CPS Medium: Humanizing Autonomous Vehicles

May 6, 2018

CPS: Medium: Humanizing Autonomy

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

CPS: Medium: Humanizing Autonomy

1 Introduction

Two fatal self-driving-car accidents in March 2018 (Uber and Tesla) have cast doubt on whether the self-driving car can become the future of personal transport. For reasons unknown, the Tesla accident happened in broad daylight in pretty much perfect driving conditions. Just days before this, a fully automated 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 perhaps partially automated fleet, established driver perceptions about their trust and comfort in various vehicle technologies is critical to our understanding as to how one needs to facilitate behavior change.

Everyone is talking about autonomous driving. From automotive manufacturers, to consumer electronic giants, to software engineers, and academic researchers, driverless cars are at the forefront of everyone's imagination. There were more autonomous driving concepts at CES 2018 than ever before [1]. 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, it's been approached from many angles. But 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 many autonomous user experiences – what about those human factors? *Understanding why the vehicle is doing what it is doing will be a critical factor in trust and acceptance moving forth.* The industry is preoccupied with the race to make their cars be as smart as possible and more safe than human drivers. What's really important though, above all else, is how this technology can enable greater human mobility and hence improve their autonomy. The auto industry's 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. As the industry gets profoundly disrupted, we firmly believe that it's not just automotive insiders who have a valuable contribution to make. What's important is for car makers and service providers to embrace this moment 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 nuance of human behavior and trust, while exploiting technological and safety benefits of an autonomous vehicle.

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. Despite the advances in technology, there is a fundamental fear and distrust of autonomous vehicles. According to a Deloitte study, trust appears to be the biggest roadblock to adoption of self-driving cars in every country surveyed. South Korea ranks the highest with 81% of people expressing safety concerns about fully-autonomous vehicles. China has the lowest figures, with 62%, but that still represents a large majority of consumers. The US falls roughly in the middle, where nearly three-quarters of consumers (74%) believe that fully-autonomous vehicles will not be safe. For the other countries in the study the percentage of consumers who believe fully-autonomous vehicle will not be safe are: Japan at 79%, Germany at 72%, and India at 64% While everyone wants an AVs to be safe, the interpretation of safety, and the nature of trust in the vehicle or the system changes from person to person. Some perceive an AV in the same light as a malfunctioning robot, running amok around town, striking fear into passengers, fellow motorists, and pedestrians. On the other hand, even the sheer amazement experienced by those witnessing a fully

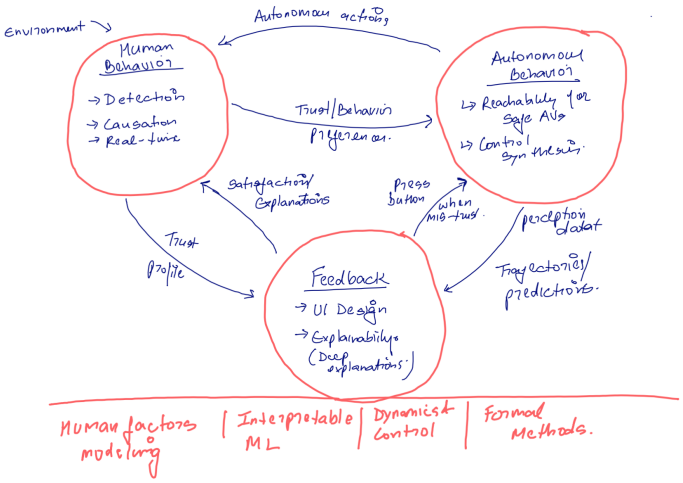

Figure 1: Placeholder overview figure

autonomous car could cause a distraction and lead to accidents. Then there are those who believe and expect that the AV will always "have your back" and is totally safe. *All of this boils down to one thing – trust*. 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 factors must be taken into consideration.

Add motivation for the emotional behavior part

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 end-to-end connected experience?

2 Research Description

2.1 Research Thrust 1: Behavior modeling

The ability to (1) detect, (2) assess and (3) control a person's emotions has been identified to be the predictor of success in relating to the people around us. By being able to read other's "emotional cues," not only we can better understand how they are feeling at a given time, but it also helps us to predict how they will respond in different scenarios. Research suggests that emotions are normally associated with specific events or occurrences (cite), and they can significantly influence our thoughts and behaviors. Additionally, we can use reason to evaluate our emotions, interpret them, and reassess our initial reactions to them. Therefore, by detecting certain events and occurrences, we may be able to assess and predict individual emotional states and use reason to soften their impact or "shift" their meaning. If the vision for the future AVs is to build "human-like" trust with the passenger(s), just like humans, AVs should be able to detect, assess, and control the passenger(s) emotions in real-time.

