Research Statement

Suguman Bansal
Assistant Professor
School of Computing, Georgia Institute of Technology
https://suguman.github.io/

Artificial intelligence (AI) is on the verge of revolutionizing the development of computational systems. In the future, software and hardware development will be driven by technologies emerging from AI and machine learning (ML). Early indications are already visible in the design of complex decision-making systems, such as NVIDIA's autonomous car and Deepmind's Go playing system. Yet, the primary hurdles of AI/ML prevent their wide-scale adoption since real-world deployment necessitates **trust in software and hardware**. Without trust, AI will not realize its potential, especially in safety-critical and security-critical domains including robotics and automation, medical diagnostics, finance and banking etc. My research vision is to make AI trustworthy through tight integration of formal methods (FM) and programming languages (PL) with AI/ML.

Towards **trustworthy computing**, my research so far has focused on simplifying the design, development, and deployment of safe and reliable intelligent systems. It is practically impossible for a system engineer/developer to predict all future possibilities in the environment and determine the correct response of the system. To this end, I have worked on **automated controller synthesis**, i.e., the automated generation of control systems from a given high-level specification. This declarative paradigm simplifies the task of designing correct controls, as it allows the designer to focus on the intent rather than the low-level details of the controller. Existing work on control synthesis for real-world systems, however, largely fail to offer formal guarantees of correctness. In fact, the synthesis problem is known to be challenging and undecidable for most real-world systems. These challenges stem from two sources: the environment and the specification. Real-world environments and specifications are complex. As either increases in complexity, assurance guarantees from synthesis become elusive.

My first research thrust addresses **complex environments**. I have co-developed scalable reinforcement learning (RL) algorithms that also offer assurance guarantees, in sharp contrast to prior work that neither scales to complex long-horizon tasks nor offers correctness guarantees [3,6]. RL is a data-driven approach that learns control policies by sampling the environment, rendering RL suitable for complex environments including ones with non-linear dynamics, are difficult to model, and may even be unknown. Using temporal logic for task specification, we develop the first algorithm and tool that generates neural-network policies in realistic continuous environments that can be formally certified [1,5]. We have evaluated the tool on state-of-the-art environments provided by OpenAI Gym + MuJoCo. My second research thrust addresses **complex specifications**. Specifications for long-horizon tasks routinely combine qualitative temporal goals with quantitative goals. Prior synthesis solutions from these richer specifications lack theoretical guarantees due to incompatibility between logic-based and numerical-methods based approaches for temporal and quantitative reasoning, respectively. To this end, I have proposed a novel purely logic-based framework for quantitative reasoning that offers rigorous guarantees of correctness [8,9]. This is a *paradigm shift* from traditional quantitative reasoning which requires numerical methods (optimization, linear programming, min-max games etc). In addition to advancing the state-of-the-art with the first sound algorithm for reactive synthesis from temporal and discounted-sum quantitative goals, algorithms built on our logic-based framework [10] have demonstrated computational efficiency and practically scalable, robust, and resilient in performance [7]. These works have found applications in IoT devices [12,13], robotics [5] and AI planning [10] so far, with future extensions to autonomous drones and vehicles.

Our work has been published at premier conferences in **formal methods and programming languages** (CAV×5, FoSSaCS, TACAS, POPL) and **artificial intelligence and machine learning** (AAAI×2, NeurIPS). My research has been supported by a USD ∼250K NSF/CRA postdoctoral fellowship under the prestigious Computing Innovations Fellow program. I was named one of the MIT EECS Rising Stars in 2021.

Thrust I. Learning-Based Synthesis for Complex Environments

Reinforcement Learning (RL) combined with neural-networks (NN), has made remarkable strides in control synthesis in real-world domains, including challenging continuous (infinite-state) environments with non-linear dynamics or unknown models. Yet, current RL approaches are poorly suited for control synthesis for long-horizon tasks for a number of reasons. First, typically control tasks in RL are specified in the form of rewards. Providing a suitable reward function for complex, long-horizon tasks can be daunting. Second, RL algorithms are inherently myopic, as they respond to immediate rewards, which may cause the algorithm to learn optimal policies for the short term but below-par policies in the long term. Lastly, the learnt NN-policies offer poor degree of assurance: Neither are these policies interpretable and nor can they be verified against the desired task.

In my research program, I have developed algorithms for RL from high-level specifications, specifically temporal logic, as opposed to rewards. The use of temporal logic resolves the first issue and reduces the burden of task specifications. We investigate theoretical and practical solutions to generate policies with high assurance.

