Updated 21 FEB 2021 at 2147

Georgia Institute of Technology / School of Interactive Computing
CS 7648 Interactive Robot Learning

Instructor: Asst. Prof. Matthew Gombolay <Matthew.Gombolay@cc.gatech.edu>
Credit: 3-0-3
Pre-requisites: None
Location: Remote
Time: Mon/Wed 12:30 PM – 1:45 PM

Teaching Assistant:	Andrew Silva	Andrew.Silva@gatech.edu (Head TA)
	Letian "Zac" Chen	letian.chen@gatech.edu
	Zheyuan "John" Wang	pjohnwang@gatech.edu

Text Books: Students do not need to purchase any textbooks for this course (see below).

Required:

- Chernova, S., & Thomaz, A. L. (2014). Robot learning from human teachers. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, *8*(3), 1-121. Morgan & Claypool publishers offers this required textbook free, online to GaTech students at this link.
- Russell, S.J. and Norvig, P., 2016. *Artificial Intelligence: A Modern Approach*. Malaysia; Pearson Education Limited. Link

Optional:

- Mitchell, T. M. (1997). Machine learning. *McGraw Hill*.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. MIT press.
- Montgomery, D. C. (2017). *Design and analysis of experiments*. John Wiley & Sons.

Description: Humans can easily learn a new task by observing a demonstrator, asking a few questions, and gaining experience from a brief period of trial, error, and refinement. The field of Learning from Demonstration (LfD) has sought to endow robots with this human-like ability to learn – through observation. In this course, we will cover challenges such as how to 1) learn from novice users (i.e., noisy, sparse demonstrations), 2) intelligently query the demonstrator for additional information (i.e., active learning), 3) compose hierarchical models to learn both high-level tasks and low-level motion primitives, and 4) evaluate the design of LfD systems via human-subject experimentation. The course culminates in a final research project in which students develop and demonstrate their own LfD technique, which will enable them to contribute to the democratization of robotic technology in the home and workplace.

Targeted Students: This course is designed to be accessible to graduate students of all levels and targeted to students who have basic knowledge of calculus, linear algebra, and statistics. LfD draws upon techniques in machine learning, including supervised, unsupervised, and reinforcement learning. While completion of a course dedicated to these topics would be helpful, e.g. CS 7641 and CS 4649, this course will briefly review these topics. This course can benefit students in the College of Computing, School of Psychology, ME, and ECE.

Objectives and Expected Outcomes: This course aims to introduce the basic principles and techniques of LfD. By the end of the course, the students should be able to:

1. Articulate the value of LfD relative to more traditional forms of programming robots,
2. Identify and discuss key results and open problems from prior work in LfD,
3. Apply a variety of machine learning techniques to enable robot LfD,
4. **4.** Design and conduct an experiment to train and validate the design of their own LfD system.

Grading:

Attendance/Participation:	10%
Student lecture/discussion	10%
Homework	20%
Midterm	10%
Term Project	
• Proposal	10%
• Project update	15%
• Report/presentation	25%

Attendance/Participation: 100% attendance is expected to provide students with the best opportunity to learn and contribute to the learning of their peers through constructive interaction. *Excused* absences are permitted and require official institute documentation. We will also allow absences for job interviews on a case-by-case basis; contact Andrew Silva for permission with documentation of the interview *before* the interview occurs. As a component of your attendance grade, we will give pop quizzes during the lecture period that will consist of a single, multiple choice question. A link to the webpage where you can take the quiz will be given out during the lecture, and you will have a short amount of time to complete the quiz before we resume the lecture. You will get 0 points for not trying the quiz; 1 point for trying but not getting the answer correct, and 2 points for trying and getting the answer correct. The purpose of the quiz will be to help keep students engaged with the material. The quiz grades for the entire semester will be averaged and will count for 40% of your attendance/participation grade (i.e., 4% of your overall grade). I will interrupt the lecture to give you time to complete the quiz.

Student-Delivered Lecture/Discussion: Students will be required to find a group of 1-3 students (depending on enrollment), select a date and associated paper from the topical outline (below), and deliver a lecture on the paper, and guide a discussion of that paper. The students should expect to lecture for 40 minutes and lead an in-class discussion of the paper for 35 minutes. Students' lectures will be graded according to the Stanford Checklist for Effective Lecturing.

