Research Statement

Thomas J. Walsh

My core research in Artificial Intelligence and Machine Learning attempts to bridge the gap between classical hand-designed AI representations that scale well in many domains, and modern machine learning where models are built from a limited amount of data. The bulk of this work has been within the *reinforcement learning* (RL) paradigm, where an agent must actively explore its environment and use its experience to determine a policy for maximizing its utility or completing a task. The foundations of my research in this area can be summarized by the following properties of agents:

- Agents should **build compact general models** of their environment based on data. For instance, a soccer-playing robot might learn the outcome of kicking a ball in one part of the field, and apply that model in planning how to kick the ball into the goal. This stands in contrast to many traditional representations used in reinforcement learning, where these two situations would need to be learned independently.
- Learning should be **sample efficient**, that is there should be a practical theoretical bound on the amount of time an agent spends *exploring* its environment.
- When possible, agents should **make make use of teachers**, either humans or other agents, to learn their tasks faster. This channel of experience often goes unused in many domains, but my research has shown it to be of immense practical and theoretical importance.

In the next phase of my research career, I want to work in larger and more complex real-world environments and applications, while still applying the basic tenets listed above. Below I give a general outline of my research up to this point and a discussion of my future goals.

In reinforcement learning, an agent observes aspects of its environment and can take actions to change its state and receive rewards, with the long term goal of maximizing some function of these reward terms (e.g. their discounted sum), perhaps by completing a task. Within RL, my research has focussed on algorithms that instantiate the properties I mentioned above: compact models, sample-efficient learning, and integrating teachers. The compact representations I have focused the most on are relational operators [3, 5], such as stochastic versions of STRIPS rules from classical AI. Such operators describe the possible conditions and effects for an action such as $pickup(X)$, which might have conditions based on the weight and shape of X that determine the probability of the agent holding X after executing the action. This is in contrast to common propositional RL models, where the conditions for picking up every object in the world might need to be learned independently.

In my work, agents must actively explore their environment and learn the dynamics of such operators (the conditions and probability distributions of effects) in a *sample efficient* manner. Sample efficiency in RL (see [1]) is a theoretical measure that bounds the number of steps (amount of real world data) where an agent is not acting optimally (while it is exploring the world). This can be seen as an analogue to the standard computer science analysis of computational complexity, but with a focus on limiting real-world steps, rather than computational operations. Traditionally, agents with such theoretical guarantees have learned only propositional models (e.g. MDPs, DBNs) of their environment, but my research has shown that relational operators, like $pickup(X)$, can be learned with similar guarantees.

My dissertation research [2] showed that the effects of operators like *pickup*(X), as well as the conditions and probability distributions for these effects (perhaps a block made of foam has a higher likelihood of actually being picked up than one made of plastic) can all be learned efficiently by autonomous agents. The algorithms and theoretical bounds for learning such operators are non-trivial, as an agent has to reason about multiple hypotheses for the possible conditions causing different effect distributions, and has to learn the effect distributions themselves from ambiguous data. For instance, suppose an agent executes *pickup*(X) on a oddly shaped heavy block and the action fails. The agent has to consider the possibility that the shape, the weight, or a combination of these two properties, influenced the outcome, and has to try to classify the outcome with others it has previously seen. Also, the agent needs to make non-trivial choices between exploring new action conditions versus exploiting the knowledge the agent already has. The algorithms I design to learn under these conditions use contemporary machine learning techniques, (for instance architectures for modeling multiple possible conditional probabilities while learning, and the use of "optimism in the face of uncertainty") but adapt them for use with relational operators, a classical AI representation. The resulting behavior is often quite natural– agents seek out new configurations of the world (for instance, trying to pick up both heavy and light objects) to try and find conditions that affect action outcomes, but cease once their models allow them to complete their tasks quickly.

To scale these algorithms up to real-world domains, I have shown that adding a human teacher who can demonstrate a task, leads to provably (and empirically verified) faster learning [4]. Agents learning from teachers can avoid harmful or unfruitful trajectories, and are able to bootstrap their models faster than their autonomous counterparts. This has allowed me to apply RL agents in complex tasks, including a "delivery" robot and tasks involving the composition of *web services*. In the latter, agents learned how to interact with web services from Amazon and Google based on traces of people using these services to complete a task (for instance looking up their recent travel in their calendar, and looking up the per-diem information in a spreadsheet, and using that to fill out a reimbursement form). The resulting model showed how the inputs and outputs of services were related (for instance always using the per-diem information from the date range of the travel) and allowed an agent to complete the same tasks with new input data. More recently, I have been making use of demonstrations provided by human teachers to help learn models of verb-phrase meanings (understanding commands like "go around the couch"), simulated UAV dynamics, and models of student development for an intelligent tutoring system. The representations used in these domains resemble the relational operators I described above, but the use of human teachers makes learning these models of complex real-world situations much more practical.

In my future research, I will continue to develop machine-learning algorithms that learn fast and generalize well. In terms of theoretical work, I continue to develop algorithms for exploring compact relation models, such as the relational Bayes Nets we have recently been using as models in or intelligent tutoring system work. On the more practical side, I plan to continue investigating how human teachers, compact representations, and exploration techniques from sequential decision making can lead to practical learning agents for very complex domains. I am particularly interested in pushing my research further into areas such as medical applications, Internet applications, and robotics. In medical applications, the amount of available electronic data (for instance records of hospitalized patients) has skyrocketed in recent years, but researchers are still struggling to build adequate models of diagnosis and care from this information. I believe machine learning research, and specifically the use of compact sequential models, can play a large role in bridging this gap. I also believe agents that learn how to interact with services and other applications on the Internet, such as the web service task learning agents I described earlier, are a promising and under-explored research direction. Such agents have access to a huge amount of potential data, but must explore it intelligently, learn from demonstrations, and represent it with compact models, all of which are core properties of my research. In robotics, the kinds of problems I am interested in are again high level-tasks involving relations between objects, and how they affect dynamics, akin to the relational operator learning I described earlier, but with a grounding in the physical world.

References

[1] Lihong Li, Michael L. Littman, and Thomas J. Walsh. Knows what it knows: A framework for self-aware learning. In *ICML*, 2008.

[2] Thomas J. Walsh. *Efficient Learning of Relational Models for Sequential Decision Making.* PhD thesis, Rutgers University, Piscataway, NJ, USA, 2010.

[3] Thomas J. Walsh and Michael L. Littman. Efficient learning of action schemas and web-service descriptions. In *AAAI*, 2008.

[4] Thomas J. Walsh, Kaushik Subramanian, Michael L. Littman, and Carlos Diuk. Generalizing apprenticeship learning across hypothesis classes. In *ICML*, 2010.

[5] Thomas J. Walsh, Istvn Szita, Carlos Diuk, and Michael L. Littman. Exploring compact reinforcement-learning representations with linear regression. In *UAI*, 2009.