TL;DR: I don't see a lot of concrete direction, nor do I know why you actually want to work with these professors.

Statement of Objectives – Massachusetts Institute of Technology PhD in Computer Science

Eisuke Hirota

Objective: ~~In~~ the past few years ~~alone, the public~~ has seen tremendous growth in the capabilities of Artificial Intelligence (AI). While these results are the culmination of 100+ years of combined effort from academia, industry, and government, the everyday layman would only guess that this came out of nowhere. The public fears of the day when AI replaces the workforce or when Terminators and Metal Gear run the streets. This scare is justified; many leading scientists and organizations have voiced for updates in law and regulation revolving AI research and production. Furthermore, just recently in February 2024, California state Senator Scott Wiener proposed Senate Bill 1047, aiming to regulate developer production of AI models. It's clear that as the years progress, innovation will blossom, but without regulation, at what point will innovation backfire and crumble upon itself?

~~I concur with these pushes for AI regulation, and through my research position, I aim to push science tangentially to these wishes.~~ Thus, I am highly interested in developing robot learning methods that enable high value alignment through easy user interaction. In other words, I aim to develop robots which intelligently follow human commands without causing harm or danger. Predominantly, these robots must have a strong comprehension of the task itself – like a human – but limited in privileges to prevent any cautionary behavior. ~~These challenges have motivated my research interests in:~~ "To do so, I aim to solve three core problems in my thesis:..."

1) **Reward Learning through Multiple Modalities:** Current reward learning paradigms focus on only one modality of human feedback ~~(e.g. solely using preference comparison feedback).~~ How can we use multiple modalities together to enable a more natural, user-friendly interface for providing feedback that also reduces the number of feedback queries asked to the user?
2) **Safety-Guaranteed Language-Guided Reward Learning:** Generally speaking, language can be classified into "facts" and "opinions," or "laws" and "suggestions." How can we incorporate "law"-based language feedback into reward learning ~~(perhaps through planning languages or Control Barrier Functions)~~, where current paradigms only use "suggestion"-based language feedback, to guarantee safe robot behavior?
3) **Play Data for Feature Representation Learning:** Human-teleoperated robot play is a cheap source of collecting a task-agnostic dataset. ~~Despite so,~~ is it possible ~~that we can~~ to extract implicit feature representations to pretrain a reward model and understand general human nuances?

~~My interest in these future directions were cultivated through experiencing various research opportunities which I describe below. These past opportunities have prepared me for a path in PhD.~~

Research Experience: "...my research experience has uniquely prepared me to solve these problems..."

~~I first started my RL journey working in the AI4CE Lab under Professor Chen Feng:~~

~~Localization + Planning~~: Traditionally, when tasked in a dynamic environment, a robot must effectively localize itself, accounting for the movement of landmarks independent of the robot. In civil engineering tasks such as construction however, each robot building step causes a shift in the environment, directly causing the environment to become dynamic. Hence, a significant challenge for robots is to simultaneously localize and plan a sequence of actions in an environment that changes dependent on the robot's actions. In the AI4CE Lab under Dr. Chen Feng, ~~my work focused on benchmarking current **state-of-the-art Reinforcement Learning** (RL) algorithms in such a task~~. Observing poor performance from these algorithms (on-average 28% success rate), my work theorized that an end-to-end model is insufficient in achieving high performance. Hence, we proposed **DRQN+L-Net**, an algorithm that **decouples the localization and planning** into two separate models, which performed almost twice as better (52% success rate) than baselines. This work was accepted at **ICLR 2023**, and I presented the **poster** and **video** for this conference.

~~From the accumulated experience in RL, I continued applying my knowledge and gain more experience in the AIR Lab under Professor Shiqi Zhang, working on the following research directions:~~

~~Force Estimator:~~ I helped develop a novel **force estimator** to estimate the external forces applied onto a quadruped (achieving up to 80% higher accuracy than the onboard accelerometer), trained using **supervised learning**. Through

the addition of this force estimator, our system accomplishes two things: (1) we feed the estimated forces into the policy, enabling a **robust, force-tolerant quadruped controller**, and (2) we can **classify the direction to which the force** is being applied. Through (2) we developed a **guide-dog system for the visually impaired** emulating how real guide dogs work. The user and the quadruped bidirectionally communicate: the user applies tugs to a leash connected to the quadruped to communicate which direction to navigate, and the quadruped locomotes certain directions, pulling the user away from obstacles. This work was accepted at **CoRL 2023**, and I had the opportunity to present a **poster** at the **SUNY AI Symposium 2023**. Additionally, our work had positive media coverage (e.g. Wired, Daily Beast, Spectrum News, etc.) and enabled us the concurrent opportunity to collaborate with Guiding Eyes for the Blind, the biggest guide dog association in New York.