♣[2] In this project, we aim to better understand what environmental factors influence passenger(s) emotions and how these emotions influence driving behaviors. Specifically, we are interested in formally charac-

[2] AH: add a transition sentence

terizing emotions through a "context-dependence" approach, where spatial and temporal information can be automatically detected, analyzed, and interpreted. Through instrumentation of a number of manual and semi-autonomous ♣[3] vehicles, we will collect environmental data with ambient-condition sensors (i.e., thermal, indoor and outdoor noise levels, and lighting) in parallel with user-specific behavioral, emotional and physiological traits through cameras, wearables (e.g. smartwatch), pressure sensors (seat and steering wheel). Through facial recognition algorithms as well as application and social media data (e.g., Spotify, Pandora) we will monitor and identify changes in passenger emotions and behaviors. Through statistical analysis and machine learning techniques, we will identify models of behavioral and emotional traits in specific contextual and environmental settings. The specific research questions that will be addressed through this approach include:

1. What is the relationship between environmental/contextual settings and passenger(s) emotions and behaviors?
2. What is the taxonomy of behaviors and emotions in driving?
3. How behavior and emotions can non-intrusively and with least number of sensors be accurately detected?
4. What are environmental and social factors that can gain your attention? (research on signs)
5. Do people "trust" certain specific behaviors or conditions? ♣[4]?

As a first approach, we plan to observe trends in our data that both reveal (1) factors that influence specific human emotions and (2) the downstream consequences of particular emotions on their behaviors. For example, we will aim to understand 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. Through this approach, we will also be able to develop a database of emotions, behaviors and environmental changes. As it may be difficult to often identify triggers to particular emotions (i.e., passenger may appear already experiencing a particular emotion), we plan to evaluate which behavioral outcomes can be used as cues for particular emotions. For example, we may learn that when participants are feeling sad, they are more likely to deviate from their normal accelerating and decelerating behaviors or more likely to listen to certain genre of music. With this information, we will then assess if we can reliably predict particular emotions, given that a passenger behaved in a particular way.

After having a better understanding of these factors, we aim to develop psychological interventions to (1) reduce the negative outcomes of experiencing particular emotions ♣[5] and (2) reduce the likelihood of passengers' experiencing triggers that lead to emotions that have negative consequences (e.g., safety, trust).

With this information as a cue, we can understand that there might be a need to intervene or provide some feedback to the passenger at that time.

In this research, we will conduct real-life as well as simulation-based experimental studies to identify understand (1) the association between human emotional he causation of 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.

♣[6]

[3] AH: can we say autonomous cars? or should we just leave it at manual cars

[4] AH: for instance, do you trust your own behavior more than another type of driving behavior

[5] AH: this will be updated

[6] AH: Modify this table if needed

Type	Factors	Sensor/Algorithms
Environmental Sensing	Speed	Automatic Pro
	Location	GPS/Automatic Pro
	indoor and outdoor conditions	cameras
	Passengers' identity and count	cameras
	Weather Conditions	Camera and weather database
Human Sensing	Face features	eye tracking, camera, facial recognition
	Noise levels	noise-level sensors
	Passenger voice	Voice recorder, Natural Language Processing
	Brake and acceleration rate	Automatic Pro
	Grip on steering wheel	pressure sensor
	Music and other social media use	APIs (e.g., Spotify)
	Social interaction	Camera and voice recorder

♣[7]

Spill the magic dust Arsalan

2.2 Research Thrust 2: Reach-ability analysis based safety region estimation

This is all you Nicola

Lu adds subsection on control synthesis

The goal of this section is to define a method to quickly assess safety (and perhaps trust) online and then use this reachability analysis to close the loop and correct the system to maintain safety and increase trust

Here discuss recent work on reachability analysis for safety then add the new ideas on online reachability using machine learning and control to predict future states while guaranteeing safety and adapt.

Approach: ♠[8] This thrust is divided in two main parts, as outlined in Figure♠[9]. We start by proposing a reachability-based approach in which we leverage knowledge about the system dynamics and process, sensor, and environmental uncertainties to determine the future states over a finite time horizon. The computed reachable sets will be used to determine the first time that a safety-critical event may occur. More specifically, in this work we are interested in the fist time that the AV may: i) collide with any obstacle, both static and dynamics ii) deviate from the desired trajectory, and/or iii) suffer a failure ♠[10]. Based on these events, we will then design a safe reinforcement learning approach to adapt and correct the behavior of the system to maintain safety and increase trust ♠[11]. Here we propose also techniques for fast verification and monitoring ♠[12].

