Eric Brown and *Watson* take on Jeopardy! challenge

Question Answering (QA) has been an active area of research for several decades. Instead of retrieving whole web pages in response to keyword queries, as is typical for web search engines, a QA system retrieves answers to questions. Eric Brown (Ph.D. '96) has been involved in QA since 1999, when he developed a custom search engine at the IBM T. J. Watson Research Center for one of the first QA research projects at IBM. Leveraging his experience as a student of Distinguished Professor Bruce Croft and a member of the Center for Intelligent Information Retrieval, Brown implemented a semantic search engine to support an approach called Predictive Annotation, where named entities are indexed according to their semantic type, then retrieved as candidate answers if they match the answer type detected in the question.

After conducting research in this area for several years with a relatively small team, IBM Research identified QA as an important technology of the future, with the potential for transforming how humans interact with computers. To highlight this technology and accelerate its advancement, IBM embarked on the Jeopardy! Grand Challenge, where the goal is to build a QA system that can play the popular television quiz show Jeopardy! and beat a human champion. In late 2006, Brown joined a dozen other researchers at the Watson Research Center and, under the lead of principal investigator David Ferrucci, took on this challenge in earnest.

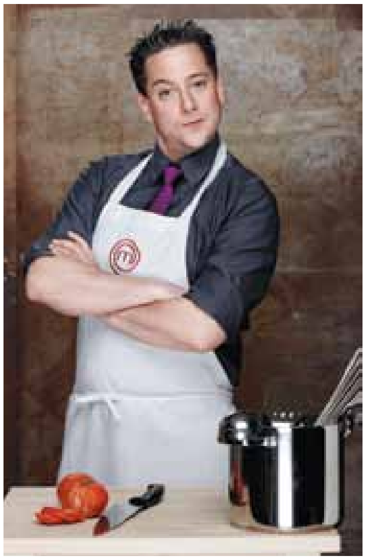
At first glance, building a system that can play Jeopardy! may seem like a strange project for IBM, says Brown, who is currently an IBM Research Staff Member and Manager, Unstructured Information Management Systems. Closer inspection, however, reveals that playing Jeopardy! requires solving several hard problems that make it an ideal benchmark for advancing QA technology. First, Jeopardy! covers a huge variety of topics, including history, geography, science, sports, the arts, entertainment, current events, and more. Nearly any topic you can imagine is fair game. Second, the questions, or "clues" as they're called in Jeopardy!, are expressed using complex natural language, and may include tricky phrasing, puns, or even puzzles. Third, to answer a Jeopardy! clue the player must respond with a precise answer – not a ranked list of web pages, but rather the precise word or phrase that answers the clue. Moreover, the player must determine a confidence in their answer and decide whether or not to ring in and even attempt to answer the clue. The reason for this is simple; if the player answers correctly, the value of the clue is added to the player's score, but if the player answers incorrectly, the value of the clue is subtracted from the player's score. And to top it all off, speed is essential. A player typically has just three or four seconds to come up with his/her answer and decide if he/she is confident enough to ring in and attempt the clue.

"When the idea of building a computer system to play Jeopardy! was first proposed, many people thought it was impossible," says Brown. "But we knew that if we were successful, we would accomplish something truly amazing." To tackle this challenge, the IBM team had to revisit the state-of-the-art in QA and develop new approaches to solving the problem. The core architecture that emerged, called DeepQA, is built on a fundamental philosophy of combining a large number of different natural language processing analytics that operate on existing unstructured information (e.g., documents) to generate candidate answers and evaluate evidence. Underpinning this approach is a probabilistic machine learning framework that has enabled the team (now nearly two dozen researchers) to independently develop a large number of analytics for analyzing a wide variety of information and evidence sources, and then combine those analytics and their results into a single, integrated system. The application of DeepQA that actually plays Jeopardy! is a system called *Watson* (in honor of IBM's founder Thomas J. Watson, see **www.ibm.com/watson** for more).

Brown's role on the project has spanned a number of areas, including responsibility for the initial architecture and data model, specifying and acquiring hardware for the team's development environment, leading the systems architecture and engineering team, leading several "task force" teams to address specific classes of questions, and coordinating many of the team's academic collaborations and relationships. In particular, the IBM team has collaborated with Professor James Allan at UMass Amherst to leverage Indri as one of the search engines used in DeepQA.