Non-prehensile Quadruped Manipulation: I lead my own research direction surrounding **non-prehensile quadruped manipulation.** Given a quadruped with capabilities to push around objects in its environment, we can deploy the quadruped for rearrangement tasks. Quadrupeds have the agility to traverse environments that are harmful or inaccessible to disabled or elderly people, such as glass-scattered floors or tight spaces underneath tables. Thus, we can train a controller using **curriculum learning** and **Hierarchical RL** – a low-level employee policy that controls its legs based on end-effector position, commanded by a high-level manager policy that perceives its surroundings using the onboard RGB-D camera to develop a plan in manipulating objects. (This project is currently on-pause due to our robot breaking… how to phrase??)

Knowledge-Based RL: In quadruped locomotion, **different gaits** (foot-floor contact patterns) have trade-offs between speed and energy. Given we as humans have such knowledge, we can integrate this into the RL training schema to enable easier training. I helped train such gaits using **Markov Decision Processes with Reward Machines** (MDPRMs), which adds a **simple finite-state automata** to the traditional MDP to output extra rewards when certain sequences of foot contacts occur, which enabled more **Sample-Efficient RL**. This work was accepted at **ICAPS 2024**.

From here, I wanted to branch out – go down a specific path in RL. I came to realize that traditional RL requires much engineering effort in designing a well-working reward function. At the same time, I was self-studying Inverse RL (and more specifically Preference-Based RL), hence, I decided to step into this new road in LiraLab under Professor Erdem Bıyık.

Active Reward Learning with Comparative Language Feedback: In Active Reward Learning, a robot queries humans for feedback regarding its performance in a certain task and aims to ask the least amount of queries possible to keep cost low. Previous solutions consist of Volume Removal and Information Gain, however, they only account for scenarios where the human provide comparison feedback, a type of feedback that requires a human to rank two or more performances. Following up on a paper from our lab recently accepted at CoRL 2024, the user can now give language feedback after only seeing one performance; hence, my work extends previous work in which I **derived an Information Gain solution for language feedback**. This solution requires the use of **Approximate Bayesian Inference**; thus, we test four different algorithms (Metropolis-Hastings, Gibbs, Laplace, Expectation Propagation) and evaluate our results in both simulation and in user studies, observing a **performance boost of [some metric].** Currently a journal paper is under review in **RA-L** on which I am the first author. (this project is not done yet, but it should be finished by latest, Early December)

VLMs for Identifying Gestures as Commands: Vision-Language Models (VLMs) have been trained on tons of data and can thus be leveraged for Human-Robot Interaction (HRI) purposes. The goal of this project is to use a VLM to recognize a gesture, and then to **command robots through such gestures**. I trained a robot policy using **offline RL** and **motion primitives** which follows the user commands. From this project, I have realized that there exist other types of human feedback that I have yet to consider. This work is currently in progress, and we hope to submit to IROS 2025.

Future: My interests are highly aligned with professors **Andreea Bobu**, **Julie Shah**, and **Brian Williams**. My overarching goal is to introduce methods that probabilistically guarantee desired behavior from robots in a user-friendly, intuitive manner. ~~MIT facilitates an exceptional community dedicated to value-aligned robot learning, and I hope to be a part of it.~~ I appreciate your consideration.

Research and Technical
What field(s) of study would you like to do your graduate research in, and what are your reasons for that choice? Please discuss the problems/issues that you would like to address, the type of research that interests you (e.g., experimental, theoretical, or issue oriented), and which faculty/PIs within the MIT AA Department could help support this research. (250 words)

Recently, the public has seen tremendous achievements from AI. While these results stem from the combined effort of academia, industry, and government, the everyday layman worries that this came out of nowhere. As a result, many have voiced for updates in regulation, e.g. Senate Bill 1047. Innovation will blossom in the upcoming years, but without regulation, when will it all collapse?

~~I concur with these pushes~~ for regulation, and ~~through my position, I aim to further science tangentially to these wishes. Thus, I am highly interested in developing~~ robot learning methods that enable high value alignment through easy user interaction. ~~This has motivated my inte~~rests in: — I aim to develop

The key challenges in this area are....

^ this start is not good enough anymore. For NSF GRFP sure, because it's practically a grant. But for PhD admissions, I don't think this is good. I should go straight into value alignment.

Let's see a faculty talk or something, that way I can see exactly how to formulate a "vision"
Probably something like:

- Provocative, and no misleading. First sentence must direct audience towards what we are talking about
- Mention a clear gap in current value alignment or reward learning. Not something incremental, but something bigger. To get to that something bigger, what steps can we take?

Try to integrate how Julie shah's work intertwines with our ideas. Rn it seems like these 3 ideas are good for andreea but maybe not Julie or the third prof.