[7] AH: add about how cars can enhance the driver emotions as well but selecting routes or choosing music or behaving in a certain way...

[8] NB: need to improve this section

[9] NB: add figure

[10] NB: what do you all think? Can we add somethng related to trust?

[11] NB: it would be nice to connect trust here

[12] NB: I will elaborate and rephrase

Online Prediction

Reachability-based Approach: Traditional model-predictive or finite horizon controllers (MPC) [2, 3, 4] can estimate future states and inputs of a system, given a correct knowledge of its model. On the other hand, reachability analysis (RA) [5, 6, **?**] 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 [7, 8, 9].

In this work 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 and other vehicles configuration) in which the AV is operating dictates the reaction time to switch to a different mode of operation. Thus, predicting as accurately as possible the different situations that can occur in the near future will allow a better safety assessment, ♠[13] and will guide a better planning of actions to perform. Let us 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.

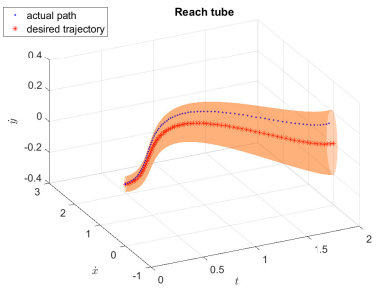

Figure 2: Velocity Reachable tube for a quadrotor following a straight line trajectory for 2s.

In this work we propose reachability analysis techniques for safety monitoring AVs and present 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 $\Delta T = [0, T]$, [10, 11]. 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:

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

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. 2, the $[\dot{x}, \dot{y}]$ velocity reachable tube for an autonomous quadrotor aerial vehicle following a straight line trajectory for 2s ♠[14], 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.

To deal with unmodeled and stochastic system dynamics, we propose a methodology to compute and maximize the probability of maintaining the state of the system within a certain safe region and decide to

[13] NB: THUS MORE TRUST

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

represent the system as a stochastic hybrid system whose dynamics can be influenced by a control input [12, 13]. Unlike previous approaches, the safe set can be time-varying [14, 15]. 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. These are the set of initial conditions for the system, such that there exists a Markov policy capable of maintaining the state of the system within the safe set with a probability greater than a prescribed safety level [16, 17, 12]. If the objective is to minimize the probability that the system will exit the safe set, then we can formulate the reachability problem as a stochastic optimal control problem, but this time with a cost that is the maximum of a function of the state over the time horizon. Here, again, DP will be considered to determine probabilistic maximal safe sets for Markov policies, [12].

Runtime Verification for Safety-Aware Autonomous Vehicle Operations:

Offline verification of autonomous vehicles operations may not scale to larger systems where exploring all possible contingencies can be computationally prohibitive. For example, an autonomous car can continuously estimate clear roadways and potential obstacles for which the planned trajectory must be verified for safety. Performing offline mission verification would require considering all possible environments (roadways, weather conditions, obstructions, *etc*.). Moreover, virtually all prior research in the area of formally verifying complex mission planning (expressed using temporal logics) assumes the plan is developed off-line [**?**, **?**, **?**]. However, it is imperative to monitor – at runtime – the vehicle to check that the assumptions and properties are indeed satisfied and will continue to hold as the mission evolves. In this thrust, we propose to enable safety-aware learning by performing runtime verification of the control components for mission-critical scenarios where the systems may violate safety constraints.

Runtime verification [**?**] is a lightweight formal methods technique for online monitoring of evolving traces of observations from a system execution with respect to a formal specification. Like most other established formal methods, runtime verification techniques target traditional safety-critical systems, developed with respect to specifications, fixed during the system design. In AVs, it is possible that monitoring specifications will be changing dynamically. For example, in the presence of external disturbance (like changes in surface type or wind and other atmospheric disturbacnes acting on the vehicle), the system may need to switch mode of operation, reconfigure, adapt its control strategy, and replan its trajectory to fight the disturbance and achieve its desired goal, all while maintaining a certain level of safety assurance. Thus, online prediction for runtime verification is critical to AVs. 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. In this thrust we consider two complementary approaches to runtime verification for safety-aware learning.

Our first approach will be to develop techniques for online verification of missions with AVs. While such problems are computationally very hard, we will leverage our prior work that relies on satisfiability modulo theories (SMT) solvers to address the system of constraints that mix Boolean variables with real variables [**?**, **?**]. One of the major challenges we will address is to verify reactive plans that dynamically compose primitives in order to quickly respond to changes in the environment or mission. 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.

