CPS Medium: Humanizing Autonomy - Towards Behavior Guided Autonomous Vehicles

May 8, 2018

CPS: Medium: Humanizing Autonomy - Towards Behavior Guided Autonomous Vehicles

Autonomous cars are not just about the technology. They are about freedom of mobility, and a whole set of experiences that will literally and figuratively move people in new ways. While the promise of self-driving cars is attractive, applying it in a meaningful and coherent way still remains a major challenge. Recent accidents involving autonomous (or semi-autonomous) cars that resulted in fatalities of either the driver, or in one case in one case the pedestrian, have exposed major holes regarding the interaction between the car and its driver/passengers. We know now, that it may not always be possible for the driver to take over the control from the autonomous vehicle (AV) at any stage.Therefore, before it's too late, we need to reevaluate our approach towards autonomy, by seeking answers to questions that put humans at the center stage. As journeys become fully automated, the experience itself will need to become more human.

The goal of the proposed research is to utilize both human and machine advantages to humanize autonomy ♠[1] by instilling the beneficial nuance of human behavior and trust, while exploiting technological and safety benefits of an AV. The proposed research aims to bring human factors such as emotions, behaviors, and trust into the autonomous loop, where AVs can enhance the passenger(s) experience, safety, and comfort. Human behavior and emotions are highly dynamic and are different among individuals based on their previous experiences, environmental factors (e.g., weather, lighting), societal factors, and internal factors (e.g., physiological changes). Currently, AVs (as well as many other autonomous systems) lack in have any sensing and optimization capability according to passenger(s) real-time behavioral and emotional changes. Additionally, trust is another core human needs that the AV must establish and defend. However, the nature of trust in the vehicle or the autonomous system varies from individual to individual as well. In this research, we will conduct real-life as well as simulation-based experimental studies to identify (1) the association between human emotional attributes and contextual interaction observed from each individual, (2) a taxonomy of emotional and behavioral traits as they relate to the internal and external triggers as well as personalized traits and (3) perspectives of trust in autonomous vehicles from real people, rather than working on assumptions, and (4) the intervention and communication strategies that could be automated by AV to meet passengers(s) need and enhance their ”driving/journey” experience.

Transparency and communication are critical to building trust. To establish user understanding of the system and its capabilities the interface must communicate clearly, and transparently - by revealing what the car sees, what the system is currently doing, what it intends to do in response to environmental conditions and why. Safety is not primarily just a functional consideration, it is also emotional. We propose that the issues of functional and emotional design for autonomous vehicles should be tackled together. For example, emotional down-regulation could be used when passengers might be facing an upsetting or frustrating situation – for instance, a delay in travel. Here the AV could sense the frustration and then down-regulate through voice prompts. When you're jumping in and out of different AVs, a consistent and personal experience will be vital for successful adoption. We're concerned with a person's ability, and even right, to make their own decisions, and come and go as they please. Not about how clever cars are without human drivers.

Intellectual Merit: The contributions of this proposal are as follows:

1. Behavior guided autonomy

   Arsalan add

   Lu add

2. Reachability analysis and control synthesis for safety region estimation

   Nicola to add brief description.

[1] NB: what does it mean to humanize autonomy? We need to be clear here

3. Feedback design: We propose building a framework, that provides both the intent of the autonomous car and an explanation for its behavior to the driver/passenger by augmenting existing user interfaces. We propose a scenario-based trust modeling method in which by varying the degrees of contextual feedback provided to the user, we can measure how the human trust in the system varies under different traffic situations. We then create a 'trust' profile for the driver and the autonomous driving behavior can be molded to conform to the trust profile of the user, but only within the confines of overall safety. We show how local interpretability can be used to explain actions of an autonomous car, the operation of which is very complex and hidden from the driver/passenger. For example the AV's action of stopping suddenly is not enough – the user wants to know why. Without any explanation or context, the user will panic. Our proposed framework can generate such explanations.

Broader Impact: Autonomous control and decision systems are forming the basis for significant pieces of our nation's critical infrastructure. In particular, autonomous vehicles present direct, and urgent safety-critical challenges. If successful, the research outcomes will have the following impacts: (a) be a valuable contribution towards increasing the overall safety of fully autonomous vehicles, which are likely to become ubiquitous in the near future, (b) the underlying frameworks of generating local explanations from sensor data, and safe operation through reachability analysis can help enhance a large scope of autonomy including but not limited to autonomous vehicles, robotics, aircraft autopilots, and automatic surgery equipment, and (c) Leveraging human trust and emotional behavior to help enhance the capabilities of autonomous vehicles and also facilitate the deployment of autonomous vehicles in the real world.
Educational Impact: The PI's will develop curriculum including course lectures and hands-on projects related to autonomous driving. The PIs are very vested in promoting and employing undergraduate researchers. They will continue developing and participating in research programs to involve K-12 students into lab research and inspire their interests in autonomy and human factors based upon the hardware and software developed in the proposed research. The PIs will also actively disseminate the research outcomes through outreach in both academia and automotive industries.

Contents

A. Project Summary A-1

B. Table of Contents B-1

C. Project Description C-1

1 Introduction **C-1**

2 Research Description **C-2**
2.1 Intellectual merit . . . C-2
2.2 Research Thrust 1: Behavior modeling . . . C-2
2.3 Research Thrust 2: Reachability analysis based safety region estimation . . . C-5
2.4 Online reachability based safety region calculation: . . . C-5
2.5 *Runtime Verification for Safety-Aware Autonomous Vehicle Operations:* . . . C-7
2.6 Behaviour-guided safe trajectory synthesis . . . C-7
2.7 Research Thrust 3: Behavior-guided feedback design . . . C-8

3 Experimentation and Evaluation Plan **C-12**
3.1 Scenario-based behavior and trust modeling . . . C-12
3.2 Test-beds for reachability and control synthesis experiments . . . C-14

4 Broader Impacts **C-15**
4.1 Improving Education on Autonomy and Cyber-Physical Systems: . . . C-16
4.2 K-12 Impacts: . . . C-16
4.3 Graduate/Undergraduate Students and Outreach Effort: . . . C-17
4.4 Dissemination Impacts: . . . C-17

5 Project Management and Collaboration Plan **C-17**
5.1 Roles and responsibilities: . . . C-18
5.2 Risks and mitigation plans: . . . C-19
5.3 Results from Prior NSF Support . . . C-19

D. Bibliography D-1

E. Data Management Plan E-1

G. Facilities, Equipment, and Other Resources. G-1

CPS: Medium: Humanizing Autonomy - Towards Behavior Guided Autonomous Vehicles

1 Introduction

Two fatal self-driving-car accidents in March 2018 (Uber [1] and Tesla [2]) have cast doubt in the general public on whether autonomous vehicles can become the future of personal transport. The Tesla accident happened in broad daylight in pretty much perfect driving conditions. Just days before that, an autonomous car prototype by Uber, equipped with LIDAR(s) completely vehicle failed to detect a pedestrian crossing the road at night. The autonomous vehicles industry, and research community, faces an uphill battle in convincing the public that self-driving cars are safer than human drivers. As the framework for mobility in the United States begins to shift from one of personally-owned, manually-driven vehicles to one of a shared and partially automated fleet, established driver perceptions about their confidence, and trust in various vehicle technologies is now more than ever critical.

Everyone is talking about autonomous vehicles. From automotive manufacturers, to consumer electronic giants, to software engineers, and academic researchers, self-driving cars are at the forefront of everyone's imagination. There were more autonomous driving concepts at the International Consumer Electronics Show (CES) 2018 than ever before [3], demonstrating the seriousness, and commitment of many industries to bring autonomous cars to market. Indeed AVs can make a meaningful difference to the world, enabling a new level of mobility, independence, and safety for all. This has been covered in reams of papers and many 1,000s of articles and news stories all over the globe. From questions of technological feasibility to thorny ethical dilemmas [4], it's been approached from many angles. However there are aspects that haven't yet been fully addressed - what do people want and need from AVs and how best to design for the user experiences - what about those human factors? For example, we hypothesize that explaining why the vehicle is doing what it is doing to the passenger can increase their level of comfort in the autonomous vehicle. The industry is preoccupied with the race to make their cars be as smart as possible and more safe than human drivers. Their approach to autonomy is imbalanced - there is too much focus on the discrete technologies that will enable it, with little regard for the powerful human factors involved. Naturally, one of the most important things to get right in AVs is safety. These vehicles must be incredibly safe, in fact safer, statistically and practically, than manually driven cars on the roads today if they're to be of much use to us. According to a Deloitte study [5], trust and human behavior appears to be the biggest roadblock to adoption of self-driving cars in every country surveyed. The US falls roughly in the middle, where nearly three-quarters of consumers (74%) believe that fully-autonomous vehicles will not be safe.

What is important is for car makers to rethink the design process to transform the entire user experience for the betterment, and safety of everyone. With that in mind - the goal of the proposed research is to utilize both human, and machine advantages to humanize autonomy by instilling the beneficial nuances of human behavior, emotion, and trust, while leveraging the technological and safety benefits of autonomous vehicles. Beyond getting in and enjoying the ride – there are far more factors, details and nuances to be considered. What are the human factors at play here, and how can we design the best autonomous driving experience?

While everyone wants AVs to be safe, the interpretation of safety, and the nature of trust in the vehicle or the system changes from person to person. Even for the same person, their perception of the trust in the autonomous vehicle varies based on their emotions and behavior. Some perceive an autonomous vehicle in the same light as a malfunctioning robot, running amok around town, striking fear into passengers, fellow motorists, and pedestrians. Then there are those who believe and expect that the AV will always "have your back" and will be totally safe. A quick YouTube search, reveals many videos where people are aware the systems have limitations, but still push them further than their intended use, operating pilot assist systems

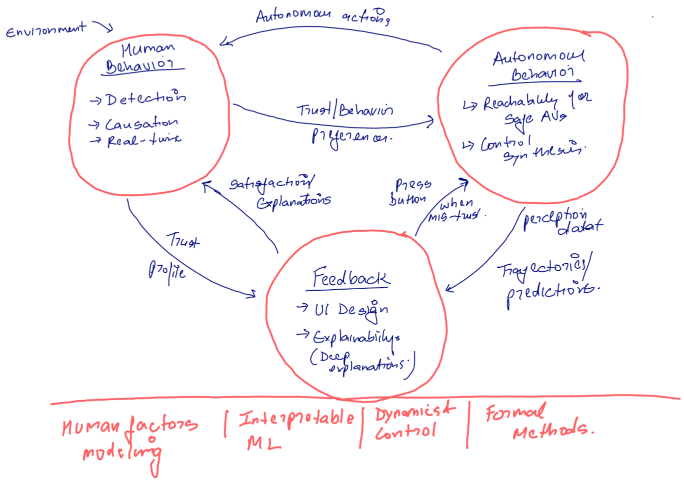

Figure 1: Placeholder overview figure

on roads or situations when they shouldn't. *All of this boils down to one thing – human behavior* Trust is one of the core human needs that the autonomous vehicle must establish and defend if the technology is to be adopted at all. For the technology to be adopted, it must be trusted, and to do that, the human behavior and emotional needs must be taken into consideration.