Lack of theoretical guarantees. We present an impossibility result on RL from temporal specifications with rigorous mathematical guarantees. We prove that one cannot learn a near-optimal policy, w.r.t. probability of satisfaction of the specification, with high confidence. In other words, there does not exist an algorithm to learn policies from specifications with rigorous PAC guarantees [3]. The negative result already holds on learning for a safety specification in an environment with finitely many states. This occurs due to the *non-robustness* of temporal specifications, i.e, small perturbations in the environment can incur drastic changes in near-optimal and optimal policies w.r.t. probability of satisfaction.

Assurance in practical solution. Given the theoretical limitations, we focus on developing practical solutions to learn policies from specifications that also render assurance. For this we develop a novel compositional RL algorithm DiRL [5] which (a). scales efficiently to long-horizon tasks even in challenging continuous domains, and (b). generates a compositional NN-policy such that each sub-policy refers to a subtask in the original task. A key feature of the learnt compositional NN-policy is that each subtask is a reachability task with safety constraints. The benefit we can derive is that existing NN verification algorithms are known to scale to such simple goals. Thus, despite the inability to learn a policy with formal guarantees of correctness, we learn a policy for which each NN-component (subpolicy) can be formally verified, thus obtaining a much higher degree of assurance than the statistical guarantees offered by testing and other prior approaches.

To address the scalability issue on long-horizon tasks, the critical insight of DiRL is to learn a policy that adheres to the structure of the temporal specification as opposed to one that is globally optimal. For this, it leverages the structure of the specification to decompose the RL problem into a high-level planning problem and a set of low-level RL control problems for a simpler subtasks. Then, it interleaves model-based high-level planning with model-free RL to compute a policy that maximizes the probability of satisfying the specification subject to its structure. The high-level planning enables DiRL tackle myopia as it rejects locally optimal subtasks in favor of locally suboptimal subtasks that better support the satisfaction of the specification in the long term.

DiRL is available as an open-source tool [1]. It has been evaluated on long-horiozn control tasks in challenging continuous environments, including realistic environments from OpenAI Gym + MuJoCo (e.g. non-linear simulations of Baxter robotic arm). Empirically, we observe that DiRL significantly improves the scalability and performance of learning from specifications: The sample complexity of DiRL scales roughly linearly in the size of the specification, whereas the existing baselines rapidly degrade in performance.

Guarantees in finite-horizon environments. We also investigate environments with finite (fixed) horizon. In this case, we prove that learning from temporal specifications reduces to learning from (undiscounted-sum) rewards. Furthermore, when the environment has finitely many states, we show learning algorithms also offer probabilistic guarantees of correctness. Focusing on competitive multi-agent systems, we present the first sound algorithm to learn near-optimal Nash equilibrium with high confidence: The learnt joint policies are near-Nash equilibrium under PAC guarantees [6]. Additionally, the equilibria are prioritized by high social welfare.

Our algorithm is based on a novel search-and-verify framework. In the search phase, we execute a compositional RL approach to learn joint policies and rank them in decreasing order of social welfare. For the verification phase, we develop the first polynomial-time PAC algorithm to check if a given joint policy is near-Nash. Within this, the verification problem is reduced to solving two-agent zero-sum games, in an unknown environment. Each zero-sum game evaluates whether an agent can unilaterally deviate profitably. Each such game is solved using a model-based self-play RL algorithm after converting temporal specifications to undiscounted-sum rewards. Empirical evaluation demonstrates that our approach computes Nash equilibrium with high social welfare, whereas state-of-the-art baselines either fail to compute the equilibria or compute ones with lower social welfare.

Thrust II. Logic-Based Synthesis from Complex Specifications

Reactive synthesis (RS) is the automated construction, from a high-level description of its desired behavior, of a reactive system that continuously interacts with an uncontrollable (adversarial) external environment. Logic-based approaches for RS from temporal logics offer rigorous guarantees of correctness and practical scalability despite its computationally hardness [11]. Real-world specifications, however, routinely combine temporal logic with quantitative objectives. A commonly appearing quantitative objective in long-horizon tasks is the *satisficing goal*, i.e., to ensure that the discounted-sum reward exceeds a given threshold value. Existing solutions for RS combining temporal and satisficing goals neither offer theoretical guarantees nor scale. The primary reason is that traditionally satisficing goals are solved using numerical methods (optimization-based algorithms) which are incompatible with logic-based methods for temporal goals.

In my research program, I have developed a novel logic-based framework for formal reasoning of quantitative goals to resolve the incompatibility by eliminating the dependence on numerical methods. We have shown the benefits of this framework in reactive synthesis and other applications in quantitative reasoning including quantitative verification, decision-making under bounded rationality, and infinite-precision arithmetic.

Logic-based quantitative reasoning. I have introduced a novel paradigm, called comparator automata (comparators, in short), in which reasoning about a large class of quantitative properties is reduced to the reasoning about logical properties [9]. When the reduction is efficient, one could then employ techniques from logic-based methods, which (a). offer theoretical guarantees and (b). enhance scalability as logic-based approaches have witnessed significant progress in scalability. This contrasts with prior work, which viewed quantitative and logic-based reasoning as disjoint where neither could benefit from the other.

