

Utku Umur Acikalin

uuacikalin@gmail.com | utkuumur.github.io | Ithaca, NY

Google Scholar | github.com/utkuumur | LinkedIn

Fourth-year Cornell CS Ph.D. student advised by Carla P. Gomes. Builds machine learning to identify unknown molecules from tandem mass spectra: neuro-symbolic spectrum prediction, learned search over molecular graphs, and retrieval over 105M structures, with discoveries confirmed by synthesis.

EDUCATION

Cornell University · Ph.D. in Computer Science

2023 - expected 2028

Advisor: Carla P. Gomes

TOBB University of Economics and Technology

Ankara, Turkey

M.S., Computer Engineering, 2019 - 2022 · GPA 4.00/4.00

B.S., Computer Engineering, 2015 - 2019 · GPA 3.74/4.00

RESEARCH EXPERIENCE

Ph.D. Researcher · Cornell University, Gomes Lab · with the Schroeder Lab (chemistry)

2023 - present

AIME & MS²KOSMOS: neuro-symbolic mass spectrometry at database scale

Co-first author · bioRxiv 2026, under review at Nature · [Preprint](#) · [Public system](#)

Outcome: synthesis-validated discovery of previously unreported microbiota-dependent polyamines.

- Co-designed and implemented **AIME**, a neuro-symbolic MS/MS predictor that explains peaks through chemically constrained fragmentation pathways; **0.83 cosine on NIST** vs. 0.74 for the strongest baseline.
- Built CPU/GPU retrieval over **105M structures and 800M+ predicted spectra** (MS²KOSMOS), with interactive queries in a few seconds and a repository-scale search of 7.14M GNPS consensus spectra.
- Co-developed the public web system for spectral search, pathway inspection, and on-demand prediction.

GEMS: structure elucidation as learned search over isomer graphs

First author · under review at ICLR 2027 · ICML 2026 workshops: GenBio (spotlight), AI4Physics · [Workshop paper](#)

- Developed a learned search that refines candidates through formula- and valence-preserving bond edits; a policy maps MIST-predicted fingerprints to atoms and bonds and scores all edits in one pass.
- Designed shortest-path supervision, enabling pretraining on PubChem structures alone and fine-tuning on spectrum-predicted fingerprints.
- Hit@1 of 29.6% on NPLIB1 and 30.2% on MassSpecGym** with the target excluded from initialization, above published baselines in comparable settings; higher accuracy than FRIGID on NPLIB1 in ~2.6 min vs. 26 h.

MoRE: Mixture of Reused Experts

Co-first author · COLM 2026 · [arXiv](#) · [Code](#)

- Co-developed a middle ground between standard MoE and looped (Universal) Transformers: adjacent layers share an expert pool with per-layer routers, giving ~18x more routing choices at no parameter cost.
- Beats a DeepSeek-v3-style MoE and MoEUT at 114M-1.15B parameters with matched parameters and FLOPs.

X²GNN: learning to explore and exploit

First author · ICLR 2025 · [Code](#)

- Unsupervised GNN solver for MIS, Max Clique, Max-Cut; trained on one refinement step, it scales with test-time compute: 1.2% from optimal on large Max Clique graphs (prior neural: >22%), beating Gurobi at similar runtime.

SELECTED PUBLICATIONS & PREPRINTS

* Equal contribution.

Intellectual property protection lost and competition: An examination using large language models.

U. U. Acikalin, T. Caskurlu, G. Hoberg, G. M. Phillips.

Journal of Financial Economics 182:104306, 2026.

Charting the small-molecule universe from mass spectra with neuro-symbolic AI.

U. U. Acikalin*, D. Feng*, A. M. Ferber*, G. J. Gouveia*, et al.

bioRxiv 2026; under review at Nature.

GEMS: Molecular Structure Elucidation as Learned Search over Isomer Graphs.

U. U. Acikalin, Y. Wang, G. J. Gouveia, T. Schwertfeger, F. C. Schroeder, C. P. Gomes.

Under review at ICLR 2027. Earlier version: ICML 2026 workshops, GenBio (spotlight) and AI4Physics.

MoRE: Mixture of Reused Experts.

E. S. Qiu*, U. U. Acikalin*, J. Lovelace, C. Belardi, A. B. Mulchandani, C. P. Gomes, K. Q. Weinberger.

COLM 2026.

Learning to Explore and Exploit with GNNs for Unsupervised Combinatorial Optimization.

U. U. Acikalin, A. M. Ferber, C. P. Gomes.

ICLR 2025.

Unsupervised Combinatorial Probabilistic Reasoning: Probabilistic Coin Change Problem.

Z. Qu, Y. Wang, U. U. Acikalin, et al.

AAAI 2026.

Earlier work (11 papers, 2019 - 2025): combinatorial optimization (INFORMS Journal on Computing, Constraints, GECCO), algorithmic game theory, NLP (Natural Language Engineering, SIGIR PatentSemTech workshop), and economics. Full list at utkuumur.github.io.

TECHNICAL SKILLS

Machine learning: PyTorch, Hugging Face Transformers, scikit-learn; graph neural networks, mixture-of-experts transformers, learned search policies, LLM-based text extraction.

Molecular ML: RDKit; MS/MS spectrum prediction and retrieval, fragmentation modeling, molecular fingerprints (Morgan, MIST), isomer-graph search; NIST20, MassSpecGym, NPLIB1, CASMI, GNPS.

Optimization and HPC: integer programming (Gurobi, CPLEX), OR-Tools, A* and heuristic search; CPU/GPU similarity search at 100M+ scale, MPI, Cerebras CSL, AWS.

Languages: Python, C++, Java, JavaScript.

PRESENTATIONS, SERVICE & TEACHING

ICML 2026, Seoul: GEMS at GenBio (spotlight) and AI4Physics; MoRE at Weight-Space Symmetries.

Reviewer: NeurIPS 2026; ICML 2026 GenBio Workshop; IEEE Transactions on Circuits and Systems I.

Teaching assistant: Introduction to Computer Programming; Design and Analysis of Algorithms; Social Networks; Combinatorics.

SELECTED AWARDS

TUBITAK 2210 M.S. Scholarship, 2020 - 2022; **TOBB ETU Graduate Special Merit Scholarship**, 2019 - 2021; **TUBITAK Undergraduate Research Project Contest**, 2nd prize, 2019.