Addressing all of the aspects that make Jeopardy! so hard (and so compelling) will have clear and direct implications for business applications of the technology. "We view this technology as going well beyond QA and into the realm of information systems that support intelligent decision making," says Brown. There are a number of application areas that could benefit from this technology, including medical diagnosis, business intelligence, customer support call centers, compliance, legal research, etc. "We've made tremendous progress in a few short years to build a competitive Jeopardy! system. Effectively adapting and applying this technology to even just a few of the many relevant business applications should keep us busy for many more years."

David Miller, MasterChef finalist



We had a chance to talk with Dave Miller (B.S. '03, M.S. '06) after he returned from his successful stint on the first season of Fox TV's MasterChef competition. Chosen from thousands who auditioned, Miller was one of the top 14 contestants on the show and was named as the runner-up in the final competition.

"Aside from getting married in Las Vegas with a bridal party donned in tuxedo t-shirts, the MasterChef experience has been the greatest of my life. Yes, 'I'm a software engineer – it's a fantastic career and I love it, but to toot my own horn: I'm a hell of a chef, too," says Miller. "To have the opportunity to take your biggest passion to the world stage for all to admire and poke fun at with no consequences to anything other than pride, well, at least it was my dream come true. In no small part, my extended stay at UMass Amherst (truthfully both in and out of the Computer Science building) shaped me into the questionably-entertaining, overconfident, eyebrow raising gourmand you know today."

Miller adds, "Now, not every software engineer wants to try out for MasterChef or be on TV at all for that matter, but it does hearken the age-old question: 'What is it, you say, you do here?' Here's something from MasterChef that I wish had made it past the cutting room floor:

Gordon Ramsay (MasterChef host): 'David Miller, what is it again that you do for a living?'

Miller: 'I'm a software engineer.'

Ramsay: 'Really? I could have sworn you were a professional face-maker'."

Miller works at Vistaprint in Lexington, MA: a predominately .NET shop, where he develops internal CRM applications to support their three international customer care centers. He has a flexible schedule, so he can be home to work in the kitchen and make dinner before his wife arrives home. "I'm sure if I worked in a restaurant all day, I'd come home and write iPhone apps for fun," says Miller. "Fortunately, things are the other way around - I can pick up a Chateaubriand and some chanterelles after work, but I don't code for fun anymore. Does that mean I've given up my passion? In all honesty, I hope not. Some people cook to live, but I live to cook. If I did it for a living, I don't know if I'd love it like I do now."

Miller adds, "Often people approach me skeptical that someone with such a technical mind can do what I do in the kitchen, and, to the surprise of many, I tell them that the two really aren't that different. During my graduate research with LASER in process modeling and simulation, I adopted the concept introduced to me by Professor Lee Osterweil that (pardon my gross oversimplification) software is process and process is software; essentially any formalized process, too, is itself software - something that has stuck with me to this day. Ironically, Lee most frequently referenced recipes as a real-world example of a formal process definition. Combine the mind of a software engineer with that of a passionate chef, and you've got one extremely process-driven individual. Thus, I don't find it that surprising that concepts like resource allocation, scale, exception handling, performance, and optimization apply in the kitchen – I'd even argue that a little creativity and flair are necessary in software as well."

If there's anything that computer scientists in general can take from his adventures on MasterChef, Miller hopes it would inspire confidence in the fact that talented people really can excel in more than one talent - no matter how seemingly dissimilar they may be. "In your own lives, please don't fall subject to the pigeonhole principle - make more holes and fill them with something tasty," says Miller.

Alum Connections

This fall, **Jody Daniels**, CS alum (Ph.D. '97) and 2010 OAA award recipient, was nominated for a promotion to the rank of brigadier general. According to U.S. DoD news release, Daniels will be assigned "as commander, (troop program unit), U.S. Army Reserve Support Command, First Army/deputy commanding general, First Army (East), Fort Meade, Md." Daniels is also the Director of Advanced Programs at Lockheed Martin.

At the ACM/IEEE 32nd International Conference on Software Engineering (ICSE 2010), **Jay Corbett** (Ph.D. '92), **Matthew Dwyer** (Ph.D. '95), and co-authors received the ISCE Most Influential Paper Award for their paper "Bandera: Extracting Finite-State Models from Java Source Code." The award is given to the author(s) of the paper from the ISCE conference 10 years ago that is judged to have had the most influence on the theory or practice of software engineering during the decade since its original publication. Corbett is currently a Senior Software Engineer at Google and Dwyer is currently the Henson Professor of Software Engineering in the Department of Computer Science and Engineering at the University of Nebraska.