Homework: Students will be given a total of five (5) take-home "Problem Sets" (PSets), which you shall do collaboratively in your project groups. In addition to these PSets, for each student-delivered lecture/discussion, every student will submit a 300-word review of the paper being discussed, which will be due at the beginning of the lecture.

Midterm: There will be one midterm (but no final exam) that will cover the material taught in the lecture-based portion of the course. The midterm will happen during the lecture time.

Office Hours: We will be using office hours as a mechanism to facilitate live interactions between instructors/TA's and enrolled students. All hours below are Eastern Time.

Instructor Office Hours:	Tuesdays	4 PM – 5 PM	https://bluejeans.com/881076957
	Fridays	9 AM – 10 AM	https://bluejeans.com/484213740
TA Office Hours:	Tuesdays	2 PM – 3 PM	https://bluejeans.com/156515754
	Thursdays	12:30 PM – 1:30 PM	https://bluejeans.com/450486239

Late Policy: Late assignments, except for the final project, will be accepted with 1 letter grade off per 1 day, rounding up (e.g., 1 second and 23 hours 59 minutes 59 seconds each result in 1 letter grade off). No late assignments will be accepted for the final project presentation. In the case of an excused absence, the student and teacher will work to arrange an extension under the guidelines of GaTech.

Term Project: Students will conduct a group (2-3 students) research project based on the topics in this course. Students are encouraged to propose projects relevant to their own research to bring their unique perspectives.

However, the course project must be something the student would not have done during their normal course of research and coursework. The term project consists:

- Proposal – Two-page proposal detailing the project motivation, description, data, plan to collect the data, expected outcome, identification of benchmark, and timeline. The proposal should address the "Heilmeier Catechism". The instructor will provide feedback on the proposal, which should be incorporated into the project update and final deliverable.
- IRB Application – Training and evaluation of LfD requires working with human-subjects. If you want to publish with the data you collect during this project, you *must* receive Internal Review Board (IRB) approval. Each group is encouraged to submit one application. The IRB application is ungraded and only required for students wishing to disseminate the results of their course projects in the broader academic community (e.g., a workshop, conference, or journal paper). **However, project teams who submit an IRB application will receive +1 on their final grade, and teams who get approval for their application will get an additional +1 (total of +2) on their final grade.** Please note: Dr. Gombolay is required to certify any IRB application, and he reserves the right to reject an application if it proposes unethical experimentation or is otherwise deficient.
- Project Update – A one-page summary and a 10-minute, in-class presentation detailing the progress to date on executing the project proposal. Presenters' peers will provide feedback. The final project presentation grade will be influenced by how well students incorporate helpful, reasonable feedback into their projects.
- Project Presentation – An 8-page conference-style paper and a 10-minute, in-class presentation detailing students' contributions for their research project. The paper should have an abstract, introduction, related works, methods, results, discussion, and conclusion. The results section must include a benchmark that the students applied to their data set.

Accommodations: If you are a student with learning needs that require special accommodation, contact the Office of Disability Services at (404) 894-2563 or http://disabilityservices.gatech.edu/, as soon as possible, to discuss your needs and to obtain an accommodations letter. Please e-mail me as soon as possible to set up a time to discuss your learning needs.

Academic Integrity: Georgia Tech aims to cultivate a community based on trust, academic integrity, and honor. Students are expected to act according to the highest ethical standards. For information on Georgia Tech's Academic Honor Code, please visit this link. Any student suspected of cheating or plagiarizing will be reported to the Office of Student Integrity.

Student-Faculty Expectations Agreement: At GaTech, we believe that it is important to strive for an atmosphere of mutual respect, acknowledgement, and responsibility between faculty members and the student body. See this agreement for an articulation of expectations you can have of me and I have of you. Respect for knowledge, hard work, and cordial interactions will help build the environment we seek.

Statement of Intent for Inclusivity: As members of the Georgia Tech community, we are committed to creating a learning environment in which all students feel safe and included. Because we are individuals with varying needs, we are reliant on your feedback to achieve this goal. To that end, we invite you to enter into dialogue with us about the things we can stop, start, and continue doing to make our classroom an environment in which every student feels valued and can engage actively in our learning community.

Amendments: Your instructors reserve the right to make changes as severe weather and other factors necessitate. Any changes will be accompanied by advanced notice from the instructors.