2 Research Description

2.1 Intellectual merit

The intellectual merit of the proposed work is to investigate models and algorithms, that can adapt autonomy in appropriate ways by leveraging human behavior. Using autonomous driving as the basis of this research the intellectual merit of this work lies in:

1. the behavioral, and emotional modeling for human drivers [Thrust 1]
2. a novel online reachability analysis for safe path planning, and control synthesis for autonomous vehicles which can take user preferences into account [Thrust 2] , and
3. deep neural networks for generating natural language explanations of the actions of the autonomous vehicle, and a user interface design for conveying the intentions of the vehicle to the passengers [Thrust 3] .

The research spanning across the three thrusts is innovative in that lies at the confluence of human factors modeling, control synthesis, formal methods, reinforcement learning, and deep learning, to close the loop between human behavior and emotions and autonomous planning and control. The three research thrusts and their interconnections are described next.

2.2 Research Thrust 1: Behavior modeling

Being able to read other's emotional cues not only allows us to understand how people are feeling at a given time, but it also helps us to predict how they will respond in different scenarios. If an autonomous vehicle can

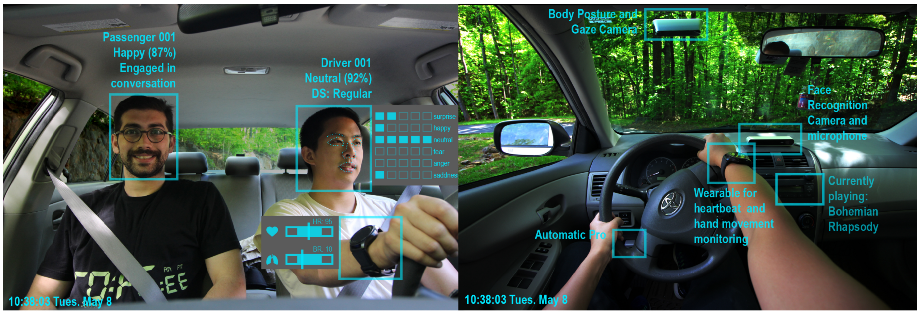

Figure 2: Overview of instrumentation of vehicles and the collected information from participants

detect the emotional cues of the driver, and more importantly respond to them, then it can develop "human-like" trust with the driver. Research suggests that emotions are normally associated with specific events or occurrences (cite), and they can significantly influence our thoughts and behaviors. This research thrust aims to enhance driver experience, safety, and trust in the AV. We do so by automatically and accurately detecting, assessing, and managing the passengers behaviors and emotions.

Several Naturalistic Driving Studies (NDS) have studied human behavior in vehicles in different safety-related conditions, with the specific emphasis towards crash or near-crash incidents. In these studies, cars were equipped with sensors and devices that continuously monitor various aspects of driving behavior such as vehicle acceleration, deceleration, location, and speed, driver-related information such as eye, head and hand movement, and environmental conditions such as traffic and weather conditions (CITE). For instance, in (CITE - European paper) researchers identified that certain driver-personality traits are prone to commiting speeding violations; or bad behaviors may cluster together, where drivers who perform one type of risky behavior are more likely to engage in other types of risky behaviors. Although these studies provide significant insights on the driver behavior in different scenarios and contextual settings, they specifically focus only on safety-related behaviors that result in crashes or near-crash incidents; Due to complexities related to human behaviors (i.e., dynamic changes in human behaviors) many of the NDS were limited to short-term monitoring and as a result, lacked in identifying how the driver behaviors may change based on various contextual, environmental, and social settings. In a recent study researchers are using computer vision and voice-processing techniques to monitor drivers body posture, eye movement, and emotions to predict driver's frustration and gaze region that may result in unsafe behaviors while driving semi-automated and fully-automated vehicles (CITE behavioral impact of drivers roles and other MIT work).

At higher levels of autonomy, behavioral and emotional information not only are extremely useful for increasing safety of AVs, they can also be applied towards optimizing control actions of AVs, and most importantly, enhancing the driving-experience for the driver and/or passenger(s); for instance, by identifying that the driver is feeling "happy", AVs can optimize their route selection in order to increase the drivers positive emotions, potentially enhancing the driver's mental health. Additionally, by optimizing the decision making of automated systems around the user's specific preferential and comfort profiles, research suggests that the automated system can better gain user trust (cite). There exists a lack of understanding on the causation of behaviors as a result of environmental, emotional, physiological, and social factors. In this research thrust, we aim to understand how environmental, physiological, and social factors may influence and/or cause certain behaviors and emotions in the driver and passengers. When we break down this hypothesis further, the following are the questions of interest:

1. What is the taxonomy of driving behaviors and emotions based on environmental conditions, social

interactions, and physiological changes?

2. How can driving behavior and emotions be non-intrusively and with least number of sensors be detected?
3. Do people have differentiating "trust" profiles according to specific behaviors and contextual conditions?

While existing literature has considered each of the environmental, social, and physiological factors separately, the challenge lies in learning causal relationships between these factos, and how they influence emotion and behaviors of the drivers. Specifically, we will gather data and build models for (1) factors that influence specific human emotions and (2) the downstream consequences of particular emotions on their behaviors. For instance, we will gather data to asses what environmental (e.g., weather, thermal conditions, lighting, noise) and social factors (e.g., social interaction), physiological factors (e.g.,arm movement) trigger particular emotions. One outcome of this work will also be able to develop a library of emotions, behaviors, social interactions, and environmental changes. Our preliminary investigations have revealed the following taxonomy for driver behaviors:
[**Risky driving**] is often addressed as one of the caveats of autonomous vehicles. For example, if drivers sufficiently trust vehicles, some drivers will likely follow the front car closely or drive faster than usual.
[**Anxious driving**] style. To better understand this anxious drivers, we need to look at the plausible sources of anxiety regarding autonomous vehicles. On one hand, people can be anxious about their poor driving skill. Fully automated vehicles can solve this issue. On the other hand, people can be anxious about not being able to control over something. In this case, we can design the interface so that drivers can make the driver feel more in control.
[**Dissociated driving**] includes errors and mistakes (e.g., errors in gear shift or lights). If this stems from poor (or inexperienced) driving, again fully autonomous vehicles can solve this issue. Drivers do not need to differentiate gear shifts or calculate route themselves. However, detecting the casual effects of dissociated driving and being able to predict this behavior profile remains an open challenge.
[**Careful driving**] is referred to "better safe than sorry". To fulfill this type of drivers' motivation with any type of autonomous vehicles, we can provide more effective and robust monitoring interfaces rather than providing a number of distracting tasks. Such drivers likely want to have higher situation awareness and will be satisfied with the more completed monitoring mechanisms.

We plan to gather the required data both through high fidelity simulations on a full scale diving simulator (Section **??**) and by instrumenting participant-owned vehicles, we will collect and analyze the impact of different factors on driver/passenger behaviors and emotions. To collect environmental-related information, Automatic Pro (CITE) will be used to monitor the car's speed and location. Automatic Pro will be also used to collect behavioral information such as driver's acceleration and deceleration rates Cameras will be used to capture outdoor conditions such as traffic, weather, pedestrians, signs etc. as well as indoor conditions such as social interactions, drivers and passenger(s)'s body posture, face, head and hand/arm movement. Ambient-condition sensors will be used to measure thermal settings, indoor and outdoor noise levels, and lighting levels. Audio recorders will be used to capture social interactions as well as any other audio sources (i.e., music, radio, audio books, etc.). Wearables such as Samsung Gear or Fitbit will be used to measure the driver and passengers physiological data (i.e., heart rate and hand/arm movement). Through face recognition software (i.e., OpenSmile, iMotion, and Google API) recorded videos will be processed to create a time-stamped emotional states for the driver and passengers; The audio information as well as physiological information can also be used to identify the driver/passengers emotional states. Figure 2 provides an overview of the sensors and data that will be collected from the drivers and passenger(s).

By having an understanding about the causation of certain behaviors and emotions, behavioral and emotional models can be developed and integrated into AVs systems where they can optimize the control actions as well as feedback/communication strategies. We will utilize data-driven algorithms that aggregate the information extracted from vision and audio with physiological and environmental information. Through clustering techniques and machine learning algorithms (i.e., CNL and HMM), we will define the behavioral models described above, as a function of emotional traits, environmental conditions, social interactions, and physiological measures.The resulting behavior models will be passed on as a driver profile/preference input to the trajectory planning module of the autonomous vehicle in Thrust 2 (Section 2.3), and also to trigger a natural language explanation for the car's actions according to Thrust 3 (Section 2.7) While we will largely rely on exiting data-driven and machine learning models and techniques for individual factors modeling and classification, the novelty of this human factors research lies in establishing causal relationships between environmental, emotional, social, and physiological data and driving behavior, and to be able to predict and/or classify driver behavior in real time as an input to the autonomous driving and feedback system.

2.3 Research Thrust 2: Reachability analysis based safety region estimation

Safe driving behavior is not always a binary decision, but rather can be considered as a region. In this thrust we propose to develop novel methods of calculating safety regions of the autonomous vehicle in and online manner, and developing safe trajectories for the autonomous vehicle which can take the human emotion, and behavior preferences into account. Consider the following example: while driving in the left most lane on a freeway, next to an concrete barrier and in the absence of a left shoulder, we notice that majority of human drivers, do not drive in the center of the lane, but rather drive towards the outer edge (away from the barrier). This is natural, since we understand the momentum of the car and do not want to drive too close to the barrier. However, a semi-autonomous vehicle today which is capable of lane-keep assist, will tend to drive the car in the middle of the lane. While it is technically, safer to drive in the middle of the lane, it makes many of the drivers/passengers of the semi-autonomous vehicle anxious of the car being too close to the wall. In this situation, if the human preference is to be away from the wall, the autonomous car could take that into account while planning trajectories, provided it is safe to do, i.e in this case when the adjacent lane is not occupied. To achieve this we need a way to analyze the safety region of the autonomous vehicle in real time, and generate trajectories in a safe and reliable manner.

This research thrust is divided in two main parts, as outlined in Figure 3. We start by proposing a reachability-based approach in which we leverage knowledge about the system dynamics, and environmental uncertainties to determine the future states over a finite time horizon. The computed reachable sets are used to determine the first time that a safety-critical event may occur. Based on these events, we then design a control synthesis approach to adapt and correct the behavior of the autonomous vehicle to maintain safety and increase trust with the human.