The Massachusetts Innovation & Technology Exchange (MITX), New England's premier association for Internet business and marketing, named **Steve Vinter** (Ph.D. '85) as one of the five new members to its Board of Directors. Vinter is the Director of Google's Cambridge, MA facility.

Intronis Online Backup and Recovery appointed **Jay Bolgatz** (B.S. '85) as it Vice President of Engineering and Delivery.

Massachusetts-based GenomeQuest, Inc. announced this fall that **Richard Resnick** (B.S. '94) was appointed acting CEO of the company.

Dr. **Tom Wagner** (Ph.D. '00) was recently appointed to the position of Senior Vice President and Chief Technological Officer at iRobot.

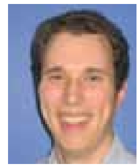
Nilanjan Banerjee: *Improved Network Consistency and Connectivity in Mobile and Sensor Systems*; (Mark Corner, Advisor); Sept. 2009; Assistant Professor, Department of Computer Science and Computer Engineering, University of Arkansas, Fayetteville.

Edge networks such as sensor, mobile, and disruption tolerant networks suffer from topological uncertainty and disconnections due to a myriad of factors including mobility and limited battery capacity on client devices. Hence, providing reliable, always-on consistency for network applications in such mobile and sensor systems is non-trivial and challenging. However, the problem is of paramount importance given the proliferation of mobile phones, PDAs, laptops, and music players. This thesis identifies two fundamental deterrents to addressing the above problem. First, limited energy on client mobile and sensor devices makes high levels of consistency and availability impossible. Second, unreliable support from the network infrastructure, such as coverage holes in WiFi degrades network performance. We address these two issues through client- and infrastructure-end modifications. The first part of this thesis proposes a novel energy management architecture called Hierarchical Power Management (HPM). HPM combines platforms with diverse energy needs and capabilities into a single integrated system to provide high levels of consistency and availability at minimal energy consumption. We present two systems, Triage and Turducken, which are instantiations of HPM for sensor net microservers and laptops, respectively. The second part of the thesis proposes and analyzes the use of additional infrastructure in the form of relays, mesh nodes, and base stations to enhance sparse and dense mobile networks. We present the design, implementation, and deployment of Throwboxes—a relay system to enhance sparse mobile networks and an associated system for enhancing WiFi based mobile networks.

Patrick Deegan: *Whole-Body Strategies for Mobility and Manipulation*; (Roderic Grupen, Advisor); May 2010; Senior Robotics Engineer, Heartland Robotics.

The robotics community has succeeded in creating remarkable machines and task-level programming tools, but arguably has failed to apply sophisticated autonomous machines to sophisticated tasks. The dissertation introduces the uBot-5—a mobile manipulator concept to support new robotic applications in our culture that require fully integrated dexterous robots in unstructured environments. The integrated system provides dexterous modes for mobility and manipulation and control firmware that organizes these behavioral modes logically for use by application code.

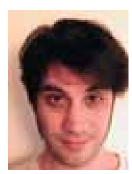
The approach chosen in this study centers around a hardware and software co-development. The platform successfully pairs motor flexibility and performance with a hierarchical embedded control framework for constructing dexterous machines. In particular, postural control underlies the uniform treatment of several mobility modes that engage different combinations of sensor and motor resources. The result is a platform for studying "whole-body" control strategies that can be applied jointly to simultaneous mobility and manipulation objectives. Furthermore, dexterous machines can express the "aptitudes" implicit in the design of the robot in the embedded firmware and hierarchically organize the behavior of the system for programming. This is a win-win situation where the quality of the embedded firmware determines how efficiently programmers (autonomous learning algorithms or human programmers) can construct control programs that are robust, flexible, and respond gracefully to unanticipated circumstances.



Andrew Fast: *Learning the Structure of Bayesian Networks with Constraint Satisfaction*; (David Jensen, Advisor); Feb. 2010; Research Scientist, Elder Research Inc.

A Bayesian network is a graphical representation of the probabilistic relationships among a set of variables and can be used to encode expert knowledge about uncertain domains. The structure of this model represents the set of conditional independencies among the variables in the data. In this thesis, I focus on learning the structure of Bayesian networks from data with constraint-based algorithms. These algorithms use a series of conditional hypothesis tests to learn independence constraints on the structure of the model.