The defining insight underlying comparators is to recognize that the atomic operation in quantitative reasoning is to compare the quantitative properties between system runs, rather than the traditional numerical-methods view of determining the "optimal" solution. Given an aggregate function, a comparator reads pairs of infinite weight sequences synchronously, and compares their aggregate values in an online manner. In laying the theoretical foundations of comparators, we establish the criteria under which aggregation functions permit comparators represented by finite-state machines, specifically Büchi automata. The crux of these finite-state representations is that they reduce the numerical problem of comparison between aggregation of weight sequences to a problem over finitely many states, inspiring logic-based algorithms that avoid numerical methods and can offer formal guarantees of correctness.

Synthesis from satisficing goals w/o temporal goals. We reduce synthesis from satisficing goals to synthesis from automata-based specification using the comparator for discounted-sum. The key insight is that the satisficing condition is to ensure the discounted-sum rewards exceeds a threshed value, which is captured by an appropriate finite-state comparator for discounted-sum. Synthesis from temporal specifications are also performed via reduction to synthesis from automata-based specification. Since automata are closed under conjunction, we obtain the first sound algorithm for synthesis from temporal and satisficing goals [7].

We render this theoretically elegant solution practically viable through the most compact (minimal and deterministic) construction of the comparator for discounted-sum aggregation function [14]. Consequently, synthesis from satisficing goals using our logic-based method is more efficient by an order of magnitude than prior numerical-method based solutions. Furthermore, using these compact constructions leads to an exponential improvement in synthesis from temporal and satisficing goals than our prior non-deterministic constructions [9].

Our algorithm has been evaluated on challenging benchmarks from robotics planning in adversarial environments and manipulation domains for human-robot interaction [10]. Empirically, our algorithm is not only more efficient, its performance is more robust and resilient than numerical methods. Finally, the soundness guarantee is particularly valuable as the benchmarks are so complex that they preclude an analysis of correctness.

Other applications. Our logic-based framework results in theoretical and practical improvements in a myriad of applications involving quantitative reasoning. Using comparators, we resolve a 15 year old open problem on the complexity of the *discounted-sum inclusion problem*, which is a fundamental problem in verification of quantitative properties [9]. Using comparators, we reduce quantitative inclusion to qualitative (language) inclusion, which is the cornerstone problem of model-checking. The associated implementation QuIP [8] has been shown to be more space- and time-efficient than prior approaches using numerical methods. It is even more efficient with the compact construction of comparator it avoids an inefficient determinization steps. We have used QuIP in verification of in multi-agent games with rational players such as online auction systems, repeated games, and cryptocurrency protocols. Comparators with discounted-sum (integer discount factors) can also be used for infinite-precision arithmetic and in synthesis in multi-agent systems with rational-agents.

Future Vision and Outlook

My long-term research vision is to make AI trustworthy, safe, and reliable for applications in safety- and security-critical domains such as robotics and automation, medical diagnostics, finance etc. Towards this goal, I will continue to bridge the gap between FM/PL and AI/ML to amalgamate their complementary strengths of well-understood guarantees and versatile-cum-practical algorithms, respectively. Within this theme, my cross-disciplinary research hitherto has made noteworthy developments towards assured autonomy through principled development of practical algorithms with assurance guarantees, systematically alleviating the primary hurdles of AI. In addition to extending my work on assured autonomy for real-world impact, moving ahead I will work towards broadening the spectrum of formal reasoning to other emerging technologies in AI and novel notions of formal guarantees. Addressing these challenges will combine expertise in foundational development and contributions from practitioners. I will continue to seek academic and industrial collaborations. I will pursue this vision through (a) advances in automated synthesis using existing and emerging AI technologies, (b).

advances in formal verification and certification in AI, and (c). advances in new notions of formal guarantees, such as interpretability and generalizability. Few concrete themes have been described below.

Synthesis of Cyber-Physical Systems. My work on control synthesis is evidence that formal methods has a significant role in safe and trustworthy computing. In my work so far, a crucial assumption is that the high-level task is *fully specified*. In the next steps, I will investigate synthesis from partial forms of specifications, as it may be challenging to write a complete formal specification in real-world systems. A possible solution could be to synthesize from specifications obtained using supervised or unsupervised learning on (simulations of) system executions. In these cases, synthesis will have to account for noisy and possibly adversarial data in the specification. In another line of work, I will investigate synthesis under partial information and asynchronous models of computations. For example, in training an autonomous car in simulations, say an autonomous F1/10 racing car, in simulations we can assume the input to the controller are the position coordinates of the car. In reality, however, the input may be less precise perceptions in the form of LIDAR signals or images. This requires us to develop verification and synthesis algorithms for NN-based controllers under the assumption of partially available information from the environment. I have already begun work on synthesis under partial information and asynchrony in logic-based approaches [12,13]. I will work towards extending these to the learning setting, by incorporating partial information and asynchrony into the high-level model-based planning. I will seek interdisciplinary collaborations with robotics, controls researchers, and electrical engineers to automate the Sim-to-Real pipeline with as strong assurance guarantees as possible.