2.4 Online reachability based safety region calculation:

Traditional model-predictive or finite horizon controllers (MPC) [6, 7, 8] can estimate future states and inputs of a system, given a correct knowledge of its model. On the other hand, reachability analysis (RA) [9, 10, **?**] provides regions that can be reached by the system when applying a sequence of inputs and assuming different uncertainties like process noise, model uncertainties, sensor noise and jitter, delays, and external disturbances. For this reason RA is typically used to monitor and assess if and when safety conditions can be violated [11, 12, 13]. We propose to research the use of RA to predict reachable states of vehicles under different uncertainties and assess safety constraints like collision avoidance and trajectory tracking. The key idea is that the context (e.g., obstacle, human behavior, and other vehicles configuration) in which the AV is operating dictates the reaction time to switch to a different mode of operation. Thus, predicting the situations

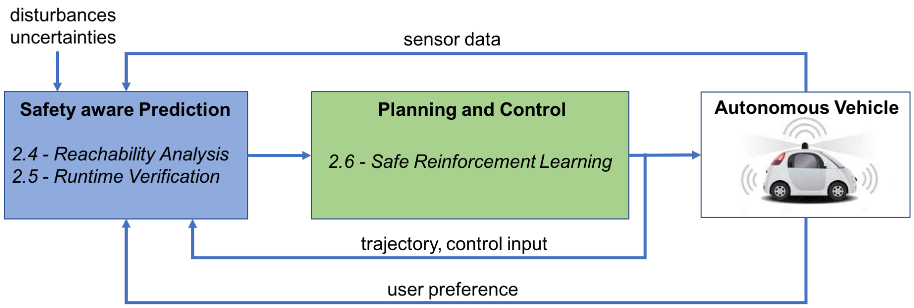

Figure 3: Overview of the proposed reachability based safety assessment and recovery

that can occur in the near future will allow a better safety assessment, and lead to better planning of actions which can conform the user preferences provided as an outcome of Thrust 1 (Section 2.2).

Consider a vehicle modeled generally as a non-linear system of the form $\dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{u}_m)$ with $\boldsymbol{x}$ the state vector of the system and $\boldsymbol{u}_m$ the input vector. Assume that the system can be linearized around a point $(\boldsymbol{x}^*, \boldsymbol{u}^*)$ (that can be the current operating point) and obtain the following, discretized with sampling time t_s, state space representation $\boldsymbol{x}(k+1) = \boldsymbol{A}\boldsymbol{x}(k) + \boldsymbol{B}\boldsymbol{u}(k) = \boldsymbol{A}\boldsymbol{x}(k) + \boldsymbol{B}_m\boldsymbol{u}_m(k) + \boldsymbol{B}_I f(\boldsymbol{x}^*, \boldsymbol{u}^*) + \boldsymbol{B}_d\, \boldsymbol{d}$ where $\boldsymbol{d}$ is an external disturbance. The challenge is to compute the future states of the system using RA in an online manner, and under various uncertainties. In this work we propose a machine learning-based approach to minimize runtime computation time. A reachable set computed at time t_0 for a future time t_f and represented by $R(\boldsymbol{x}_0, \boldsymbol{u}(t), t_f)$ is an ellipsoid ϵ that contains all the states $\boldsymbol{x}$ reachable at a future time $t_f > t_0$ where the initial set $\boldsymbol{x}_0 \in \epsilon(\boldsymbol{x}_0, \boldsymbol{X}_0)$ is bounded by an ellipsoid with center $\boldsymbol{x}_0$ and shape matrix $\boldsymbol{X}_0$ and the input $\boldsymbol{u}(t) \in \epsilon(\boldsymbol{u}(t), \boldsymbol{U})$ is bounded by an ellipsoid with center $\boldsymbol{u}(t)$ and shape matrix $\boldsymbol{U}$. A reachable tube $R(\boldsymbol{x}_0, \boldsymbol{u}(t), [0, T])$ is then defined as the set of all reachable sets over the time interval $\Delta T = [0, T]$, [14, 15]. The external bound for the reach set at time t starting from time t_0 is calculated based on the initial state ellipsoid, the plant model, and the input ellipsoid, as follows:

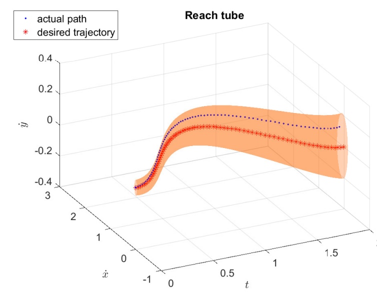

Figure 4: Velocity Reachable tube for a quadrotor following a straight line trajectory for 2s.

$$R^+(t, t_0, \epsilon(\boldsymbol{x}_0, \boldsymbol{X}_0)) = \boldsymbol{\Phi}(t, t_0)\epsilon(\boldsymbol{x}_0, \boldsymbol{X}_0) \oplus \int_{t_0}^{t} \boldsymbol{\Phi}(t, \zeta)\boldsymbol{B}(\zeta)\epsilon(\boldsymbol{u}(\zeta), \boldsymbol{U})d\zeta$$

where $\boldsymbol{\Phi}(t, t_0) = e^{\boldsymbol{A}(t-t_0)}$ and $\boldsymbol{A}$ and $\boldsymbol{B}$ are the state and input matrices related to the AV. Thus, with this approach, reachable sets can be calculated to capture the uncertain motion of the AV tracking a given trajectory. In Fig. 4, the $[\dot{x}, \dot{y}]$ velocity reachable tube for an autonomous quadrotor aerial vehicle following a straight line trajectory for 2s ♠[2], considering disturbances and measurement and input noise, is shown. The blue dotted curve shows the path of the quadrotor whereas the red star curve shows the desired trajectory. The desired trajectory is different from the actual one due to the presence of wind disturbance, however the actual trajectory is contained inside the reachable tubes, since system and sensor uncertainties as well as disturbances are considered when calculating such reachable tubes.

[2]NB: I'll create one for a car

The proposed stochastic reachability methodology is based on formulating the reachability problem as a stochastic optimal control problem. Based on the expression of the probability that the state of the controlled system will evolve within the safe region as a multiplicative cost, dynamic programming (DP) can be used to compute the Markov policy maximizing the cost, and also the maximally safe sets corresponding to different safety levels.

2.5 *Runtime Verification for Safety-Aware Autonomous Vehicle Operations:*

Having determined the safe reachable set of states, we need to ensure the planned trajectory of the autonomous vehicle is also safe. However, offline verification of autonomous vehicles operations do not scale well as exploring all possible contingencies can be computationally prohibitive. The main challenge here lies on how to perform runtime monitoring online and efficiently maintaining computation time bounded while guaranteeing system's safety. To this end, we propose to develop techniques for prediction of autonomous vehicles safety-critical properties (e.g., probability of colliding with a vehicle) that can be autonomically verified at runtime. We will develop techniques for online verification of missions with AVs. While such problems are computationally very hard, we will leverage prior work by PI Bezzo, that relies on satisfiability modulo theories (SMT) solvers to address the system of constraints that mix Boolean variables with real variables [**?**, **?**]. In this approach, our aim is to generate guarantees based on a probabilistic temporal logic framework so that we can achieve verified autonomy in unknown environments with learning in the control loop. Once completed, this will be the first blend of probabilistic formal methods that can reason about probabilistic abstractions of machine learning models in an autonomous system operation. To enable online efficient prediction, hence minimizing computation time and allow real-time monitoring applications, we propose to cast the prediction problem as a two-point boundary value problem (2PBVP) [16] and use support vector machine learning algorithms to approximate the reachability boundary using a nonlinear classifier, separating training queries into reachable and non-reachable sets. The idea is to generate training data by solving a large number of 2PBVPs as presented in [16] for various randomly generated query pairs. The classifier is then used online to estimate if new query points are reachable. The advantage of such approach is that it can be applied to any system with minimal computation time at runtime since training can be executed offline a priori.

2.6 Behaviour-guided safe trajectory synthesis

The safety assessment introduced above is used to guide the AV actions to maintain safety, i.e., to maintain the AV within safe reachable sets. The next step is to select appropriate steering and control actions for the autonomous system and maintain safety at all the time. The innovation here is that we consider a behavior-guided approach in which the preferences from the user profiles in Section 2.2 are taken into consideration while determining the proper actions of the AV without violating safety constraints. We propose a reinforcement learning based method for determining the control actions of the autonomous vehicle while taking the user profiles into consideration.

2.6.1 Safe Reinforcement Learning-based Adaptation:

We propose a reinforcement learning-based approach to compute the desired control policy. Reinforcement Learning (RL) is an area of machine learning concerned with learning how an agent should behave in a given environment in order to maximize some form of cumulative reward. To this end, an agent cycles through a series of transitions which consist of going from one state to the next state by applying actions to the environment and receiving rewards as a consequence of his actions. The goal of RL is to derive a policy, which, given a state, provides the action to take in order to maximize the cumulative reward. A central problem of RL is thus that of properly assigning the reward to the actions that lead to such reward (a notion known as credit-assignment [17, 18]). In our case the reward is a combination between user preference and safety, where safety is determined using the reachability set analysis as explained in Section 2.4. Reinforcement learning is particularly well-suited to problems with a long-term versus short-term reward trade-off. A key challenge is about enabling online learning while maintain safety [19, 20, **?**] If the uncertainties of the system's model are high, for example in the event of a failure like a motor outage, then we can think to use a machine learning approach to determine the optimal policy and control inputs to provide to the system to maximize a given reward (e.g. go-to-goal) while learning the new model.

In our case the AV is assumed to act according to an optimal policy for a Markov decision process M^{S}. The system knows its current state and action sets as well as the initial transition probabilities for M^{S}. Due to unforeseen disturbances and unpredicted behaviors (e.g., environmental disturbances, high traffic) the transition probability model changes, but the reward function remains the same. The model is not completely known and need to be refined as the system is running. This problem is cast as an MDP problem in which transition probabilities are changed and adapted at every iteration to handle unknown environments and systems dynamics. Current and historic measurements and inputs will be used to refine the current model and update transition probabilities. We can think to leverage reachability analysis also in this task to incorporate uncertainties as the RL-based approach converges to a model and policy and to predict accordingly the possible future states that the vehicle may cover.

The outcomes of the proposed research in Thrust 2 are the calculations of a safety region for the autonomous vehicle in real time, and a way to synthesize safe control decisions and trajectories which take the user emotion and behavior preferences in Thrust 1 (Section 2.2) into account. Both of these are then used to provide feedback and situational awareness to the passenger. This is described next in Thrust 3.

