Sequential decision-making under uncertainty is a fundamental problem in artificial intelligence. This problem is faced by autonomous agents (e.g., physical systems, including robots, softbots, and automated manufacturing systems) embedded in complex environments that need to choose actions to achieve long-term goals efficiently. It is also a core problem in many biological systems, examples of which include navigation, medical diagnosis, driving, and general everyday activities, from going to get lunch, to going to graduate school.

For example, in a medical diagnosis task the physician (agent) cannot directly observe the internal state (hidden state) of the patient (environment) but rather perceives the symptoms (observations). Yet, after observing every symptom he has to issue a treatment or test (action), which in turn produces a new symptom. Treatments can produce side effects because they can change the internal state of the patient in an unexpected manner. The goal of the physician is to come up with a treatment plan that will “heal” the patient (i.e., cause a transition in the patient’s “state” from being “ill” to being “healthy”). The need to decide what is the best treatment that should be issued now is an instance of the problem of sequential decision-making under uncertainty.

My research objectives concern the development of computational models of learning and sequential decision-making under uncertainty, which are essential traits of any intelligent or rational agent. Computational rationality can be achieved by modeling an agent and its interaction with its environment through actions, perceptions, and rewards. Rational agents choose actions after every perception, such that their long-term reward is maximized. A widely adopted paradigm for modeling this interaction is the partially observable Markov decision process (POMDP) framework. In practice, however, the uncertain outcomes of the agent’s actions and the fact that the true world state may not be completely observable, make learning of POMDP models difficult and using them for sequential decision-making algorithmically infeasible. Scaling learning and planning to large sequential decision-making tasks has been the primary focus of my research.

In my Ph.D. dissertation, I proposed and investigated a multiple resolution hierarchical partially observable Markov decision process (H-POMDP) framework to scale learning and planning in partially observable environments. I applied the new framework in real, large-scale robot navigation and observed various advantages over traditional flat approaches. In scaling planning, the key insights were spatial and temporal abstraction. In spatial abstraction, grouping of lower level states into abstract higher level states (locations in the environment are grouped into corridors and junctions) reduces the uncertainty of the agent. In temporal abstraction, the agent uses longer-duration macro actions instead of just primitive one-step actions. Macro-actions are crucial, in that they produce long traces of experience, which help the agent to disambiguate perceptually aliased situations. In scaling learning, hierarchy has advantages in ease of maintenance, interpretability, and reusability. In particular, the hierarchical nature of the model allows us to train submodels separately and then combine them into an overall hierarchical model (a technique we call “reuse”) which fits the data better and faster than uniform resolution POMDPs. In addition, the hierarchical model allows us to view the environment in terms of junctions and corridors, and thus enables us to infer the spatial structure of the environment. There are two reasons for this: first at higher
levels of abstractions there are fewer spatial relations to consider; second, higher levels of abstraction capture relationships between locations in the environment which are spatially distant.

In my current research, I am continuing to study computational models for learning and sequential decision-making under uncertainty. In particular, I am exploring methods of learning the initial structure of the H-POMDP models, alternative representations of H-POMDPs such as Dynamic Bayesian networks (DBNs), and developing a more rigorous theoretical framework for hierarchical planning in POMDPs. I believe that in order to be able to solve real-world sequential decision problems we need models that incorporate the inherent uncertainty of the real world. Furthermore, to scale up to applications with large state spaces we need structured multi-level representations that allow agents to learn and use them efficiently.

It is also my conviction that many real world problems which do not enjoy efficient solutions yet, can be casted as sequential decision-making tasks under uncertainty. One such example is the problem of computer vision. A popular approach for tackling this problem is classical supervised machine learning. Unfortunately, a major obstacle to this approach is feature selection, or dimensionality reduction. For instance, in an image classification problem we usually first reduce the feature space and then do the classification in one-shot decision-making. However, we could think of image classification as a sequential decision-making task where the class of the image depends on a sequence of observations and actions (where to look at next in the image or which set of features to use next). Learning what actions to take would specify the relevant features required. The feature space can be reduced even more if our actions are long-term macro actions rather than small one-step primitive actions. Other examples can be found in domains such as bioinformatics (e.g., recognition of gene sequences), biometrics (e.g., finger print matching), web data mining (e.g., document classification), and general domains that involve, diagnosis, process monitoring, informational retrieval, and recognition of sequential data (e.g., speech, text).

Given the plethora of real world problems that can be casted as sequential decision-making tasks under uncertainty, I intend to continue exploring and inventing methods for scaling up sequential decision-making models. I would like to apply these models to rich domains, where the state space is exponential in the number of variables, the state variables are partially observable, the agent needs to achieve multiple goals and when the agent may need to interact with other agents. To build an agent that can successfully learn and plan in such rich environments means to have solved a large piece of artificial intelligence.