To deploy robots in our everyday lives, they must comprehend our preferences and act accordingly to our expectations. If one were to request a robot to clean the room, the robot should not trash a sweater solely because it was on the floor. Similarly, robots must not perform harm to life solely because such action is optimal according to some evaluation.

Reward learning methods learn a human's reward function to ensure robot behavior is aligned with human expectations. To bridge the gap between our current robots and the robots we wish to see in the future, I will tackle the following lines of work: } 3 works

goal

1) **Easy Reward Learning through Multiple Modalities:** Current reward learning paradigms leverage only one modality of human feedback. How can we string multiple modalities together for friendlier interaction and to improve sample efficiency?
2) **Safe Easy Reward Learning through CBFs:** Can we incorporate Control Barrier Functions with multiple modalities to ensure safe, easy reward learning?
3) **Play Data for Feature Representation Learning:** Human-teleoperated robot play is a cheap source of collecting a task-agnostic dataset. Is it possible to extract feature representations from these datasets to pretrain a reward model and understand general human nuances?

My interests are highly aligned with professors **Andreea Bobu**, **Julie Shah**, and **First Last**. My overarching goal is to push the theory behind value alignment. ~~MIT facilitates an exceptional community dedicated to value-aligned robot learning, and I hope to be a part of it. I~~ appreciate your consideration.

Professional Experience and Objectives
Discuss your academic, research, work, and/or field experiences to date, and explain how those have helped shape your objectives for graduate school and your long-term professional goals. (250 words)

(should I be shaping this to talk about aeroastro? Like how it impacts that. Cuz tbh while my statements are solid, do they really matter to aeroastro -> probably not)

Things you can write in here:
- What are promising techniques for approaching your challenges?
- Which parts of your proposed advisors' research aligns with your goal?

By the end of my career, ~~During retirement and reflecting on my life's work~~, I hope to have pushed value alignment theory and influenced policies for AI regulation. These goals are the culmination of my research experiences, tackling different domains and exchanging ideas with diverse mindsets.

Working with Professor Chen Feng at NYU and Professor Shiqi Zhang at Binghamton University, I encountered issue-oriented research about robots for civil engineering and guide dog domains, respectively. Such work, especially ~~robotic guide dogs for the visually impaired~~, persuaded me to develop robots that directly impact human lives, promoting a safe and accessible world. Currently, I work with Professor Erdem Bıyık at USC and have been exposed to a theoretical approach to robot learning – here, I derived a novel Active Reward Learning equation that integrates language feedback. Through this experience, I gained a keen interest in value alignment and robots that intelligently meet human expectations.

This exposure to robots that ~~(1)~~ help human lives and ~~(2)~~ act aligned with values have shaped my goals for my PhD career and onwards. Not only will I develop novel, impactful contributions towards value alignment, but I will also use my research as demonstrations that spark discussion in AI regulation. Only just recently in 2024 has AI regulation started with Senate Bill 1047 – two years after ChatGPT became publicly available. Clearly, our current policies and systems cannot keep up with the exponential growth of AI, thus, I will leverage my research in both PhD and onwards as a platform to spur modernization in administration.

My main objective is to lay down more foundation behind value alignment for robotics. During my PhD and career afterwards, I aim to ensure that robots can easily comprehend humans and behave safely. These goals are the culmination of my various research experiences, tackling problems in different domains and exchanging ideas with diverse mindsets. Throughout my undergraduate career, I applied Reinforcement Learning (RL) methods for robotics. Working with Professor Chen Feng at NYU and Professor Shiqi Zhang at Binghamton University, I encountered issue-oriented research in civil engineering and guide dog domains, respectively. ~~Such research is crucial in deploying robotics to everyday life,~~ but I soon ~~found myself glancing~~ towards contributions in theory. Regardless of the domain, whether simulation or real-world, innovations in RL theory ~~was,~~ in my mind, foundationally more crucial. Concurrently, I encountered obstacles in a self-lead project regarding RL for quadrupedal non-prehensile manipulation, inspiring me to expose myself to different schemas for robot learning – I then had the opportunity to work with Professor Erdem Bıyık at USC. Here, I derived a novel Active Reward Learning equation that theoretically guarantees faster reward learning using language feedback. ~~Immersed in this theory-backed approach~~, I have decided to continue pursuing this methodology for my PhD. I aim to continue leveraging this style, pursuing an industrial research career in the future. A long-term goal of mine is to deploy robots in homes, and achieving this is the most feasible through collaborating with industry and ensuring such robots are safe and intelligent.

Do you have more words? I think a good architecture would be:

- I learned a Big Lesson in undergrad.
- Because of that, I care a lot about a Hard Problem.
- The Ph.D. will allow me to apply Special Technique to advance Hard Problem.
- In the long term, I hope to continue solving Hard Problem.
- My work will have Big Impact on Society™.