2.7 Research Thrust 3: Behavior-guided feedback design

The purpose of this research thrust is to develop and validate what feedback should the autonomous vehicle provide to the passengers. If we detect the passengers are in distress, or confused about the car's driving behavior and actions, then providing appropriate feedback and meet their emotional and behavioral needs. While cars have become significantly more usable — particularly with regard to reliability and safety over the past twenty years — thanks to the introduction of new technologies such as electronic fuel injection, ABS, airbags, stability control etc., many of these technologies have succeeded out-of-sight of the humans behind the wheel. Yet when it comes to newer technologies - like advanced driver assistance systems (ADAS) , we see a much less successful integration of technology, vehicle, and user. Furthermore, as soon as driving "feels" even partly autonomous, people switch off, they become disengaged from the process of driving — and fail to monitor the system, which can lead to disastrous consequences for semi-autonomous cars. We hypothesize, that for autonomous vehicles trust comes from two factors: *predictability*, and *explainability*. If a user expects a car to drive in a certain way in a certain situation, and the car conforms to his expectation, the user will tend to trust it more. Occasionally, when the AV's action surprise/confuse the user- as long as there is an explanation provided for it, the user can again gauge her level of trust in the system. Given a proile of the passenger's behavior and emotional needs, from Section 2.2 the goal of this research thrust is to:

1. Develop an automated way to provide explanations for the autonomous vehicle's actions to the passenger - this caters to the *explainability* aspect of a passenger's trust, and
2. Develop user interfaces which convey the intended actions of the vehicle to the passenger so they can gauge the *predictability* component of their trust in the autonomous vehicle.

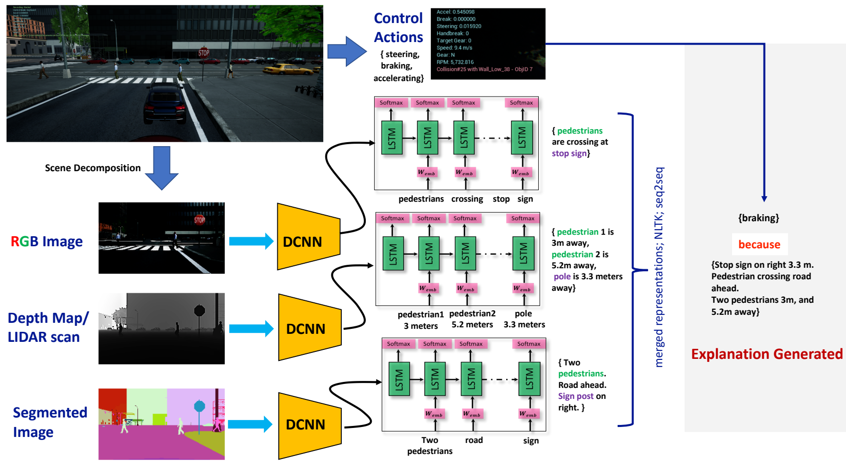

Figure 5: Deep-Explanation generation: Each dimension of the scene decomposition is used as an input to caption generation. Representation matching, and seq2seq are then used to generate a likely explanation for the predominant action stream.

2.7.1 Explainability via Deep-Explanations

Automatic image description generation is a challenging problem that has recently received a large amount of interest from the computer vision and natural language processing communities [21, 22, 23, 24, 25] The task of automatic image description involves taking an image, analyzing its visual content, and generating a textual description (typically a sentence) that verbalizes the most salient aspects of the image. Despite the difficult nature of this task, there has been a recent surge of research interest in attacking the image caption generation problem. In particular, deep neural networks have been shown to form new grammatically correct sentences as opposed to the template based models and their limited generalization capability to a novel image. Aided by advances in training deep neural networks and the availability of large classification datasets, recent work has significantly improved the quality of caption generation using a combination of convolutional neural networks (convnets) to obtain vector representation of images and recurrent neural networks to decode those representations into natural language sentences.

In the proposed research we extend current image caption generators to work with multi-modal data-sets.

1. Instead of generating captions from RGB images alone, we will also generate captions from LIDAR data, depth sensor images, and segmented images.
2. The captions are then combined with information about the control decision (steering, acceleration, and braking) made, to create an explanation (description) of the scene for the user.
3. We test this approach in simulation, using the photo-realistic Airsim [26] simulation.

Consider the scene shown in the Figure 5; in this situation multiple modalities of the scene are available - namely, an RGB image from a center mounted camera, a depth map (or it could be a point cloud) using a depth sensor or a LIDAR, a segmented image (usually obtained by running a deep convolutional neural network on the RGB image), and the action state of the vehicle - steering, acceleration, braking, etc. We first train a single joint model that takes an image I as input, and is trained to maximize the likelihood $p(S|I)$ of producing a target sequence of words $S = S1, S2, ...$ where each word S_t comes from a given dictionary, that describes the image adequately. The inspiration for our work comes from recent advances in machine translation, where the task is to transform a sentence S written in a source language, into its translation T in the target language, by maximizing $p(T|S)$. For many years, machine translation was also achieved by a series of separate tasks (translating words individually, aligning words, reordering, etc), but recent work has shown that translation can be done in a much simpler way using Recurrent Neural Networks (RNNs) [27, 28, 29] and still reach state-of-the-art performance. An "encoder" RNN *reads* the source sentence and transforms it into a rich fixed-length vector representation, which in turn in used as the initial hidden state of a "decoder" RNN that *generates* the target sentence. By replacing the encoder RNN with a CNN has been shown to work well when the inputs are images [30]. For the RGB image captioning we will use Neural Image Captioning method, as described in [25].

2.7.2 Captioning from point clouds, depth maps, and segmented images

To properly process the world, an autonomous car needs to take raw sensor information (like point cloud) and figure out what it's seeing. Arguably, two of the most important pieces of information are depth:*"how long until I hit this object?"* and category: *"what kind of object is this?"*. CNNs have produced incredible results on RGB images, and in this research we will show that they're very applicable to LIDAR depth data. Using annotated depth map data obtained from Airsim autonomous vehicles simulator, we will develop networks which take an input a depth image or a point cloud - where each pixel's intensity/color represents the distance of the object from the sensor. The network's objective is to directly maximize the probability of the correct description given the image by using the following formulation:

$$\theta^* = \arg\max_\theta \sum_{I,S} \log p(S|I;\theta) \tag{1}$$

Where θ are the parameters of the model, I is the depth image, and S is the correct transcription. Since S represents any sentence, its length is unbounded. Thus, it is common to apply the chain rule to model the joint probability over $S_0, \cdots, S_N$, where N is the length of this particular example as:

$$\log p(S|I) = \sum_{t=0}^{N} \log p(S_t|I, S_0, \cdots, S_{t-1}) \tag{2}$$

where we dropped the dependency on θ for brevity. During training, (S, I) is a training example pair, and we optimize the sum of the log probabilities as described in (2) over the whole training set using stochastic gradient descent. The long short-term memory (LSTM) model for sentence generation is trained to predict each word of the sentence after it has seen the image as well as all preceding words as defined by $p(S_t|I, S_0, ..., S_{t-1})$. Both the image and the words are mapped to the same space, the image by using a vision CNN, the words by using word embedding W_e. The image I is only input once, at $t = -1$, to inform the LSTM about the image contents. This process is also shown in Figure 5. Finally, the output of the CNN classification, is used to report the average distance of the classified object, as part of the sentence. For instance, when the CNN detects a pedestrian, it uses a bounding box around the detected pedestrian, to compute the distance of the centroid of the depth pixels, or the average of the depth pixels inside the bounding box.

2.7.3 Explanation generation via captioning merging

Using the CNN and LSTM networks on each of the input modalities, we obtain a caption for each. The thing to note is that each of the networks uses the same dictionary, and therefore, each network is likely to result in a caption which refers to the same object. Going back to the example in Figure 5, each of the captions refers to a pedestrian. We use this commonality to merge the captions into a single explanation. We will also explore architectures where the output of one caption generation can influence the input of the other networks. To merge the captions together we rely on an idea similar to sequence-to-sequence matching [29]. Sequence-to-sequence (seq2seq) models have enjoyed great success in a variety of tasks such as machine translation, speech recognition, and text summarization. Specifically, an NMT system first reads the source sentence using an encoder to build a vector, a sequence of numbers that represents the sentence meaning; a decoder, then, processes the sentence vector to emit a translation. This is often referred to as the encoder-decoder architecture. In this manner, NMT addresses the local translation problem in the traditional phrase-based approach: it can capture long-range dependencies in languages, e.g., gender agreements; syntax structures; etc., and produce much more fluent translations. For sentence merging, we use the encoder part of the NMT to translate each caption into the same vector space. We then search for common word embedding (say corresponding to a pedestrian, or a road, or a sign). The problem becomes that of searching for common vectors. We simply, merge or combine the sequences corresponding to the common vectors and then use the combines sequence as the input to the decoder part of the NMT (which is simply the inverse of the encoder as we are not performing any translation). To summarize, we take all the captions from each modality, map it to a common vector space, look for common words in the vector space, which are present in the dictionary, and decode back into a sentence. The resulting sentences are human readable. By combining them with the actions being taken by the car, they can serve as meaningful explanations, providing insights into what the autonomous vehicle "sees" and what action it takes. PI Behl has previously conducted research on interpretable machine learning and control synthesis algorithms for building automation [31, 32]

2.7.4 Predictability via user interface design

It is not enough to generate an explanation, we must also optimize the manner in which the explanations are presented to the human driver. This research will develop a user interface that is able to communicate intentions, and movements of autonomous vehicle to the user. The ultimate goal is to increase the confidence felt by the user to this system, encourage inclusion, acceptance and all the benefits it could bring to user mobility, and emotions. We envision a scenario in which car is not only a tool to get from point A to point B, but a complex device that receives and provides information, it can recognize emotional states (Section 2.2), hypothetical distractions and driver physical problems (sleep or tiredness), and provide feedback to the driver. Present day dashboards/instrument clusters/user interface for semi autonomous vehicles present a high level overview or a wireframe view of what the car sees as its driving. Its typical to highlight lane markings and road signs, and even highlight potential hazards. However, what is lacking is a design to transparently communicate the intended actions of the vehicle. Such transparency is important, because the algorithms by which autonomous cars make decisions are largely invisible to the driver. If something unexpected occurs, the driver can only speculate what happened. In our autonomous vehicle simulator (Section **??**) we will test graphical interfaces design which conveys the intention of the

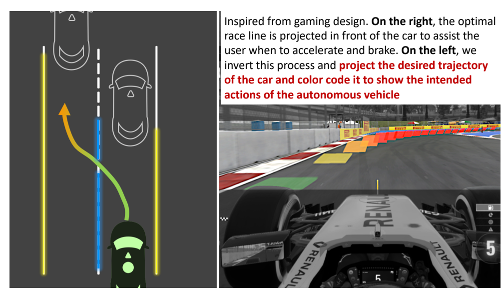

Figure 6: Inspired from gaming, projecting the intended trajectory of the car in a color coded manner can realy critical feedback to the user about what actions the car intends to take in any given situation.

autonomous car to the passenger. We are not proposing designing production ready user interfaces for autonomous cars, but rather a functional study for testing, determining visual variables complexity, information hierarchy and topology, and cognitive load for certain functional UI elements. The UI elements will show immediate future movements and choices. "*Why is the car turning right?*", "*Which way is it taking?*", "*Can the car detect the cyclist in front of me?*". To present a example, we take inspiration from design for motorsport gaming. As shown in Figure 6(right), many games project a path in front of the race car to guide the user which is the fastest race line on the track. The color of the race track tells the gamer where to accelerate (green) or slow down (amber). Our idea is to invert this concept for an autonomous car where the UI (Figure 6(left)) shows the intended trajectory of the autonomous vehicle and the color of the trajectory coveys if the car will speed up of slow down (brake). We hypothesize that such feedback can help passengers gauge their "predictability" metric and hence their confidence, and trust in the vehicles.