Updated 21 FEB 2021 at 2147

Tentative Topical Outline:

	Date	Topic	Due
Week 01	01/18	*MLK Holiday*	
	01/20	Introduction to Interactive Robot Learning	
Week 02	01/25	Situated Learning	C. & T. – Chapters 1-3; Russel & Norvig Chapters 17 and 21 Thomaz, A., & Breazeal, C. (2008)
	01/27	Supervised Learning, Behavior Cloning (BC), and Dataset Aggregation (DAgger)	Ross, et al. (2011); Spencer et al. (2020) **Pset1 Due 01/29 @ 23:59 EST**
Week 03	02/01	Unsupervised Learning in Learning from Demonstration	Nikolaidis et al. (2015); Gombolay et al. (2017)
	02/03	Inverse Reinforcement Learning – Max Margin, Max Entropy, and Bayesian IRL	Abbeel, P., & Ng, A. Y. (2004) Ziebart, B.D. et al. (2008) Ramachandran & Eyal (2009) **Pset2 Due 02/05 @ 23:59 EST**
Week 04	02/08	Deep, Generative, and Adversarial Versions of Imitation Learning and IRL	Fu et al. (2017) Ho & Ermon (2016)
	02/10	Learning from feedback and advice (TAMER & COACH)	C. & T. – Chapter 6; Knox & Stone (2009); Celemin & Ruiz-del-Solar (2015) **Pset3 Due 02/14 @ 23:59 EST**
Week 05	02/15	Active Learning for LfD	Chernova & Veloso (2009); Schrum & Gombolay (2020)
	02/17	Evaluating LfD	C. & T. – Chapter 7 **Pset4 Due 02/21 @ 23:59 EST**
Week 06	02/22	Low-level Skills (Dynamic Motion Primitives)	C. & T. – Chapter 4
	02/24	Future Directions & Project Kickoff!	C. & T. – Chapter 8 **Pset5 Due 02/28 @ 23:59 EST**
Week 07	03/01	High-level Skills (Hierarchical Task Networks) *Guest Lecture: Dr. David Kent*	C. & T. – Chapter 5 351_david_kent20.pdf on Canvas
	03/03	Natural Language-based Robot Learning *Guest Lecture: Andrew Silva*	TBD
Week 08	03/08	**Midterm**	
↓↓↓ Student Lectures Begin ↓↓↓			
	03/10	Imitation in Human Development	Meltzoff, A. N. (2005) **Project Proposals due 03/12 @ 2359**
Week 09	03/15	Natural Methods for Robot LfD	Nicolescu & Mataric (2003)
	03/17	LfD via Skill Trees	Konidaris, G., et al. (2012)
Week 10	03/22	LfD with Table Tennis	Mülling, K., et al. (2013)
	03/24	***Mid-semester break from instruction***	
Week 11	03/29	Policy Gradient in Imitation Learning	Sun, W., et al. (2017)
	03/31	Active Learning via Imitation Learning	Bullard et al., (2019)
Week 12	04/05	Project Updates – Day I	**Project updates due 04/02 @ 2359**
	04/07	Project Updates – Day II	
Week 13	04/12	One-shot imitation	Duan, Y., et al. (2017)
	04/14	One-shot Imitation via Meta-learning	Yu, T., et al. (2018)
Week 14	04/19	Multi-style reward distillation	Chen et al. (2020a)
	04/21	Self-supervised Reward Regression *Guest Lecture: Letian Chen*	Chen et al. (2020b)
Week 15	04/26	Personal Neural Trees *Guest Lecture: TBD*	Paleja et al. (2020)
	04/28	***No Class (Reading Period)***	
Week 16	04/30	Project Presentations during the exam slot from 11:20am-2:10pm on April 30th	**Project presentations due to Prof. Gombolay by 4/30 @ 0800**

Reading List:

Abbeel, P., & Ng, A. Y. (2004, July). Apprenticeship learning via inverse reinforcement learning. In *Proceedings of International Conference on Machine learning*.

Bullard, K., Schroecker, Y. and Chernova, S., 2019. Active learning within constrained environments through imitation of an expert questioner. arXiv preprint arXiv:1907.00921.

Celemin, C. and Ruiz-del-Solar, J., 2015. COACH: learning continuous actions from corrective advice communicated by humans. In *Proc. International Conference on Advanced Robotics (ICAR) (pp. 581-586).*

Chen, L., Paleja, R., Ghuy, M. and Gombolay, M., 2020. Joint goal and strategy inference across heterogeneous demonstrators via reward network distillation. In Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (pp. 659-668).