While recent advances in SMT solvers can be leveraged to perform *fast* verification of controllers, the

worst-case complexity of such verifications remain troublesome. This motivates our second approach to run-time verification wherein we abstract the safety conditions such that monitoring can be performed efficiently using predictive techniques (e.g., reachability presented in the previous section). 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) [18] 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 [18] 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.

Safe Reinforcement Learning-based Adaptation:

The safety assessment introduced in the previous section is used to guide the AV actions to maintain safety, i.e., to maintain the AV within safe reachable sets. To this end we propose a safe reinforcement learning approach to select appropriate actions for the autonomous system and maintain safety at all the time. The innovation here is that we consider a human in the loop approach in which the desired actions from the user are taken into consideration while determining actions of the AV without violating safety constraints. ♠[15] For example, ♠[16] ♠[17] 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 [19, 20]). In our case the reward is a combination between user preference and safety ♠[18] Another main issue of learning behavioral policies in domains that are unknown at first (especially in autonomous vehicles) is that of efficiently bringing the agent from a tabula-rasa state to a condition where the agent is acting as close to optimality as possible. This notion is also known as regret minimization [19, 20] and is closely related to the topic of trading off exploration of the environment (to sample previously unseen parts of the state-action space) with exploitation of the knowledge accumulated so far. An agent that enters the world would therefore need to explore the environment by applying actions and gather data which are as informative as possible so that a policy, encoding the knowledge accumulated so far, is derived. The goal is to go from a situation where the system is mostly exploring the environment to a situation where it is mostly exploiting the accumulated knowledge. When the transition probabilities and rewards of a Markov Decision Process (MDP) are known, an agent can obtain the optimal policy without any interaction with the environment. However, exact transition probabilities are difficult for experts to specify and may not be known completely a priori or change over a mission due to aging of the system or unforeseen disturbances and faults. With these considerations in mind, one option left to an agent is a long and potentially costly exploration of the environment. One such algorithm is the E3-algorithm presented in [21] in which the basic idea is to repeatedly apply an exploration policy, i.e., one that tries to visit state-action pairs whose transition dynamics are still inaccurately modeled. After a polynomial number of iterations, it will deem itself to have modeled enough of the MDP accurately. Then, it will apply an exploitation policy,

[15] NB: need to rephrase but this is the concept we need to convey here

[16] NB: let's find a proper example here like a user that likes to drive always in the center lane of a highway and is in a rush so it's pushing but there is traffic.

[17] NB: still working on this section...I'm going to use some of this material but the key idea is to be able to adapt your actions or the AV actions based on how safe the system believes to be

[18] NB: we need to define safety somewhere

which (given the current MDP model) tries to maximize the sum of rewards obtained over time . In [22] the authors demonstrate that the E3-family of algorithms [21] is often unacceptable because it requires executing policies that explore also unsafe parts of the workspace, including parts that would lead to a crash of the CPS (e.g. crash of the helicopter discussed in [22]). The same authors consider reinforcement learning in systems with unknown dynamics and propose an apprenticeship learning approach, in which an initial teacher demonstration of the task to be learned is presented and then by using only exploitation policies over the training data from the teacher to learn the optimal policy. Farther work from the same authors in [23] deal with scalability issues in reinforcement learning offering a hybrid algorithm that requires only an approximate model and a small number of trials to obtain a near-optimal performance in a real system. In [24], the authors propose another alternative to the suboptimal exploration of the state space presented in the previous works: given initial (possibly inaccurate) specification of the MDP, the agent determines the sensitivity of the optimal policy to changes in transitions and rewards. It then focuses its exploration on the regions of space to which the optimal policy is most sensitive. This technique named Active reinforcement learning enables this type of exploration: it uses sensitivity analysis to determine how the optimal policy in the expert-specified MDP is affected by changes in transition probabilities and rewards of individual actions. This analysis guides the exploration process by forcing the agent to sample the most sensitive actions first. It is shown that the proposed exploration strategy performs well on several control and planning problems. 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.

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. More specifically in this task we propose an iterative reinforcement learning (RL) framework to determine the control policy and refine the model of the system. The abstract setting can be described in the following terms: a 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 can be 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.