We envision that these deep explanations can offer insight into the neural networks responsible for the perception and scene understanding for self-driving cars. Such black box networks soon may also play a role in planning and control for the autonomous cars [33], making it even more important to obtain explanations for their actions/predictions. By providing readable explanations of the actions along with and correctly designed user interfaces, we provide context and feedback to the passenger, enabling an increase in the trust between the passenger and the vehicle. The trajectories and safety regions computed in Thrust 2, will be incorporated as a part of the feedback design. Likewise, the explanations will be generated for instances when the emotion detection from Thrust 1 predicts the driver is confused due the the behavior of the car.

3 Experimentation and Evaluation Plan

For each of the proposed research thrusts in Sections [2.2],[2.3], and [2.7], we will conduct extensive experimentation using high fidelity automobile simulators like PreScan and AirSim, a full scale diving simulator, smaller scale autonomous vehicle test-beds, and a full scale car fitted with sensors for monitoring human driving behavior and emotions. We next describe in detail, the different automotive cyber-physical systems tesbeds, and a cross cutting experimentation plan.

3.1 Scenario-based behavior and trust modeling

Behind the wheels of a self-driving car, everyone of a sudden becomes a backseat driver. For thrust 1 on monitoring driver behavior and emotions (Section 2.2, we want to observe and record driver activity in a controlled environment. For thrust 3 on feedback design (Section 2.7), we want to measure how passengers react to varying degrees of feedback and explanations provided to them by the UI. We setup different traffic scenarios in our full scale driving simulator, to run experiments to gather the data required to build models for human behavior, and test the effectiveness of different kinds of feedback. We can simulate multiple scenarios – a four way stop sign, safely passing bicyclists on a narrow road, overtaking a large semi on a freeway, driving too close to a barrier in the absence of a shoulder on the left lane, etc.

3.1.1 Participants

We are proposing to conduct user studies and driving experiments for Thrusts 1, and 3. The design of a subject study protocol is being undertaken at the time of the witting of the proposal. We are closely working with the Institutional Review Board for Social and Behavioral Sciences at the University of Virginia. The data management plan describes in detail how user data will be recorded, used, distributed, and maintained.

Figure 7: We recreate real world traffic situations in simulation using PreScan. During experimentation, we invite human drivers to first manually drive through each scenario (mode 1), followed by an autonomous driving mode where limited feedback and car intentions are provided to the driver, and finally in mode 3, the users experience fully autonomous driving with full feedback and situational awareness provided. We then survey the participants about what feedback enables their trust, which UI elements do they prefer, which mode made them emotionally satisfied.

A total of at least 100 participants from the University of Virginia will be recruited for participation in this study over 3 years.

3.1.2 Setup

Figure 7 shows the full driving scale setup which is partially built by PI Feng and will complete during the proposed effort. The hardware platform is based on the Force Dynamics 401CR driving simulator. This four-axis motion platform can pitch, roll, yaw, and heave, to simulate the experience of being in a vehicle. We expect to collect data about realistic human response during the driving. The driving simulation software is called PreScan, which is a software tool used by the industry for Advanced Driver Assistance Systems (ADAS) development. Sensors will be used for both high-level inference of human's intent and preferences and low-level monitoring of human behavioral, mental and physiological states. These sensors include EEG for neural signals, EKG for heart activity, EMG for muscle activity, a camera for head tracking, eye tracking suite and cloud-based speech recognizer.

3.1.3 Design and procedure

As shown in Fig. 7, we have re-created 4 real world driving situations in Pre-Scan - passing bicyclists on a narrow back road, a 4 way intersection, overtaking a semi truck at freeway speeds, and driving on a freeway in the left lane in the absence of a left shoulder. For each of these scenarios, we will conduct the following user experiment:

1. **Mode 1 - Manual driving**: In this mode, we ask the subject to manually drive through a given scenario. While they drive through the scenario, we monitor the person's behavior and emotional cues as well as log their driving behavior. This driving behavior (steering, brake, acceleration) can be considered as their baseline profile. As an example, we will ask the subject to drive through the four way stop sign scenario, where they have to self-asses their right of way as they approach a stop

sign. We can then intentionally program one of the agents in the simulation to drive3 out of the order for the right of way and see how the driver adapts.

2. **Mode 2 - Autonomous driving with limited feedback:** Next, we ask the same subject to this time, sit back and relax, as the car drives by itself in the simulator. Everything in the scenario is identical to last time, except that the driving is fully autonomous. We present the driver with a visualization of a virtual dashboard which depicts what the car sees, very similar to what most of the dashboards for semi-autonomous vehicles depict today. Once again, we monitor the drivers behavioral and emotional cues to interpret their trust in the system. We also give them the option to press a button on the steering wheel, every time they think the car did something that they did not anticipate, or if they mistrusted the car's autonomous driving actions. "flagging" these events in a scenario provides us with a basis for designing UI feedback.
3. **Mode 3 - Autonomous driving with full feedback:** Lastly, we have the autonomous car drive through the same scenario one more time, but this time we provide full situational awareness of the scene to the user along with cues about the car's intended actions. In the stop sign example, we project the car's understanding of its right of way, and how that dynamically updates as other vehicles drive through the intersection. We project the cars intended trajectory so the user know that the car will make a left turn and will yield to incoming traffic. More generally, for the parts of the simulation which were flagged by the user, we run the Deep Explanations engine to provide natural language explanations for the car's actions.

Following the experiment, we survey the subjects to understand and gather data on which explanations help increase the trust of the subject, which UI elements work, and to what extent. Each participant will also be given a questionnaire, adapted from [34], to learn about their levels of trust in the autonomous car. We will use scales which empirically define the feeling of trust in the autonomous system from [35] which has been used in many studies about trust [36].

3.1.4 Expected results

Upon the completion of data collection, the data will be subjected to a factor analysis to reveal the underlying factor structure of the experimental scale. The simulated scenario based experimentation will enable creating a library of emotional profiles, and models, of drivers; identify the mapping between the emotional state and the control action of the autonomous car, and vice-versa, collect valuable data for training the Deep Explanations networks for natural language feedback; and by deploying the same subject in a real car and monitoring their preferences and driving behavior, we will also validate the realism of the simulation study. Although, the simulation setup has a distinct advantage, i.e. repeatability of the same scenario, one of the goals of the proposed work is to develop the proposed models for emotion and driving behavior using data from real deployment in subject-owned vehicles (as described in Section 2.2).

3.2 Test-beds for reachability and control synthesis experiments

The proposed research, in particular Thrust 2 about safety assessment via reachability analysis and and reconfiguration via safe reinforcement learning will be validated using our testbed of autonomous ground vehicle available in PI Bezzo's robotics laboratory. Figure 8 shows one of the autonomous ground vehicle equipped with the same sensors available in real autonomous cars which include velodyne lidar, stero cameras, imus, wheel encoders, differential GPS, and onboard i7 cpu. The vehicle's equations of motion and dynamics are similar to real autonomous vehicles with reduced speed. This platform and especially the sensor testbed will be used offline to create a library of primitives for the reachability analaysis and then

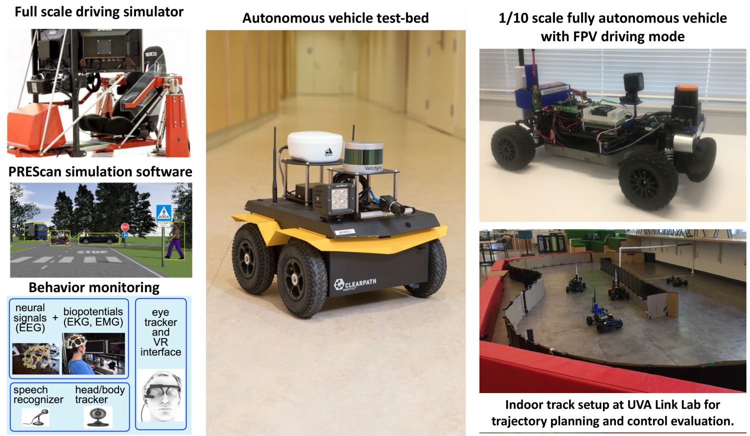

Figure 8: Automotive Cyber-physical System (CPS) test-beds which will be used for conducting the experimentation to support the proposed research. In addition, we will also utilize a real car to gather driving profile data as describe in Thrust 1.

used at runtime in different scenarios to predict and validate the safety assessment presented in Thrust 2. We also plan to use virtual reality goggles (also available in the PI's lab) to simulate the presence of a human on board the vehicle. PI Behl has developed a 1/10 scale autonomous racing car platform called F1/10 [37] (Figure 8. The cars are a realistic representaiton of a full scale car, with similar dynamics, just different parameters. It carries a suite of sensors similar to that carrier by current full scale prototypes of autonomous vehicles, including a LIDAR, a stereo camera, depth camera, an IMU, and a high performance GPU board running the Robot Operating System (ROS). PI Behl has built a fleet of F1/10 cars and has access to an indoor track facility for testing the control synthesis algorithms in a safe environment but at high and realistic traffic speeds. The PI also has access to a 1/10 scale car which has been modified in that it can be driven in first person view (FPV) using steering wheels and pedals, while wearing a headset which relays the onboard camera view to the human driver, thereby allowing us to conduct both UI and driving style behavior experiments at 1/10 the scale.

