

Nitay Alon

nitay.alon@mail.huji.ac.il • Google Scholar profile • +972 (54) 4894004 •

SUMMARY

As a PhD candidate in the intersection of Computer Science, Game Theory and Computational Cognitive Science, my research integrates Multi-Agent Reinforcement Learning and probabilistic models to model Theory of Mind. I am seeking a post-doctoral position to expand on my thesis work, with the long-term career goal of leading a lab focused on computational models of social behavior.

EDUCATION

Hebrew University, Jerusalem, IL

MPI for biological cybernetics, Tübingen, GER

- PhD in Computer Science Oct 2020 – Feb 2026 (Expected)
 - Dissertation: The role of Theory of Mind in partially observed mixed-motive problems
 - Advisers: Prof. Jeff Rosenschein (HUJI), Prof. Peter Dayan (MPI)
 - Focus: Theory of Mind in Mixed-Motive Partial Information Bayesian Games

Tel Aviv University, Tel Aviv, IL

- M.Sc. in Statistics and Machine Learning Oct 2015 – Oct 2019
 - Thesis: The effect of drift on the Skorokhod-embedded distribution of stopped Brownian Motion, Hidden Markov applications to Finance, Adviser: Prof. Isaac Meilijson
 - **Magna Cum Laude**, Thesis final grade 98
- B.S. in Statistics and Operations Research Oct 2012 – Oct 2015
- B.A. in Economics Oct 2012 – Oct 2015

PUBLICATIONS

JOURNALS

- Mitigating Deception in a Cognitive Hierarchy with Off-Policy Counterfactual Anomaly Detection 2026
 - *Proposed an off-policy counterfactual mechanism to detect and mitigate deception in multi-agent interactions. The method enhances agent robustness against manipulative adversaries in hierarchical cognitive settings.*
 - Authors: **Alon N.**, Barnby, J. M., Sarkadi S., Schulz L., Rosenschein J. S. & Dayan, P.
 - Alon, Nitay, et al. "Detecting and deterring manipulation in a cognitive hierarchy." *Accepted, JAIR* 2025.
- (Mal)adaptive Mentalising in the Cognitive Hierarchy, and Its Link to Paranoia 2024
 - Authors: **Alon N.**, Schulz L., Bell, V., Moutoussis, M., Dayan, P., & Barnby, J. M.
 - Alon, Nitay, et al. "(Mal) adaptive mentalizing in the cognitive hierarchy, and its link to paranoia." *Computational Psychiatry* 8.1 (2024): 159.
- A (dis-) information Theory of Revealed and Unrevealed Preferences: Emerging Deception and Skepticism via Theory of Mind 2023
 - Authors: **Alon N.**, Schulz L., Rosenschein J.S. & Dayan P.
 - Alon, Nitay, et al. "A (dis-) information theory of revealed and unrevealed preferences: Emerging deception and skepticism via theory of mind." *Open Mind* 7 (2023): 608-624.

CONFERENCES

- Theory-of-Mind in Partially Observed, Mixed-Motive Games 2026
 - Authors: **Alon N.**
 - Alon, Nitay. "Theory-of-Mind in Partially Observed, Mixed-Motive Games" *Proceedings of the AAAI Conference on Artificial Intelligence*
- Mind Your Theory: Theory of Mind Goes Deeper Than Reasoning 2025
 - Authors: Wagner E.*, **Alon N.***, Barnby, J. M. & Abend, O.
 - Wagner, Eitan, et al. "Mind your theory: Theory of mind goes deeper than reasoning." *Findings of the Association for Computational Linguistics: ACL* 2025. 2025.

WORKSHOPS

- Between prudence and paranoia: Theory of Mind gone right, and wrong 2023
 - Authors: **Alon N.**, Schulz L., Dayan P., Barnby J. M.
 - Alon, Nitay, et al. "Between prudence and paranoia: Theory of mind gone right, and wrong." *First workshop on theory of mind in communicating agents*. 2023.

PROFESSIONAL SERVICE

ToM4AI Workshop,

Initiated and co-organized the first AAAI workshop dedicated to Theory of Mind in AI, coordinating with keynote speakers and managing the review process for 60+ submissions.

- Co-organizer: 2026
Organizing ToM4AI workshop at AAAI2026
- Co-organizer: 2025
Organizing ToM4AI workshop at AAAI2025

The Israel Statistics and Data Science Association,

- ISDSA council member: 2024-2025
Co-organizing workshops on Causality and Explainability

TEACHING

The Hebrew University,

Benin School of Computer Science and Engineering

- TA: 2025 – 2026
 - Class: Algorithmic Game Theory
 - Details: Grad students course on Algorithmic Game Theory
- TA: 2021 – 2025
 - Class: Undergrad graduation project, special Data Science program
 - Details: Mentoring students in a year long course, from data collections and basic statistics to application of ML and reporting results to stakeholders
- TA: 2024 – 2025
 - Class: Introduction to Artificial Intelligence
 - Details: Undergrad introductory course for AI algorithms
- TA: 2021 – 2023
 - Class: Introduction to Machine Learning
 - Details: Undergrad introduction class to ML

Tel Aviv University,

Coller School of Management

- Instructor: 2016–2020
 - Class: Big Data Technologies
 - Details: Data Engineering, Hadoop, Spark, Kafka

MEMBERSHIPS

AAAI,

- AAAI Member: 2025-Present
Member of AAAI organization
- The Israel Statistics and Data Science Association Member: 2023-Present
Member of The Israel Statistics and Data Science Association

AWARDS

- Ernst reward for Excellency 2020
Awarded for excellency in Master program. Best Thesis award.

COMMUNITY SERVICE

Paamonim,

- Team Leader: 2015-2019
Led a 5-person volunteer team, mentoring and monitoring progress

LANGUAGES

English: Fluent Hebrew: Native language German: Elementary