I show that new algorithms inspired by constraint satisfaction are able to produce significant improvements in structural accuracy. These constraint satisfaction algorithms exploit the interaction among the constraints to reduce error. First, I introduce an algorithm based on constraint optimization that is sound in the sample limit, like existing algorithms, but is guaranteed to produce a DAG. This new algorithm learns models with structural accuracy equivalent or better to existing algorithms. Second, I introduce an algorithm based constraint relaxation. Constraint relaxation combines different statistical techniques to identify constraints that are likely to be incorrect, and remove those constraints from consideration. I show that an algorithm combining constraint relaxation with constraint optimization produces Bayesian networks with significantly better structural accuracy when compared to existing structure learning algorithms, demonstrating the effectiveness of constraint satisfaction approaches for learning accurate structure of Bayesian networks.



Stephen Hart: *The Development of Hierarchical Knowledge in Robot Systems*; (Roderic Grupen, Advisor); Sept. 2009; Postdoctoral Researcher, Italian Institute of Technology, Genova, Italy.

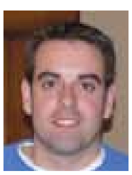
I investigate two complementary ideas in the literature on machine learning and robotics—those of embodiment and intrinsic motivation—to address a unified framework for skill learning and knowledge acquisition. "Embodied" systems make use of structure derived directly from sensory and motor configurations for learning behavior. Intrinsically motivated systems learn by searching for native, hedonic value through interaction with the world. Psychological theories of intrinsic motivation suggest that there exist internal drives favoring open-ended cognitive development and exploration. I argue that intrinsically motivated, embodied systems can learn generalizable skills, acquire control knowledge, and form an epistemological understanding of the world in terms of behavioral affordances.

I propose that the development of behavior results from the assembly of an agent's sensory and motor resources into state and action spaces that can be explored autonomously. I introduce an intrinsic reward function that can lead to the open-ended learning of hierarchical behavior. This behavior is factored into declarative "recipes" for patterned activity and common sense procedural strategies for implementing them in a variety of run-time contexts. These skills form a categorical basis for the robot to interpret and model its world in terms of the behavior it affords. Experiments conducted on a bimanual robot illustrate a progression of cumulative manipulation behavior addressing manual and visual skills. Such accumulation of skill over the long-term by a single robot is a novel contribution that has yet to be demonstrated in the literature.

Manjunatha Jagalur; *Discovery of Complex Regulatory Modules from Expression Genetics Data*; (David Kulp, Advisor); May 2010; Bioinformatics Scientist, Pacific Biosciences.

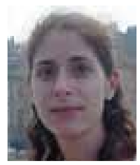
Mapping of strongly inherited classical traits has been immensely helpful in understanding many important traits including diseases, yield and immunity. But some of these traits are too complex and are difficult to map. Taking into consideration gene expression, which mediates the genetic effects, can be helpful in understanding such traits. Together with genetic variation data such a dataset is collectively known as expression genetics data. Presence of discrete and continuous variables, observed and latent variables, availability of partial causal information, and under-specified nature of the data make expression genetics data computationally challenging, but potentially of great biological importance. In this dissertation the underlying regulatory processes are modeled as Bayesian networks consisting of gene expression and genetic variation nodes. Due to the under-specified nature of the data, inferring the complete regulatory network is impractical. Instead, the following techniques are proposed to extract interesting subnetworks with high confidence.

The network motif searching technique is used to recover instances of a known regulatory mechanism. The local network inference technique is used to identify immediate neighbors of a given transcript. Application of these two techniques often results in identification of hundreds of individual networks. The network aggregation technique extracts the most common subnetwork from those networks, and identifies its immediate neighbors by collapsing them into a common network. In all the above tasks, simulation studies were carried out to estimate the robustness of the proposed methods and the results suggest that these techniques are capable of recovering the correct substructure with high precision and moderate recall. Moreover, manual biological review shows that the recovered regulatory network substructures are typically biologically sensible.



Jeffrey Johns; *Basis Construction and Utilization for Markov Decision Processes using Graphs*; (Sridhar Mahadevan, Advisor); Feb. 2010; Computing Innovation Fellow, Department of Computer Science, Duke University.