4 Broader Impacts

As the development of AVs progresses at faster rates than ever utilizing a diverse set of approaches, it is essential that frameworks for human centered design be developed in close coordination with academic and industry partners. We our working closely with automotive manufacturers like Toyota, as well as Perrone Robotics - an autonomous driving company based in Virginia. Letters of collaboration form both are included in the grant application. The research outcomes will have the following broader research impacts: (a) be a valuable contribution towards increasing the overall safety of fully autonomous vehicles, which are likely to become ubiquitous in the near future, (b) minimize the risk of economic loss, damages, and injuries due to failures and uncertainties, (c) the underlying frameworks of generating local explanations from sensor data, and safe operation through reachability analysis can help enhance a large scope of autonomy including but not limited to autonomous vehicles, robotics, aircraft autopilots, and automatic surgery equipment, (d) provide valuable and scientific insights to automotive manufactures and stakeholders about user interface design for sled-driving cars, and user expectations about fully autonomous cars, and (e) Leveraging human behavior, emotions, and trust to help enhance the capabilities of autonomous vehicles and also facilitate the deployment of autonomous vehicles in the real world.

4.1 Improving Education on Autonomy and Cyber-Physical Systems:

There is a significant gap in the way we conduct interdisciplinary Cyber-Physical Systems (CPS) research, and the way we train students about CPS. Students coming out of higher education are expected to solve 21st-century CPS problems and enter into occupations that haven't even been imagined yet. The PIs teaching mirrors the inter-disciplinary approach towards research. The PIs will develop new courses to ensure that students cultivate a holistic view of life-critical, and safety-critical system development by drawing stronger connections between systems theory, formal methods, machine learning, human factors, and hands-on development. New courses for the graduate and undergraduate teaching will be devleoped:

1. **Principles of Modeling in Cyber-Physical Systems** : This course will provide a solid foundation for understanding different modeling paradigms, and explore them through a deep dive and hands on implementation for three CPS domains: Energy, Medical, and Automotive cyber-physical systems. Students will come out of this course with advanced and transferable knowledge of model-based design methods and tools, and will be ready for tackling multi-disciplinary systems projects.
2. **F1/10 Autonomous Racing** - Principles of Perception, Planning, and Control. - PI Behl teaches a course where teams of students build, drive, and race a fleet of 1/10 scale fully autonomous vehicles, while learning about principles of perception, planning, and control. The F1/10 platform facilitates a wide range of research, education, and training in autonomy. The course material developed by PI Behl for F1/10 is open-source, and publicly available on f1tenth.org. It has been used by a dozen university around the world to build their own versions of the 1/10 autonomous cars. Course materials available online have been adapted for teaching CPS and Autonomous Systems courses at UT Austin, and Clemson University.
3. The PIs Bezzo and Heydarian, will also co-teach a new course on **Robots & Humans** which explicitly focuses on trust and behavior modeling of drivers.
4. PI Bezzo teaches also an undergraduate course on **CPS Simulations**. The proposed research and results will be included in the curriculum activities for this course to include safety assessments during simulations for autonomous vehicles conducted on PI Feng's full scale simulator.

All the PIs reside within the Cyber-Physical Systems Link Lab at UVA and will jointly work together towards creating a template for CPS education centered around behavior guided autonomy, and autonomous vehicles.

4.2 K-12 Impacts:

The PIs have previously helped to organize several outreach programs and workshops (Iridescent, C-TECH2 Workshop, NASA INSPIRE Workshop) promoting STEM studies to underrepresented and minority students in K-12. We plan to run robotics summer camps for local middle and high schools to engage them to the problem of safety in autonomous vehicles. PI Heydarian collaborates with the "Girls Can Change the World with Science" outreach program to inspire young female students across Virginia elementary and middle schools to consider futures in science and engineering fields. We use the findings of this research to demonstrate how some of our engineering-problems can only be solved by integrating engineering fields (e.g., computer science, system engineering) and other domain(s) together (i.e., psychology).

4.3 Graduate/Undergraduate Students and Outreach Effort:

The PIs are committed to recruiting and nurturing minority and local high-school students by actively participating in local programs such as Women in Computer Science (WICS), Open-house visitation days, and 1-on-1 career-related advising. Every year, for the past 3 years, PI Behl has been organizing the International F1/10 Autonomous Racing Competition, where teams from all over the build 1/10 scale cars using open source instructions on how to build, drive, and race these vehicles in a battle of algorithms. The competitions are held in conjunction with a premier venue such as CPS Week (2018), SenSys (2017), and ES-Week (2016, and 2018 Fall). In additional to the organizing the competition, the PI has regularly held F1/10 tutorials to teach undergraduates, and graduate students about autonomy. PI Feng organized an N^2Women Luncheon at the CPS Week 2015, and is co-organizing a Mentoring Workshop at FLoC 2018. The 4 PIs collectively advise 7 PhD students and have mentored over two dozen undergraduate students. We will also engage in outreach activities aimed to recruit graduate students from neighboring Historically Black Colleges and Universities, particularly Virginia State University.

4.4 Dissemination Impacts:

The scientific community will be made aware of the research project and opportunities for collaboration through the project and Link-Lab webpages, seminars and conferences such as CPS Week, IROS, RAS, ICML, HCI, and SenSys PI Behl will maintain a website dedicated to the research focus and publish all results, presentations, and videos there, consistent with the Data Management Plan. In addition, the PIs are regularly invited by Toyota, US DoT, Virginia Department of Transportation (VDoT), to give invited talks to the AV community. This research will lead to the public release of a number of models, data-sets, and tools:

1. Model and data-set library for driving behaviors and emotions - both from simulated and real world experimentation.
2. Annotated training data for learning natural language explanations from multimodal autonomous vehicle sensor data.
3. A UI toolkit with Unity elements for designing feedback for autonomous vehicles - with user study statistics assigned to each element in the kit.
4. Github repository for code and tutorials related to deep explanations neural networks, and control synthesis algorithms.

We also plan to disseminate our results in industrial venues such as industry-focused conferences and workshops (e.g., TED, Siggraph, xxxxxxxx). The PIs are fully committed to these education and outreach activities and beleieve these will be a catalyst to attract more students to pursue a career in STEM.

5 Project Management and Collaboration Plan

This is an ambitious proposal which addresses challenging problems in autonomous vehicles, but we have an excellent team with broad and complementary skills, the facilities to support it, and a strong basis to start from. The proposed work will be performed by the PIs Behl, Heydarian, Bezzo, and Feng, and 4 graduate students.

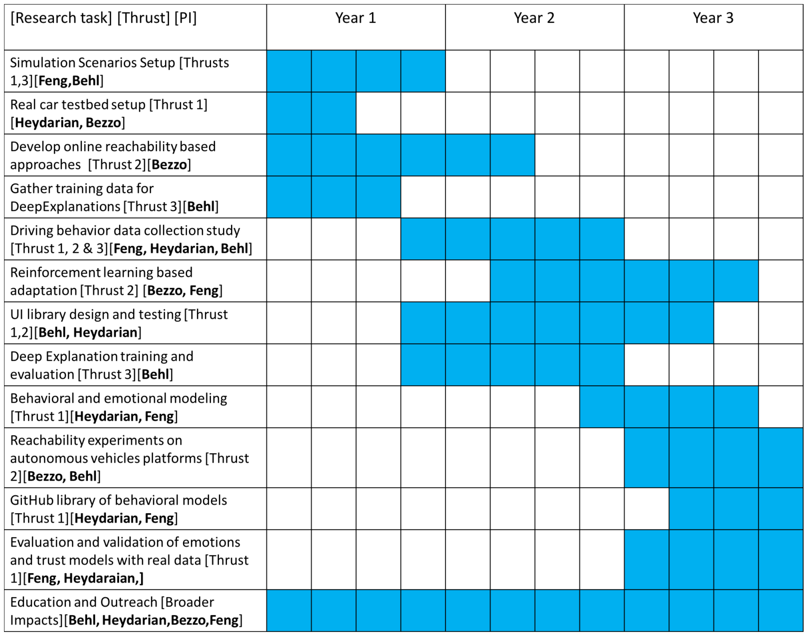

Figure 9: Description of project tasks, roles, and timeline for the proposed research

5.1 Roles and responsibilities:

All 4 PIs are members of the Cyber-Physical Systems Link Lab research center at the University of Virginia. PI Behl is an expert in machine learning modeling and data-driven control for CPS. His F1/10 autonomous racing car platform has been built by over 20 universities, and has resulted in 2 International autonomous vehicles competitions at CPS week and ES-Week. He serves on the editorial board of the SAE International Journal of Connected and Automated Vehicles. PI Behl's research won the best energy paper award at the American Control Conference (ACC) 2017. He is also the recipient of the 2011 Richard K. Dentel Memorial Prize awarded by the University of Pennsylvania for research in urban transportation.

Co-PI Heydarian's research focuses on integrating user-centered considerations (i.e., human behavior and emotional models) into the design and operation of automated systems (Human-in-the-Loop CPS). He has experience conducting a number of experimental studies (over 700 consenting participants) in the past and will lead the research effort in Thrust 1.

Co-PI Bezzo's expertise falls under the control and planning of autonomous mobile robots. His research focuses on the design of resilient controller and adaptive planning techniques to improve safety, security, and performance of autonomous aerial and ground vehicles in uncertain situations and environments. He has received several awards including the 2016 Robotics and Automation Magazine Best Paper Award, the Best Paper Award at the 2014 ICCPS.

Co-PI Feng's expertise is on temporal logic planning for autonomous robots. She has received several awards including James S. McDonnell Foundation Postdoctoral Fellowship, Rising Stars in EECS, UK Engineering and Physical Sciences Research Council Scholarship, and Cambridge Trust Scholarship.

The timeline for the proposed work is shown in the Figure 9.

5.2 Risks and mitigation plans:

Our excellent team and access to various automotive cyber-physical systems test-beds gives us the confidence that the proposed research can be carrier out without any major hassles. While we expect some difficulties may arise along the way, these would not sabotage the effort. Examples include component failures (these can be replaced), slow runtimes of software (this will be optimized), and realism of the simulated scenarios (these will be validated by making the same subject drive both in simulation and real world and compare the behavior models). The proposed work involves the use of human subjects for social and behavioral studies and that may be perceived as a risk but the team is experienced with working with IRB.

5.3 Results from Prior NSF Support

Madhur Behl: (NSF-1735587); $\$2,499,238$; 09/1/2017-08/31/2021 Title: CRISP Type 2: dMIST: Data-driven Management for Interdependent Stormwater and Transportation Systems. Role: Co-PI. **Intellectual Merit:** To create a novel decision support system denoted dMIST (Data-driven Management for Interdependent Stormwater and Transportation Systems) to improve management of interdependent transportation and stormwater infrastructure systems. **Broader impact:** The research is intended to have broad impact related to national economic and security interests due to its focus on sea level rise. Paper submitted to 11^{th} International Conference on Urban Drainiage Modeling.