♠[19] An even reacher extension to this planned work will consider disturbance and uncertainties in the model and environmental dynamics as an adversarial player with an unknown reward that we would like to learn. This specific problem can be solved with a number of inverse reinforcement learning algorithms. For this case we can still model the decision problem as an MDP M^{D} in which reward is structured around misclassification cost experienced in terminal states that correspond either to allowing the AV to proceed to its intended terminal state or to change control law. The state definition for M^{D} includes all of the vehicle state elements as well as a belief function for the vehcile reward function. Because it includes a belief state, MDP will generally suffer from the curse of dimensionality and be intractable from the perspective of full-width planning methods. To address this issue, we propose to make use of work on Monte-Carlo planning in large partially-observed MDPs (POMDPs) [25]. The basic idea is to use a simulator as a generative model of

[19] NB: I may remove the following section

the decision process. The simulator is used to generate sequences of states, observations, and rewards that are used to update the value function. Following [25] and others, sampling will be done according to a Monte-Carlo trees search algorithm. Other methods, such as DESPOT [26], and recent POMDP computational methods [27, 28, 29, 30] will also be considered.

Correct-by-construction controller synthesis

In co-PI Feng's prior work [**?**, **?**], we proposed a stochastic game based approach to automatically synthesize correct-by-construction control protocols (mission plans) for an unmanned aerial vehicle (UAV) while accounting for the uncertainty of human operators. By modeling the UAV as a player and the human operator as another player in the stochastic game, we reduced the problem to finding a winning strategy for the UAV against all strategies (including the worst-case) of the human operator. The synthesized UAV control protocol is guaranteed to satisfy the mission objectives (expressed as temporal logic specifications) imposed on the stochastic game. Furthermore, using different human operator models, our approach enables the synthesis of operator-dependent optimal mission plans for the UAV, highlighting the effects of operator characteristics (e.g., workload, proficiency, and fatigue) on the mission performance. The proposed research will build on our prior work and develop new methods for synthesizing control protocols for autonomous vehicles while accounting for the behavior models leanred in Thrust I. We will propose an approach build upon techniques of strategy synthesis from stochastic multi-player games [**?**], which often involve the computation of winning strategies based on value iteration algorithms. We will take into account cognitive mechanism and develop new techniques for the automated synthesis from the logic PRTL*. In addition, we will consider multi-objective properties that require simultaneous satisfaction of multiple safety and trust properties. For example, "agent A (human)'s trust in agent B (autonomous vehicle)'s capabilities to keep a safe distance with the front vehicle is at least 80%" *and* "the vehicle should maintain a minimal speed of 60 mph on the highway". We will develop techniques that compute an ε-approximation of the Pareto set of optimal trade-offs between the individual properties. The synthesized strategies can inform the design of autonomous vehicles, e.g., what actions should the vehicle take in order to gain human trust?

2.3 Research Thrust 3: Feedback Design

The purpose of this research thrust is to develop and validate what feedback should the autonomous vehicle provide to the passengers.

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, the seat belt, crumple zones, ABS, airbags, electronic stability control and GPS satellite navigation, 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. At the broadest level, many of the technologies available in modern cars do not appear to have been developed with a particular user-centred approach. They exist because the technology has become available to perform a specific function.

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. 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 on roads or situations when they shouldn't.

We hypothesize, that for autonomous vehicles, trust comes from two factors: *predictabilit*y, 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.

