Safe Autonomous Systems

EEL 6935

Class Periods: Mondays/Wednesdays/Fridays, 11:45am – 12:35pm

Location: Larsen Hall 330

Academic Term: Fall 2022

Instructor:
Ivan Ruchkin
iruchkin@ece.ufl.edu
Office Phone Number: 352-273-2171
Office Hours: Wednesdays, 10:30–11:30 AM starting August 31
Office Location: Larsen Hall 334B (go through the wooden doors of the ACIS lab, then turn right)

Teaching Assistant/Peer Mentor/Supervised Teaching Student: N/A

Course Description

Mathematical and algorithmic techniques and tools for building safe autonomous systems, from unmanned drones to self-driving cars to smart grids to medical devices.

The course is split into three modules:
(i) *Systems:* how to understand and formally represent a complex autonomous system? Mathematical formalisms for dynamical, probabilistic, and hybrid systems.
(ii) *Autonomy:* how to enable autonomous operation in a challenging environment? Data-driven components, probabilistic graphical models, neural networks, errors/robustness of learned models, learned perception/control.
(iii) *Safety:* how to ensure that an autonomous system behaves safely? Logical specifications of safety, verification of system models and learned components, safe planning and control, temporal and confidence monitoring.

Credits: 3.

Course Pre-Requisites / Co-Requisites

Required skills: linear algebra, basic calculus, basic probability/statistics, basic control/dynamical systems, and Boolean logic. If you are unsure about your skills in any of these areas, feel free to reach out to the instructor at iruchkin@ece.ufl.edu with a description of your background and how this course would be useful to you.

Nice to have but not required: EEE 5544 Stochastic Methods for Engineering 1, EEE 5702 Automated Hardware/Software Verification, CAP 5635 Artificial Intelligence Concepts, EEL 5840 Fundamentals of Machine Learning.

Course Objectives

What are the key challenges in designing a safety-critical autonomous system, such as a self-driving car, and how do we overcome them today? This course will teach you rigorous state-of-the-art mathematical and algorithmic techniques for building such systems. These techniques enable engineers to define what safety means in the context of this system, make a mathematical model of the system and its environment, analyze this model for safety, and implement the system using a combination of first-principles (e.g., logic-based synthesis) and data-driven (e.g., training neural networks) techniques.

By taking this course, you will gain

- A fundamental understanding of the theories and tools used to build safe autonomous systems through the readings, lectures, and in-class working sessions and discussions.
- First-hand experience of building and analyzing an autonomous system through an individualized course project, chosen and executed under the instructor's close supervision.

The course project can be an algorithm implementation, a case study in an application domain, a theoretical investigation, or a systematic review of literature identifying research opportunities in a well-defined area. Projects with an experimental component can build on a simulated environment (e.g., OpenAI gym, Carla, AirSim), an open dataset (e.g., Waymo, nuScenes), or any physical system/non-public dataset a student has access to. Project examples: analyze the safety of landing an airplane with faulty perception; apply multi-agent reinforcement learning to navigate a traffic junction; detect safety-critical occlusions in an autonomous driving dataset; review and analyze the research literature on diagnosing livelocks in learning-based systems. No single dataset, simulator, or development platform will be imposed on the whole class. Students are encouraged to formulate a project around their domain of interest or their current/future research problem.

This course is great for students who want to gain state-of-the-art knowledge in safe autonomy, deepen their understanding of rigorous design and validation of complex systems, gain practical experience with a particular type of autonomous systems, or complement their research with safety assurance.

Materials and Supply Fees

- Students are expected to use a personal computer for this course
- The required software will be either freely available or provided by the instructor

Required Textbooks and Software

- None

Recommended Materials

Most course material will be based on select chapters from the following books:

- Verifying Cyber-Physical Systems: A Path to Safe Autonomy, Sayan Mitra, 2021, ISBN: 978-0262044806
- Principles of Cyber-Physical Systems, Rajeev Alur, 2015, ISBN: 978-0262029117
- Machine Learning: A Probabilistic Perspective, Kevin Murphy, 2012, ISBN: 978-0262305242

Lectures will be supplemented with reading chapters from the above books and research papers.

Course Schedule

Every week, lectures on Monday and Wednesday introduce new material, which is practiced in an interactive working session on Friday.