Lu Feng: Co-PI of NSF grant CNS-1739333 "CPS: Medium: Safety-Critical Wireless Mobile Systems", \$800,000, 9/1/2017-8/31/2020. **Intellectual Merits**: This project aims to develop a framework for joint modeling and analysis of motion and communication in order to find provably safe coordination paths. **Broader Impacts**: The research will allow mobile systems to realize the performance benefits of wireless coordination while preserving the ability to provide provable safety guarantees. PI of NSF grant CNS-1755784 "CRII: CPS: Cognitive Trust in Human-Autonomous Vehicle Interactions", \$175,000, 4/1/2018-3/31/2020. *Intellectual Merits*: The research is to develop new formal specification and verification methods for formally expressing and reasoning about trust in human-autonomous vehicle interactions. *Broader Impacts*: The research will create new techniques for assisting in the design of safe and trustworthy autonomy into future vehicles.

References

[1] Daisuke Wakabayashi. Self-driving uber car kills pedestrian in arizona, where robots roam. *New York Times*.

[2] Tesla car that crashed and killed driver was running on autopilot, firm says. *The Guardian*, Mar 2018.

[3] Autonomous cars at ces 2018. *Techcrunch*, Jan 2018.

[4] Patrick Lin. Why ethics matters for autonomous cars. In *Autonomous Driving*, pages 69–85. Springer, 2016.

[5] Deloitte. What's ahead for fully autonomous driving. In *Deloitte's Global Automotive Consumer Study*. 2017.

[6] Daniele Bernardini and Alberto Bemporad. Stabilizing model predictive control of stochastic constrained linear systems. *IEEE Transactions on Automatic Control*, 57(6):1468–1480, 2012.

[7] Mahmoud Elnaggar, Jason D. Hiser, Tony Lin, Anh Nguyen-Tuong, Michele Co, Jack W. Davidson, and Nicola Bezzo. Online control adaptation for safe and secure autonomous vehicle operations. In *2017 NASA/ESA Conference of Adaptive Hardware and Systems (AHS)*, Pasadena, CA, July 24-27 2017. IEEE.

[8] Nicola Bezzo, James Weimer, Yanwei Du, Sang H. Son, Oleg Sokolsky, and Insup Lee. A stochastic approach for attack resilient uav motion planning. In *American Control Conference (ACC 2016)*, pages 1366–1372. IEEE, 2016.

[9] J. Ding, J. H. Gillula, H. Huang, M. P. Vitus, W. Zhang, and C. J. Tomlin. Hybrid systems in robotics. *IEEE Robotics Automation Magazine*, 18(3):33–43, Sept 2011.

[10] J. Ding, E. Li, H. Huang, and C. J. Tomlin. Reachability-based synthesis of feedback policies for motion planning under bounded disturbances. In *IEEE International Conference on Robotics and Automation*, pages 2160–2165, May 2011.

[11] Jeremy H Gillula, Gabriel M Hoffmann, Haomiao Huang, Michael P Vitus, and Claire J Tomlin. Applications of hybrid reachability analysis to robotic aerial vehicles. *The International Journal of Robotics Research*, 30(3):335–354, 2011.

[12] Jeremy H Gillula, Haomiao Huang, Michael P Vitus, and Claire J Tomlin. Design of guaranteed safe maneuvers using reachable sets: Autonomous quadrotor aerobatics in theory and practice. In *IEEE International Conference on Robotics and Automation (ICRA), 2010*, pages 1649–1654. IEEE, 2010.

[13] Matthias Althoff. Reachability analysis and its application to the safety assessment of autonomous cars. *Technische Universitat Munchen*, 2010.

[14] Esen Yel, Tony Lin, and Nicola Bezzo. Reachability-based self-triggered scheduling and replanning of uav operations. In *2017 NASA/ESA Conference of Adaptive Hardware and Systems (AHS)*, Pasadena, CA, July 24-27 2017. IEEE.

[15] A. A. Kurzhanskiy and P. Varaiya. Ellipsoidal toolbox (et). In *Proceedings of the 45th IEEE Conference on Decision and Control*, pages 1498–1503, Dec 2006.

[16] Ross E Allen, Ashley A Clark, Joseph A Starek, and Marco Pavone. A machine learning approach for real-time reachability analysis. In *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*, pages 2202–2208. IEEE, 2014.

[17] Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore. Reinforcement learning: A survey. *Journal of artificial intelligence research*, 4:237–285, 1996.

[18] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*, volume 1. MIT press Cambridge, 1998.

[19] Michael Kearns and Satinder Singh. Near-optimal reinforcement learning in polynomial time. *Machine Learning*, 49(2-3):209–232, 2002.

[20] Pieter Abbeel and Andrew Y Ng. Exploration and apprenticeship learning in reinforcement learning. In *Proceedings of the 22nd international conference on Machine learning*, pages 1–8. ACM, 2005.

[21] Justin Johnson, Andrej Karpathy, and Li Fei-Fei. Densecap: Fully convolutional localization networks for dense captioning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4565–4574, 2016.

[22] Kelvin Xu, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron Courville, Ruslan Salakhudinov, Rich Zemel, and Yoshua Bengio. Show, attend and tell: Neural image caption generation with visual attention. In *International Conference on Machine Learning*, pages 2048–2057, 2015.

[23] Cheng Wang, Haojin Yang, Christian Bartz, and Christoph Meinel. Image captioning with deep bidirectional lstms. In *Proceedings of the 2016 ACM on Multimedia Conference*, pages 988–997. ACM, 2016.

[24] Andrej Karpathy and Li Fei-Fei. Deep visual-semantic alignments for generating image descriptions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3128–3137, 2015.

[25] Oriol Vinyals, Alexander Toshev, Samy Bengio, and Dumitru Erhan. Show and tell: A neural image caption generator. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3156–3164, 2015.

[26] Shital Shah, Debadeepta Dey, Chris Lovett, and Ashish Kapoor. Airsim: High-fidelity visual and physical simulation for autonomous vehicles. In *Field and Service Robotics*, pages 621–635. Springer, 2018.

[27] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014.

[28] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*, 2014.

[29] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. In *Advances in neural information processing systems*, pages 3104–3112, 2014.

[30] Pierre Sermanet, David Eigen, Xiang Zhang, Michaël Mathieu, Rob Fergus, and Yann LeCun. Overfeat: Integrated recognition, localization and detection using convolutional networks. *arXiv preprint arXiv:1312.6229*, 2013.

[31] Madhur Behl and Rahul Mangharam. Interactive analytics for smart cities infrastructures. In *Science of Smart City Operations and Platforms Engineering (SCOPE) in partnership with Global City Teams Challenge (GCTC)(SCOPE-GCTC), 2016 1st International Workshop on*, pages 1–6. IEEE, 2016.

[32] Achin Jain, Madhur Behl, and Rahul Mangharam. Data predictive control for building energy management. In *American Control Conference (ACC), 2017*, pages 44–49. IEEE, 2017.

[33] Mariusz Bojarski, Davide Del Testa, Daniel Dworakowski, Bernhard Firner, Beat Flepp, Prasoon Goyal, Lawrence D Jackel, Mathew Monfort, Urs Muller, Jiakai Zhang, et al. End to end learning for self-driving cars. *arXiv preprint arXiv:1604.07316*, 2016.

[34] Stephanie M Merritt, Heather Heimbaugh, Jennifer LaChapell, and Deborah Lee. I trust it, but i don't know why: Effects of implicit attitudes toward automation on trust in an automated system. *Human factors*, 55(3):520–534, 2013.

[35] Jiun-Yin Jian, Ann M Bisantz, and Colin G Drury. Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics*, 4(1):53–71, 2000.

[36] Kevin Anthony Hoff and Masooda Bashir. Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, 57(3):407–434, 2015.

[37] Madhur Behl. F1/10 autonomous racing. 2018.

E. Data Management Plan

Data generated as part of this project, including research manuscripts and technical reports, instrumentation code, modeling code, training data, model descriptions, etc. will be managed using existing University of Virginia infrastructure for administering and maintaining digital research, with automatic nightly back up of source-code and documentation repositories), at no additional cost to the project. Some of the resources and data management practices are already in place, and being utilized by the PI in his research group. We will adopt and extend these practices for short-term data collection, retention and management. The PI, Madhur Behl, will have responsibility for coordinating and directing the retention and sharing of data generated through the proposed research activities. Because this work is collaborative across five PIs, many of the data management activities during the project period will reside with the coPIs and will be completed in adherence with NSF and university policies. A project website will be established as the main point of access to the generated artifacts including data, source code, publications resulting from the project, and the primary results obtained within the project. This project website will also serve to advertise undergraduate and graduate research opportunities. In addition, we will use Cyber-Physical Systems Virtual Organization (CPS-VO) to disseminate artifacts and information to the CPS/CISE community.

Expected Data

An outline of the data expected to be generated by the project is as follows:

1. *Human driving profile data*

   Data Collection Methods: Our driving behavior data collection methods include observations, interviews, surveys, and driving tests. We will obtain consent from all the participants. The driving experimentson the full scale simulator will be carried out in a safe, and private space in the PIs lab. Any observation notes will not use the participants' names or other identifying information; instead the information collected will be anonymous.

   Data Collection Tools: We will employ, photography, audio recordings, and video recordings to collect data. Our IRB protocol will explain in detail what we will record and justify the necessity for using the recording device. In our consent form we will clearly include a section informing the participants that we will be using a recording device. We will make provisions to destroy the recorded materials should the participant decide to withdraw from the study.

2. *Modeling algorithms and code:* A second major code artifact of our project will be the modeling algorithms and toolchains used to build our scenario models. We will publish the mathematical formulations of these techniques for archival by the publisher, and will maintain implementations of these techniques (*i.e.*, code) in version control. Concurrent with the publication of these techniques, we will release code implementing them using open source licenses.

Data Retention

During the project, the data associated with individual tasks will be stored by the PI using his laboratory's local servers and other resources. At regular intervals (quarterly), all the project related data will be copied and archived using the University of Virginia's digital repository. Participant data will be protected by carrying out the following measures. The team will encrypt the data and restrict access with access-level certification so that sensitive information will only be available to authorized personnel and used specifically for research purposes. All human subjects' data will be handled in accordance with the restrictions of

UVA's Institutional Review Board (IRB), which dictates the appropriate standards for protecting privacy and maintaining confidentiality of respondents. All participants will be informed transparently what data will be collected, and how these data will be reported and used. Note that any sensitive aspects (e.g., concerning human subject data), which may be provided to NSF program officers, may be withheld from public access. Conforming with IRB rules, human-subject data will only be held in anonymized form and will only be released according to IRB procedures. The PI and the research team will be in compliance with all NSF and university policies on research conduct. UVA policies govern the protection of human subjects in any research conducted at UVA, with UVA facilities, or by UVA faculty, staff, or students. Following completion of the project, the data and artifacts emerging from it will be stored for at least 5 years, and after that as long as the external website (or its successor) is maintained.