Chen, L., Paleja, R. and Gombolay, M., 2020. Learning from suboptimal demonstration via self-supervised reward regression. arXiv preprint arXiv:2010.11723.

Chernova, S., & Veloso, M. (2009). Interactive policy learning through confidence-based autonomy. *Journal of Artificial Intelligence Research*, *34*, 1-25.

Duan, Y., Andrychowicz, M., Stadie, B., Ho, O.J., Schneider, J., Sutskever, I., Abbeel, P. and Zaremba, W., 2017. One-shot imitation learning. In *Advances in neural information processing systems* (pp. 1087-1098).

Fu, J., Luo, K. and Levine, S., 2017. Learning robust rewards with adversarial inverse reinforcement learning. arXiv preprint arXiv:1710.11248.

Gombolay, M., Jensen, R. and Son, S.-H., 2017. Machine learning techniques for analyzing training behavior in serious gaming. *IEEE Transactions on Computational Intelligence and AI in Games*.

Ho, J. and Ermon, S., 2016, December. Generative adversarial imitation learning. In Proceedings of the 30th International Conference on Neural Information Processing Systems (pp. 4572-4580).

Knox, W. B., & Stone, P. (2009). Interactively shaping agents via human reinforcement: The TAMER framework. In *Proceedings of the fifth international conference on Knowledge capture* (pp. 9-16).

Konidaris, G., Kuindersma, S., Grupen, R., & Barto, A. (2012). Robot learning from demonstration by constructing skill trees. *The International Journal of Robotics Research*, *31*(3), 360-375.

Meltzoff, A. N. (2005). "Imitation and other minds: The "Like Me" hypothesis." In S. Hurley and N. Chater (Eds.), Perspectives on Imitation: From Neuroscience to Social Science (Vol 2, pg 55-77), Cambridge, MA. MIT Press.

Mülling, K., Kober, J., Kroemer, O. and Peters, J., 2013. Learning to select and generalize striking movements in robot table tennis. *The International Journal of Robotics Research*, *32*(3), pp.263-279.

Nikolaidis, S., Ramakrishnan, R., Gu, K. and Shah, J., 2015, March. Efficient model learning from joint-action demonstrations for human-robot collaborative tasks. In 2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI) (pp. 189-196). IEEE.

Nicolescu, M. N., & Mataric, M. J. (2003, July). Natural methods for robot task learning: Instructive demonstrations, generalization and practice. In *Proceedings of the second international joint conference on Autonomous agents and multiagent systems* (pp. 241-248). ACM.

Paleja, R., Silva, A., Chen, L. and Gombolay, M., 2020. Interpretable and Personalized Apprenticeship Scheduling: Learning Interpretable Scheduling Policies from Heterogeneous User Demonstrations. Advances in Neural Information Processing Systems, 33.

Ramachandran, D., and Eyal A. (2009). Bayesian inverse reinforcement learning. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)* (pp 2586-2591).

Ross, S., Gordon, G. and Bagnell, D. (2011). A reduction of imitation learning and structured prediction to no-regret online learning. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics* (pp. 627-635).

Schrum, M.L. and Gombolay, M.C., 2019. When Your Robot Breaks: Active Learning During Plant Failure. IEEE Robotics and Automation Letters, 5(2), pp.438-445.

Spencer, J., Choudhury, S., Barnes, M., Schmittle, M., Chiang, M., Ramadge, P. and Srinivasa, S. (2020). Learning from Interventions: Human-robot interaction as both explicit and implicit feedback. In *Proceedings of Robotics: Science and Systems.*

Sun, W., Venkatraman, A., Gordon, G. J., Boots, B., & Bagnell, J. A. (2017, July). Deeply AggreVaTeD: Differentiable Imitation Learning for Sequential Prediction. In *International Conference on Machine Learning* (pp. 3309-3318).

Thomaz, A. L., & Breazeal, C. (2008). Teachable robots: Understanding human teaching behavior to build more effective robot learners. *Artificial Intelligence*, *172*(6-7), 716-737.

Yu, T., Finn, C., Xie, A., Dasari, S., Zhang, T., Abbeel, P. and Levine, S., 2018. One-shot imitation from observing humans via domain-adaptive meta-learning. arXiv preprint arXiv:1802.01557.

Ziebart, B.D., Maas, A., Bagnell, J.A. and Dey, A.K., 2008. Maximum entropy inverse reinforcement learning. In *Proceedings of the Conference on Artificial Intelligence (AAAI).*