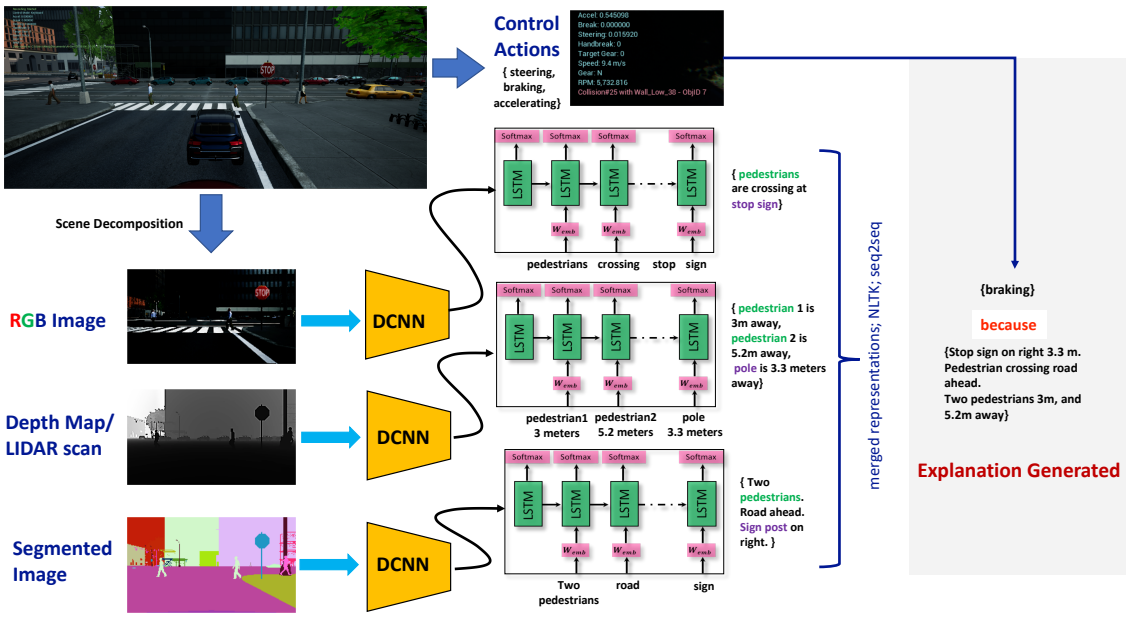

Figure 3: Deep-Explanation generation: Each dimension of the scene decomposition is used as an input to caption generation. Representation matching, and seq2seq are then used to generate a likely explanation for the predominant action stream.

Given a proile of the passenger's behavior and emotional needs, from Section 2.1 the goal of this researh 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 *predicitability* component of their trust in the autonomous vehicle.

2.3.1 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 [31, 32, 33, 34, 35] Not only must caption generation models be able to solve the computer vision challenges of determining what objects are in an image, but they must also be powerful enough to capture and express their relationships in natural language. For this reason, caption generation has long been seen as a difficult problem. 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. This requires the joint use of both Computer Vision and Natural Language Processing techniques. Yet 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. To capture the correlation between two modalities i.e. visual and natural language we need to map both these to some same space so at learn the relation between them or say we need to learn the multimodal joint

model. Models that uses different deep neural networks like convolutional neural network (CNN), long short term memory(LSTM) networks, recurrent neural network(RNN) to implicitly learn the common embedding. These by far gives the best result on all common datasets of caption generation 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 attention-based 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 [**?**] simulation.

Consider the scene shown in the Figure 3; 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 main inspiration of 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) [**?**, **?**, **?**] 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. Over the last few years it has been convincingly shown that CNNs can produce a rich representation of the input image by embedding it to a fixed-length vector, such that this representation can be used for a variety of vision tasks [**?**]. Hence, it is natural to use a CNN as an image "encoder", by first pre-training it for an image classification task and using the last hidden layer as an input to the RNN decoder that generates sentences. For the RGB image captioning we will use Neural Image Captioning method, as described in [35].

2.4 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 images, and with some small tweaks 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 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 $p(S_t|I, S_0, ..., S_{t-1})$. It is instructive to think of the LSTM in unrolled form - a copy of the LSTM memory is created for the image and each sentence word such that all LSTMs share the same parameters and the output of the LSTM at time $t-1$ is fed to the LSTM at time t. Following our exapmple, if we denote by I the input image and by $S = (S_0, ..., S_N)$ a true sentence describing this image, the unrolling procedure reads:

$$x_{-1} = \text{CNN(I)} \tag{3}$$

$$x_t = W_e S_t \quad t \in \{0, \cdots, N-1\} \tag{4}$$

$$p_{t+1} = \text{LSTM}(x_t) \quad t \in \{0, \cdots, N-1\} \tag{5}$$

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 3.

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.

2.5 Explanation generation via captioning merging

Using the CNN and LSTM networks on each of the modalities, we obtain a caption for each modality. 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 3, each of the captions refers to a pedestrian. We use this commonality to merge the captions into a single explanation. We will 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 [**?**]. Sequence-to-sequence (seq2seq) models have enjoyed great success in a variety of tasks such as machine translation, speech recognition, and text summarization. Specifically, we use a technique similar to Neural Machine Translation (NMT), widely used in language translation, but much simpler. Traditional phrase-based translation systems performed their task by breaking up source sentences into multiple chunks

and then translated them phrase-by-phrase. This led to disfluency in the translation outputs and was not quite like how we, humans, translate. We read the entire source sentence, understand its meaning, and then produce a translation. Neural Machine Translation mimics that. 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.

We envision that these deep explanations can offer insight into the neural networks responsible for the perception and scene understanding for self-driving cars, and soon may also play a role in planning and control for the autonomous cars [**?**]. By providing readable explanations of the actions of the car, we provide context and feedback to the passenger, enabling an increase in the trust between the passenger and the vehicle.

need to conclude with how this ties to thrust 1 and 2

2.6 Predictability via UI design

3 Evaluation/Experimentation Plan

- We need to describe in detail all the experiments and testbeds we will use/develop.
- this is emphasized int he solitication
- describe how we will recruit subjects
- emphasize the real world experimentation plan

This section should describe how the research concepts proposed will be demonstrated and validated. It should present metrics for success. It should identify critical experiments, and describe how the research will be demonstrated, including through simulation, prototyping, and integration with real (including sub-scale) cyber-physical systems. For Medium and Frontier projects, the validation plan must include experimentation on an actual cyber-physical system.