In reinforcement learning (RL), an agent takes actions in an environment and receives rewards. The agent must use its experience in order to learn how best to act in the future. One of the main challenges for an autonomous agent is in representing functions/features over very large and complex environments. The majority of successful, large-scale RL applications have required humans to provide such features to the agent; however, recent research suggests this process of feature construction can be automated and solved by the agent itself. Building on this idea, we propose two algorithms for scaling automatic feature construction to very large data sets. Once the features are computed, the agent must utilize those features to learn how best to behave. We introduce a new least-squares algorithm that allows for the agent to make efficient use of its experience in the environment. Furthermore, we evaluate feature selection methods that tailor the features to the agent's desired task. These feature selection methods encourage sparse solutions and provide regularization, both properties that are necessary when dealing with complex environments.



Victoria Manfredi; *Sensor Control and Scheduling Strategies for Sensor Networks*; (James F. Kurose, Advisor); Sept. 2009; Computing Innovation Fellow, Department of Computer Science, Boston University.

We investigate sensor control and scheduling strategies to most effectively use the limited resources of an ad hoc network or closed-loop sensor network. We first consider where to focus sensing in a meteorological radar network. We show that the main benefits of optimizing sensing over expected future states are when there are multiple small phenomena in the environment. We next investigate how to make sensing robust to delayed and dropped packets. We ground our analysis in a meteorological radar network and show that prioritizing sensor control traffic decreases the round-trip control-loop delay, and thus increases the quantity and quality of the collected data and improves application performance. Finally, we examine how to make routing robust to network changes. We propose a routing algorithm that selects a type of routing subgraph (a braid) that is robust to changes in the network topology. We analytically characterize the reliability of a class of braids and their optimality properties, and give counter-examples to other conjectured optimality properties in a well-structured (grid) network. Comparing with dynamic source routing, we show that braid routing can significantly decrease control overhead while only minimally degrading the number of packets delivered, with gains dependent on node density.

Sarah Osentoski; *Action-Based Representation Discovery in Markov Decision Processes*; (Sridhar Mahadevan, Advisor); Sept. 2009; Postdoctoral Researcher, Computer Science Department, Brown University.

This dissertation investigates the problem of representation discovery in discrete Markov decision processes, namely how agents can simultaneously learn representation and optimal control. Previous work on function approximation techniques for MDPs largely employed hand-engineered basis functions. We explore approaches to automatically construct these basis functions and demonstrate that automatically constructed basis functions significantly outperform more traditional, hand-engineered approaches.

We specifically examine two problems: how to automatically build representations for action-value functions by explicitly incorporating actions into a representation, and how representations can be automatically constructed by exploiting a pre-specified task hierarchy. We first introduce a technique for learning basis functions directly in state-action space. The approach constructs basis functions using spectral analysis of a state-action graph which captures the underlying structure of the state-action space of the MDP. We show how our approach can be used to approximate state-action value functions when the agent has access to macro-actions: actions that take more than one time step and have predefined policies. We describe how state-action graphs can be modified to incorporate information about the macro-actions. Finally, we describe how hierarchical reinforcement learning can be used to scale up automatic basis function construction. We extend automatic basis function construction techniques to multi-level task hierarchies and describe how basis function construction can exploit the value function decomposition given by a fixed task hierarchy. We demonstrate that combining task hierarchies with automatic basis function construction allows basis function techniques to scale to larger problems and leads to a significant speed-up in learning.

Shichao Ou; *A Behavioral Approach to Human-Robot Communication*; (Roderic Grupen, Advisor); Feb. 2010; Senior Software Engineer, Network Equipment Technologies.

This dissertation focuses on how a robot can acquire and refine expressive and receptive communication skills with human beings. I hypothesize that communication has its roots in motor behavior and present an approach that is unique in the following aspects: (1) representations of humans and the skills for interacting with them are learned in the same way as the robot learns to interact with other "objects," (2) expressive behavior naturally emerges as the result of the robot discovering new utility in existing manual behavior in a social context, and (3) symmetry in communicative behavior can be exploited to bootstrap the learning of receptive behavior.

Experimental results show that the robot successfully acquired a variety of expressive pointing gestures, and the perceptual skills with which to recognize and respond to similar gestures from humans. This illustrates the validity of the approach as a computational framework for learning increasingly comprehensive models and behavior for communicating with humans. Also, due to variations in human reactions over the training subjects, the robot developed a preference for certain gestures over others, showing that the approach can adapt to different human behavior. These results support the experimental hypotheses and offer insights for future studies.