Module	Week	Class topic	Concepts	Evaluation
Intro	1	Introduction to the course	Syllabus, organization, policies. Motivating application domains. Review of linear algebra, calculus, probability, and propositional logic.	
Systems	2	Dynamical systems	Discrete- and continuous-time systems. State machines.	
	3	Probabilistic systems	Probabilistic automata. Markov chains and decision processes.	
	4	Hybrid systems	Ordinary differential equations. Hybrid automata. Model composition.	
Autonomy	5	Data-driven setting	Probabilistic formulations of standard AI/ML problems. Types of uncertainty.	Project proposal due
	6	Learning models	Graphical models. Probabilistic	Instructor feedback on

			inference. Feedforward and convolutional neural networks.	project proposal
	7	Errors and rates	True/false positive/negative rates. Rate guarantees. Error norms. PAC.	
	8	Robustness	Adversarial robustness. Attacks on neural network inputs. Data poisoning. Confidence calibration. Distribution shift.	Midterm exam
	9	Learned autonomy	Learned perception and control. Application domains for autonomy.	
Safety	10	Specification	Temporal and metric logics. Branching tree logics. Probabilistic logics.	
	11	Verification	State-space exploration. Reachability analysis. Probabilistic model checking. Neural network safety & robustness verification.	
	12	Control	Safe control and plan synthesis. Simplex architecture.	
	13	Monitoring	Temporal specification monitoring. Falsification. Confidence monitoring.	Preliminary project due, instructor feedback
Outro	14	Student project presentations		
	15	Student project presentations		Final project due

Evaluation of Grades

Assessment component	**Total Points**	**Percentage of Final Grade**
Class participation	100	25%
Midterm exam	120	30%
Project proposal: vision, viability, clarity	40	10%
Project preliminary: soundness, clarity, progress	40	10%
Project presentation: clarity, timing	40	10%
Project final: soundness, clarity, outcome	60	15%

This class has no homeworks and no final exam. Students will deepen their understanding by reading book chapters and research papers assigned as pre-lecture and post-lecture readings.

Class participation will be assessed based on each student's thoughtful questions, activity in working sessions, and contributions to discussions, both in-class and online. Meaningful class participation is dependent on students engaging with the assigned readings.

The midterm will assess the understanding of the theoretical concepts from the first half of the course. The in-class working sessions will walk students through problems similar to those in the midterm.

The projects will be evaluated on the correctness of the technical approach (i.e., whether the technique was applied appropriately), the clarity of their report, and the outcome (i.e., whether the desired result was achieved). Students will carry out their projects individually (i.e., no teams). Before writing the proposal, students should discuss their vision with the instructor.

Grading Policy

Percent	Grade	Grade Points
90 – 100	A	4.00
80 – 89.9	B	3.00
70 – 79.9	C	2.00
60 – 69.9	D	1.00
0 – 59.9	E	0.00

Increased fractional grades may be assigned at the instructor's discretion. More information on UF grading policy may be found at https://catalog.ufl.edu/UGRD/academic-regulations/grades-grading-policies/.

Attendance Policy, Class Expectations, and Make-Up Policy
Although absence from class does not necessarily equal zero participation, students are strongly encouraged to *attend the classes in person* for the sake of better engagement and learning.

The course does not offer late days: project submissions must be submitted on time to avoid penalties. Deadlines can be adjusted based on *excused absences*, which must be consistent with university policies in the Graduate Catalog (see https://gradcatalog.ufl.edu/graduate/ for more information) and require appropriate documentation.

Course Evaluation
Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at https://gatorevals.aa.ufl.edu/students/. Students will be notified when the evaluation period opens and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via https://ufl.bluera.com/ufl/. Summaries of course evaluation results are available to students at https://gatorevals.aa.ufl.edu/public-results/.

University Honesty Policy
UF students are bound by The Honor Pledge which states, *"We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code."* On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: *"On my honor, I have neither given nor received unauthorized aid in doing this assignment."*

The Conduct Code (https://sccr.dso.ufl.edu/process/student-conduct-code/) specifies behaviors that are in violation of this code and the possible sanctions. If you have questions or concerns, please consult the instructor.

Commitment to a Safe and Inclusive Learning Environment

The Herbert Wertheim College of Engineering values broad diversity within our community and is committed to individual and group empowerment, inclusion, and the elimination of discrimination. It is expected that every person in this class will treat one another with dignity and respect regardless of gender, sexuality, disability, age, socioeconomic status, ethnicity, race, and culture.

If you feel like your performance in class is being impacted by discrimination or harassment of any kind, please contact your instructor or any of the following:

- Your academic advisor or Graduate Program Coordinator
- Jennifer Nappo, Director of Human Resources, 352-392-0904, jpennacc@ufl.edu
- Curtis Taylor, Associate Dean for Student Affairs, 352-392-2177, taylor@eng.ufl.edu
- Toshikazu Nishida, Associate Dean for Academic Affairs, 352-392-0943, nishida@eng.ufl.edu

Students Requiring Accommodations

Students with disabilities who experience learning barriers and would like to request academic accommodations should connect with the disability Resource Center by visiting https://disability.ufl.edu/get-started/. It is important for students to share their accommodation letter with their instructor and discuss their access needs as early as possible in the semester.