Scalable and Trustworthy AI. For real-world impact, trustworthiness will have to be combined with scalability. Recently, computational engines for automated reasoning and symbolic AI have witnessed impressive growth in performance. These engines hold the promise to boost the scalability of trustworthy AI algorithms. In prior work, however, we have shown naive usage as back-end engines could hurt performance since it prevents algorithms from exploiting the strengths of the front-end domain and back-end engine. We show close integration of both, exploring domain-awareness, could drive significant boost in performance. As a first step, in Lisa [2], our tool for reactive synthesis of finite-horizon tasks, we show that domain-aware integration of the logic-based front-end with symbolic AI back-end comprehensively outperforms all state-of-the-art tools on several metrics. Moving ahead, I will continue this line of work for synthesis under complex environments and complex specifications, while also exploring other aspects in trustworthy AI such as verification and security.

Generalization in AI: Looking beyond Trust. Traditionally, formal verification and synthesis have been used to amass trust from AI systems. The wide-scale adoption of AI/ML has given rise to novel problems, such as lack of generalization, non-interpretability, and non-explainability, where formal methods could make foundational contributions. In the future, I will works towards the development of a formal framework to resolve these issues. For example, the lack of generalization is rampant in learning-based techniques for synthesis. In order to generate controls that generalize to environments and specifications other than the ones the control is trained on, I will develop semantics-aware solutions for synthesis. Semantic information could be utilized to identify when environments and specifications are *isomorphic*, and to identify the transferable skills in control tasks. For example, if the control task is to jump over an obstacle, the skill to "jump" is transferable to all locations of the obstacle in the environment. One way to incorporate semantics into algorithms is through re-imagination of abstractions for specifications, such as develop skill-based abstractions as opposed to the traditional goal-based abstraction, or combinations thereof. Additionally, I will continue my collaboration on transferring/adapting an existing implementation to accomplish a new specification with rigorous guarantees of correctness [4]. Such foundational contributions could result in breakthroughs beyond trustworthy AI.

References

1. DiRL. https://github.com/keyshor/dirl.
2. Lisa. https://github.com/vardigroup/lisa.
3. R. Alur, **S. Bansal**, O. Bastani, and K. Jothimurugan. A framework for transforming specifications in reinforcement learning (under review). In *https://arxiv.org/pdf/2111.00272.pdf*, 2022.
4. G. Araman, **S. Bansal**, D. Fried, L. M. Tabajara, M. Y. Vardi, and G. Wiess. Adapting behaviors via reactive synthesis. In *Proc. of CAV*, 2021.
5. K. Jothimurugan, **S. Bansal**, O. Bastani, and R. Alur. Compositional reinforcement learning from logical specifications. In *Proc. of NeurIPS*, 2021.
6. K. Jothimurugan, **S. Bansal**, O. Bastani, and R. Alur. Specification-guided learning of nash equilibria with high social welfare (under review). 2022.
7. **S. Bansal**, K. Chatterjee, and M. Y. Vardi. On satisficing of quantitative games. In *Proc. of TACAS*, 2021.
8. **S. Bansal**, S. Chaudhuri, and M. Y. Vardi. Automata vs linear-programming discounted-sum inclusion. In *Proc. of CAV*, 2018.
9. **S. Bansal**, S. Chaudhuri, and M. Y. Vardi. Comparator automata in quantitative verification. In *Proc. of FoSSaCS*, 2018.
10. **S. Bansal**, L. Kavraki, M. Y. Vardi, and A. Wells. On synthesis from satisficing and temporal goals. In *Proc. of AAAI*, 2022.

11. **S. Bansal**, Y. Li, L.M. Tabajara, and M. Y. Vardi. Hybrid compositional reasoning for reactive synthesis from finite-horizon specifications. In *Proc. of AAAI*, 2020.
12. **S. Bansal**, K. S. Namjoshi, and Y. Sa'ar. Synthesis of asynchronous reactive programs from temporal specifications. In *Proc. of CAV*, 2018.
13. **S. Bansal**, K. S. Namjoshi, and Y. Sa'ar. Synthesis of coordination programs from linear temporal logic. In *Proc. of POPL*, 2020.
14. **S. Bansal** and M. Y. Vardi. Safety and co-safety comparator automata for discounted-sum inclusion. In *Proc. of CAV*, 2019.