Data Formats, Short-term storage and dissemination

All the data described above and its accompanying documentation will be incrementally made accessible to researchers and the general public as mentioned before. All the computer codes will be implemented and made accessible to the scientific community in the form live web-tools. Selected source code, associated input/output files and documentation will also be released as the project and this computational modeling technology matures. The researchers retain rights to access and utilize data in whatever format but will not limit requester's ability to re-use or re-distribute processed data or materials. The data will be deposited in established repositories, for example the UVA institutional repository Libra. Libra is an open repository with public access. Therefore, care will be taken to ensure confidential and sensitive data are not shared through Libra. We reserve the right to delay release of project data for a period of time to allow for publication of research results. This period will not exceed five years following the project end date.

Long-term Data Storage and Preservation of access

The UVA Libra system provides servers, backup procedures, and other policies to minimize the change of data loss. In accordance with the University of Virginia policy RES-002, "Policy: Laboratory Notebook and Recordkeeping," the data will be preserved for a minimum of five years upon completion of the project. However the current preservation plan for Libra will be to preserve the data indefinitely. The Libra backup plan provides for data redundancy including off-site storage. If the Principal Investigator resigns from the university, the department chairperson for the lead department will become the custodian of the data and will assume all the responsibilities for data management, control and dissemination on behalf of the University of Virginia.

Policies and Provisions for Reuse and Distribution

The research team will share the research data, wherever appropriate, with the general public through Internet access (including social media), news articles, or reports. The university will regulate this public access in order to protect privacy and address any confidentiality concerns, as well as to respect any personal, proprietary or intellectual property rights. The research team will consult with the university's legal office to address any concerns on a case-by-case basis, if necessary. Terms of use will include requirements of attribution along with disclaimers of liability in connection with any use or distribution of the research data, which may be conditioned under some circumstances.

G. Facilities, Equipment, and Other Resources.

The research project will take place at the The Link-Lab - the Cyber-Physical Systems lab located on the University of Virginia's (UVA) campus. The facilities are partitioned into several laboratories that provide a complete environment for the design, fabrication, and testing of prototype hardware/software systems from initial concept to final implementation. These facilities include sufficient computing and prototyping resources for the proposed research, as described below:

A. UVA Link Lab - The PI and Co-PIs are members of the Link-Lab at UVA – this space is a new collaborative initiative on Cyber-Physical Systems (CPS) research and education at the University of Virginia. The lab is called "Link Lab" because it "links" multiple engineering departments through cross-cutting mechanisms such as shared lab space, staff, and conference rooms that house faculty and students from multiple disciplines. It houses approximately 20 faculty, 125 students, 3 research scientists, 3 staff, and 6 postdocs. An open floor plan promotes cross-pollination between research groups while at the same time using furniture and layout to provide sound insulation and reduce interruptions.The Link Lab has a large 3000 ft2 open space with moveable tables called the "Arena" that is designed for equipment staging, testbeds, experimental work, as well as a shared common space for seminars, presentations, or social events. It also includes a 2000 ft2 hardware prototyping lab. By locating the space immediately adjacent to the student and faculty desks, and close to the door, this layout will facilitate daily collaboration on this project. Figure 1, illustrates the Link-Lab floor plan.

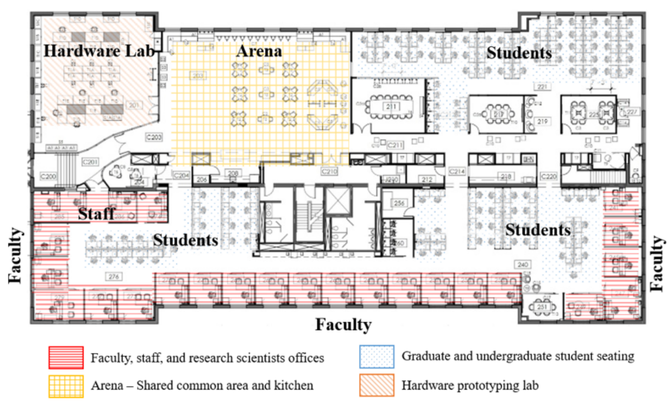

Figure 10: Link-lab floor plan

B. UVA Viz Lab The Viz lab is a facility at UVA designed to help faculty, staff and students explore various 3D visualization tools, such as virtual and augmented reality head mounted displays, for research and education purposes. The staff at Viz lab provide assistance with developing and evaluating virtual and augmented reality environments.

C. Autonomous Mobile Robots and CPS Lab: Part of the proposed autonomous vehicles testing will be carried out using the testbed available in Co-PI Bezzo's "Robotic and CPS Lab" which has a large, dedicated, state-of-the-art facility for mobile robotic systems development, prototyping, and control. The lab has been recently renovated and includes support for diverse system development activities with electronics workbenches, flexible space for assembling and experimenting with large demo systems, and secure storage in addition to the computing resources. This space covers an area of more than 900 ft2 with high ceilings and includes:

- Latest generation Vicon Motion Capture System: 8 Vantage cameras (@350 fps) with Lock system and Vicon Tracker software. This system allow sub-millimeter precision localization and tracking of multiple objects moving within the volume of the Lab space
- 2 Ascending Technologies Pelican quadrotors with carbon fiber body and equipped with 3rd generation Intel Core i7 CPU, IMU, pressure sensor, GPS, Lidar, stereo cameras, 2.4 GHz XBee link and WiFi - max airspeed 16 m/s, max climb rate 8 m/s, max thrust 36N, max payload 650g.
- 1 Ascending Technologies Firefly hexarotor with carbon fiber body and equipped with 3rd generation Intel Core i7 CPU, IMU, pressure sensor, GPS, Lidar, stereo cameras, 2.34GHz XBee link and WiFi (max airspeed 15 m/s, max climb rate 8 m/s, max thrust 36N, max payload 600g.)
- 3 Ascending Technologies Hummingbird quadrotors with carbon fiber body and equipped with IMU, pressure sensor, GPS, and 2.4 GHz XBee link - max airspeed 15 m/s, max climb rate 5 m/s, max thrust 20N, max payload 200g.
- Several Crazyflie 2.0 nano quadrotors equipped with IMU, pressure sensor, bluetooth, and Qi inductive charger.
- Several Parrot Bebop quadrotors equipped with IMU, two cameras, GPS, and sonar.
- 2 Clearpath Jackal unmanned ground vehicles, equipped with 3rd generation Intel Core i7 CPU, IMU, NovAtel SMART GPS, Velodyne 3D Lidar, Point Grey Flea3 camera, and WiFi - max speed 2 m/s, max payload 20 kg.
- 1 Black-i LandShark military grade 6 wheels UGV equipped with Intel Core i7 CPU, automated turret, Moog Quickset GeminEye, 100x zoom camera, thermal imager, 2 fisheye cameras, 1 camera, 2 Microstrain IMUs, 12 sonar rangers, 12 IR rangers, 2 Hokuyo UTM-30LX Lidars, GPS, and OCU (max speed 10 mph, max payload ¿ 200 lbs)
- 1 Stratasys uPrint SE Plus 3D printer
- 14 Cisco Systems, standard and high definition dome surveillance cameras with power- over-ethernet capability, along with hardware and software to support surveillance management and leading-edge video analytic completed with high-quality teleconference system and cloud capabilities

D. Computational Resources and Rivanna Compute Cluster - The university operates a Linux-based commodity cluster with a frontend named Fir. This cluster is managed by UVa Advanced Computing Services and Engagement and is open to faculty, staff, and graduate students at the University. Undergraduate students and university affiliates are eligible for accounts under faculty sponsorship. Fir is a large-memory cluster consisting of 92 nodes. Twelve of the nodes contain one dual-core 3-GHz Intel dual-core Xeon cpu with 32GB of RAM per node. Another 56 nodes are 8-core servers, with 48 GB per server. There are also 24 12-way nodes with 96 GB per node. Most of these cores are hyperthreaded, bringing the total number of logical cores that are eligible for no preemption to 1496. The interconnect for all these nodes is GigE.

Rivanna, launched in Fall 2014, is the Cray CS300, a 4800-core, high-speed interconnect cluster, with 1.4 PBs of storage. It is composed of 240 compute nodes, each with two 10-core processors and FDR Infiniband interconnect along with a parallel filesystem capable of providing about 25Gb/sec bandwidth. The Cray cluster combines large amounts of processing with large amounts of memory to provide a significant new resource for computationally-intensive research at UVA.

E. Laser Cutters - Campbell Hall has two Universal Laser Systems CO2 lasers. The 50 watt X660M has an 18”x32” bed. The 25 watt M-300 has a 12” x 24” bed capacity. Both can cut virtually any material other than metal, PVC plastics, or anything reflective. These machines can cut and engrave using both vector (lines/shapes) and raster (pixels) modes from virtually any software program.

F. 3D Printer / Rapid Prototyping - The Stratasys Dimension SST 3D printer uses Fuse Deposition Modeling (FDM) technology to build solid ABS plastic model prototypes from 3D stereolithography (stl) files. The machine has an 8”x8”x10” build envelope and builds models in layers down to 0.010 in. thickness.

G. 3-axis Miller and Routing - The MicroMill 2000 and MicroRouter from Denford, Inc. provide full 3-axis CNC and CAM machining capability. Using CAD/CAM software (EdgeCAM / MasterCAM) to generate G-code instructions for the machine, we can translate 2D profiles or 3D solid/surface geometry into machined parts. The router supports 12”x24”x2.5” travel with the ability to feed in and clamp longer stock materials from the side, while the mill supports approximately 9”x3.5”x6” of travel for a single machining operation. Common materials include wood, foam, plastic, aluminum, brass, copper, and mild steel. The machine is also capable of milling marble

H. 3D Digitizer / 3D Laser Scanner - Using technology from MicroScribe and NextEngine, the fabrication facility can both digitize and scan 3D objects and models into CAD systems. The MicroScribe point digitizer captures point, line, spline and surface information using common 3D modeling and CAD software. These can be used to generate surface and solid models of objects, models, topography, and reliefs. The NextEngine 3D Laser Scanner can scan up to a full 360 degree revolution around a small object, creating a full polygon mesh model for export into any 3D CAD software.

I. Software Resources - The University of Virginia has site licenses for a variety of software for basic computing needs as well as modeling and data analysis, including ANSYS, LabView, Mathcad, and MATLAB.