi.nlm.nih.gov/41027945/