In-Class Recording

Students are allowed to record video or audio of class lectures. However, the purposes for which these recordings may be used are strictly controlled. The only allowable purposes are (1) for personal educational use, (2) in connection with a complaint to the university, or (3) as evidence in, or in preparation for, a criminal or civil proceeding. All other purposes are prohibited. Specifically, students may not publish recorded lectures without the written consent of the instructor.

A "class lecture" is an educational presentation intended to inform or teach enrolled students about a particular subject, including any instructor-led discussions that form part of the presentation, and delivered by any instructor hired or appointed by the University, or by a guest instructor, as part of a University of Florida course. A class lecture does not include lab sessions, student presentations, clinical presentations such as patient history, academic exercises involving solely student participation, assessments (quizzes, tests, exams), field trips, private conversations between students in the class or between a student and the faculty or lecturer during a class session.

Publication without permission of the instructor is prohibited. To "publish" means to share, transmit, circulate, distribute, or provide access to a recording, regardless of format or medium, to another person (or persons), including but not limited to another student within the same class section. Additionally, a recording, or transcript of a recording, is considered published if it is posted on or uploaded to, in whole or in part, any media platform, including but not limited to social media, book, magazine, newspaper, leaflet, or third party note/tutoring services. A student who publishes a recording without written consent may be subject to a civil cause of action instituted by a person injured by the publication and/or discipline under UF Regulation 4.040 Student Honor Code and Student Conduct Code.

Software Use

All faculty, staff, and students of the University are required and expected to obey the laws and legal agreements governing software use. Failure to do so can lead to monetary damages and/or criminal penalties for the individual violator. Because such violations are also against University policies and rules, disciplinary action will be taken as appropriate. We, the members of the University of Florida community, pledge to uphold ourselves and our peers to the highest standards of honesty and integrity.

Student Privacy

There are federal laws protecting your privacy with regards to grades earned in courses and on individual assignments. For more information, please see: https://registrar.ufl.edu/ferpa.html

Campus Resources

Health and Wellness

U Matter, We Care:
Your well-being is important to the University of Florida. The U Matter, We Care initiative is committed to creating a culture of care on our campus by encouraging members of our community to look out for one another and to reach out for help if a member of our community is in need. If you or a friend is in distress, please contact umatter@ufl.edu so that the U Matter, We Care Team can reach out to the student in distress. A nighttime and weekend crisis counselor is available by phone at 352-392-1575. The U Matter, We Care Team can help connect students to the many other helping resources available including, but not limited to, Victim Advocates, Housing staff, and the Counseling and Wellness Center. Please remember that asking for help is a sign of strength. In case of emergency, call 9-1-1.

Counseling and Wellness Center: https://counseling.ufl.edu, and 392-1575; and the University Police Department: 392-1111 or 9-1-1 for emergencies.

Sexual Discrimination, Harassment, Assault, or Violence
If you or a friend has been subjected to sexual discrimination, sexual harassment, sexual assault, or violence contact the Office of Title IX Compliance, located at Yon Hall Room 427, 1908 Stadium Road, (352) 273-1094, title-ix@ufl.edu

Sexual Assault Recovery Services (SARS)
Student Health Care Center, 392-1161.

University Police Department at 392-1111 (or 9-1-1 for emergencies), or http://www.police.ufl.edu/.

Academic Resources

E-learning technical suppor*t*, 352-392-4357 (select option 2) or e-mail to learning-support@ufl.edu. https://elearning.ufl.edu/.

Career Connections Center, Reitz Union, 392-1601. Career assistance and counseling; https://career.ufl.edu.

Library Support, https://uflib.ufl.edu/. Various ways to receive assistance with respect to using the libraries or finding resources.

Teaching Center, Broward Hall, 392-2010 or 392-6420. General study skills and tutoring. https://teachingcenter.ufl.edu/.

Writing Studio, 302 Tigert Hall*,* 846-1138. Help brainstorming, formatting, and writing papers. https://academicresources.clas.ufl.edu/.

Student Complaints Campus*:* https://sccr.dso.ufl.edu/policies/student-honor-code-student-conduct-code/; https://care.dso.ufl.edu.

On-Line Students Complaints*:* https://flexible.dce.ufl.edu/student-complaints/.