M.S. Raunak; *Resource Management In Complex And Dynamic Environments*; (Leon J. Osterweil, Advisor); Sept. 2009; Visiting Assistant Professor, Department of Computer Science, Loyola College.

Resource management is at the heart of many diverse science and engineering areas. Often a relatively simple model of resources can suffice for work in a number of domains. The problems of resource specification and management become much more challenging, however, when working with a complex real-life domain, such as the emergency department of a hospital, with many heterogeneous resource types and intricate constraints on their utilization. This dissertation proposes an approach for modeling and managing resources in complex and dynamic environments, and presents an architecture that focuses on appropriate separation of concerns. To evaluate this approach we developed ROMEO, an implementation of the general approach proposed in the dissertation. ROMEO supports execution and simulation of complex real-world processes. We have studied the effectiveness of ROMEO's well modularized separation of concerns by examining how well ROMEO supports execution and simulation of a wide variety of different real-world processes such as hospital emergency department processes, online dispute resolution processes, and web services development processes. Our studies suggest that our choices of concerns to separate offer some important advantages, such as ease of modification, and the ability to represent important fine-scale details.

Bruno Ribeiro; *On the Design of Methods to Estimate Network Characteristics*; (Donald F. Towsley, Advisor); May 2010; Postdoctoral Research Associate, Department of Computer Science, University of Massachusetts Amherst.

Social and computer networks permeate our lives. Large networks, such as the Internet, the World Wide Web, and wireless smartphones, have indisputable economic and social importance. These networks have non-trivial topological features, i.e., features that do not occur in simple networks such as lattices or random networks. Estimating characteristics of these networks from incomplete (sampled) data is a challenging task.

This thesis provides two frameworks within which common measurement tasks are analyzed and new, principled, measurement methods are designed. The first framework focuses on sampling directly observable network characteristics. This framework is applied to design a novel multidimensional random walk to efficiently sample loosely connected networks. The second framework focuses on the design of measurement methods to estimate indirectly observable network characteristics. This framework is applied to design two new, principled, estimators of flow size distributions over Internet routers using (1) randomly sampled IP packets and (2) a data stream algorithm.

Timothy Richards; *Generalized Instruction Selector Generation: The Automatic Construction of Instruction Selectors from Descriptions of Compiler Internal Forms and Target Machines*; (J. Eliot B. Moss, Advisor); Feb. 2010; Visiting Assistant Professor, Department of Computer Science, Trinity College.

One of the most difficult tasks a compiler writer faces is the construction of the instruction selector (IS). The IS is the part of the compiler that translates compiler intermediate representation (IR) into instructions for a target machine. Unfortunately, implementing an IS by hand is difficult, time consuming, and error prone. The details of the IR and target instruction set is carefully considered in order to generate correct and efficient code. This requires an expert in compiler internals as well as the target machine. In this dissertation we describe the instruction selector problem, cover previous attempts at solving it, and identify what we believe to be the most prominent factor inhibiting their widespread adoption.

This dissertation proposes a generalized approach toward generating instruction selectors automatically. We propose CISL, a common machine description language for specifying compiler IR and target instruction semantics, and GIST, a heuristic search procedure that discovers equivalent instruction sequences between compiler IR and target instructions. GIST leverages CISLs well-defined semantics to discover IS patterns automatically. Adapter programs use GIST-generated selector patterns to output compiler specific implementation code. Our experiments show that IS patterns can be discovered automatically and independent of a particular compiler framework or target machine.

Alicia P. Wolfe; *Paying Attention to What Matters: Abstraction in Partially Observable Domains*; (Andrew G. Barto, Advisor); Feb. 2010.

Autonomous agents may not have access to complete information about the state of the environment. For example, a robot soccer player may only be able to estimate the locations of other players outside the scope of its sensors. However, even though all the information needed for ideal decision making cannot be sensed, all that is sensed is usually not needed. The noise and motion of spectators, for example, can be ignored in order to focus on the game field. Standard formulations do not consider this situation, assuming that all the can be sensed must be included in any useful abstraction. This dissertation extends the Markov Decision Process Homomorphism framework to partially observable domains, focusing specifically on reducing Partially Observable Markov Decision Processes (POMDPs) when the model is known. This involves ignoring aspects of the observation function that are irrelevant to a particular task. Abstraction is particularly important in partially observable domains, as it enables the formation of a smaller domain model and thus more efficient use of the observed features.