Lu describes the full scale driving simulator

Figure 4 shows the setup of an academic-scale testbed that co-PI Feng's group has partially built 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. Thus, we expect to collect data about realistic human response during the driving. The human

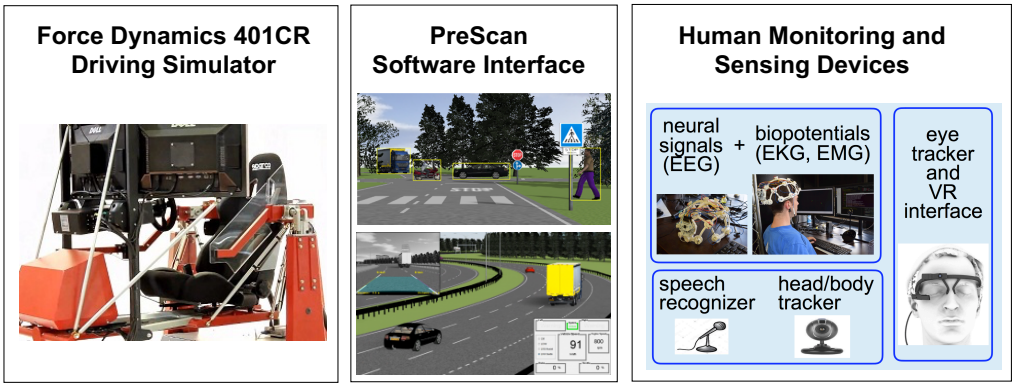

Figure 4: Our partially built testbed of human-autonomous vehicle interactions. The human is monitored by a number of sensors, and interacts with the driving simulator through the PreScan software interface.)

is interfaced to the hardware platform through PreScan, which is a software tool designed as a development environment for Advanced Driver Assistance Systems (ADAS) and Intelligent Vehicle (IV) systems. These are systems with sensors that monitor the vehicle's surroundings and that use the acquired information to take action. Such actions may range from warning the driver of a potentially dangerous situation to actively evading hazards by means of automatic braking or automatic steering. PreScan can be used for designing and evaluating ADAS and IV systems that are based on sensor technologies such as radar, laser, camera, ultrasonic, GPS and C2C/C2I communications. The "human monitoring and sensing" block in Figure 4 encloses the sensors that 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.

Nicola describes the experiment setup for reachability analysis experimentation

Madhur describes F1/10 - could be used for reachibility + control synthesis

4 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, and 4 graduate students.

4.1 Roles and responsibilities:

4.2 Project tasks and time-line:

4.3 Risks and mitigation plans:

4.4 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. Behl's role on this project is the development of a novel modeling and control framework called Data Predictive Control (DPC) for assisting decision makers in understanding and managing interdependent critical 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.

Arsalan Heydarian:

i dont have anything

Nicola Bezzo:

Lu Feng:

5 Broader Impacts

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. 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) 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, (c) 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 (d) 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.

5.1 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. PI Behl has pioneering two new courses for the graduate and undergraduate teaching:

1. **Principles of Modeling in Cyber-Physical Systems** : This course provides 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 come out of this course with advanced and transferrable knowledge of model-based design methods and tools, and will be ready for tackling multi-disciplinary systems projects. All the lectures are available online for anyone to view.

2. **F1/10 Autonomous Racing** - Principles of Perception, Planning, and Control. - 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 free and 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. PI Behl's course materials available online have been adapted for teaching CPS and Autonomous Systems courses at UT Austin, and Clemson University.

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.

5.2 K-12 outreach Impacts:

5.3 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.

5.4 Dissemination Impacts:

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. PI Behl serves on the editorial board of the SAE International Journal of Connected and Automated Vehicles.

6 Team

Arsalan adds Bio

Nicola adds Bio

Lu adds bio

Lu Feng is an Assistant Professor at the Department of Computer Science and Department of Systems and Information Engineering at the University of Virginia. She is also a member of the Link Lab - the

center of research excellence in Cyber-Physical Systems at the University of Virginia. Previously, she was a postdoctoral fellow at the University of Pennsylvania and received her PhD in Computer Science from the University of Oxford. Her research focuses on assuring the safety, trustworthiness and performance of cyber-physical systems, drawing on formal methods, machine learning and control. She has received several awards including NSF CISE CRII Award, James S. McDonnell Foundation Postdoctoral Fellowship, Rising Stars in EECS, UK Engineering and Physical Sciences Research Council Scholarship, and Cambridge Trust Scholarship.

Dr. Feng's contributions are on temporal logic planning for autonomous robots. Her recent work has focused on synthesizing control protocols for robots that interact with human operators [**?**, **?**] and providing counterexamples as diagnostic information for robotic planning [**?**, **?**]. She has also develop novel learning-based approaches for the compositional reasoning of probabilistic models with respect to temporal logic specifications [**?**, **?**, **?**, **?**].

Madhur adds bio

References

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

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

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

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

[10] 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.

[11] 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.

[12] Alessandro Abate. *Probabilistic reachability for stochastic hybrid systems: theory, computations, and applications*. University of California, Berkeley, 2007.

[13] Abraham P Vinod, Baisravan Homchaudhuri, and Meeko MK Oishi. Forward stochastic reachability analysis for uncontrolled linear systems using fourier transforms. In *Proceedings of the 20th International Conference on Hybrid Systems: Computation and Control*, pages 35–44. ACM, 2017.

[14] Dimitri Bertsekas. Infinite time reachability of state-space regions by using feedback control. *IEEE Transactions on Automatic Control*, 17(5):604–613, 1972.

[15] B Picasso-A Bicchi. Control synthesis for practical stabilization of quantized linear systems. *Politecnico di Torino*, page 397, 2005.

[16] Claire Tomlin, John Lygeros, and Shankar Sastry. Synthesizing controllers for nonlinear hybrid systems. *Hybrid Systems: Computation and Control*, pages 360–373, 1998.

[17] Andrea Balluchi, Luca Benvenuti, Maria Domenica Di Benedetto, Guido M Miconi, Ugo Pozzi, Tiziano Villa, Howard Wong-Toi, and Alberto L Sangiovanni-Vincentelli. Maximal safe set computation for idle speed control of an automotive engine. In *Hybrid Systems: Computation and Control*, volume 1790, pages 32–44. Springer, 2000.

[18] 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.

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

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

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

[22] 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.

[23] Pieter Abbeel, Morgan Quigley, and Andrew Y Ng. Using inaccurate models in reinforcement learning. In *Proceedings of the 23rd international conference on Machine learning*, pages 1–8. ACM, 2006.

[24] Arkady Epshteyn, Adam Vogel, and Gerald DeJong. Active reinforcement learning. In *Proceedings of the 25th international conference on Machine learning*, pages 296–303. ACM, 2008.

[25] David Silver and Joel Veness. Monte-carlo planning in large pomdps. In *Advances in neural information processing systems*, pages 2164–2172, 2010.

[26] Adhiraj Somani, Nan Ye, David Hsu, and Wee Sun Lee. Despot: Online pomdp planning with regularization. In *Advances in neural information processing systems*, pages 1772–1780, 2013.

[27] Hendrik Baier and Mark HM Winands. Nested monte-carlo tree search for online planning in large mdps. In *ECAI*, volume 242, pages 109–114, 2012.

[28] Arthur Guez, David Silver, and Peter Dayan. Efficient bayes-adaptive reinforcement learning using sample-based search. In *Advances in Neural Information Processing Systems*, pages 1025–1033, 2012.

[29] Arthur Guez, David Silver, and Peter Dayan. Scalable and efficient bayes-adaptive reinforcement learning based on monte-carlo tree search. *Journal of Artificial Intelligence Research*, pages 841–883, 2013.

[30] Ngo Anh Vien, Wolfgang Ertel, Viet-Hung Dang, and TaeChoong Chung. Monte-carlo tree search for bayesian reinforcement learning. *Applied intelligence*, 39(2):345–353, 2013.

[31] 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.

[32] 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.

[33] 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.

[34] 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.

[35] 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.

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.

F. Budget Justification

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.

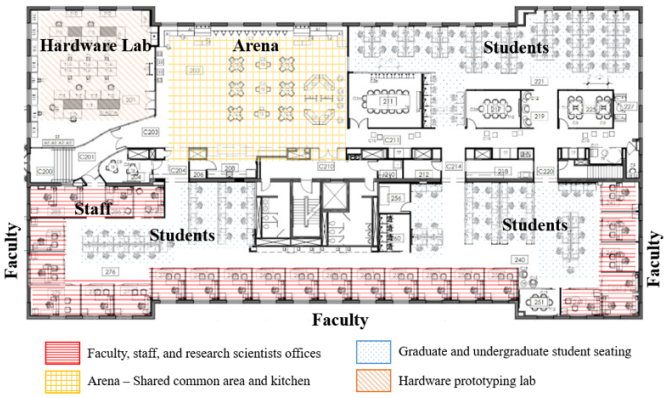

Figure 5: Link-lab floor plan

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.