

Phd thesis

Learning Domain Abstractions for Long Lived Robots

Benjamin Saul Rosman

Abstract

Recent trends in robotics have seen more general purpose robots being deployed in unstructured environments for prolonged periods of time. Such robots are expected to adapt to different environmental conditions, and ultimately take on a broader range of responsibilities, the specifications of which may change online after the robot has been deployed. We propose that in order for a robot to be generally capable in an online sense when it encounters a range of unknown tasks, it must have the ability to continually learn from a lifetime of experience. Key to this is the ability to generalise from experiences and form representations which facilitate faster learning of new tasks, as well as the transfer of knowledge between different situations. However, experience cannot be managed naively: one does not want constantly expanding tables of data, but instead continually refined abstractions of the data – much like humans seem to abstract and organise knowledge. If this agent is active in the same, or similar, classes of environments for a prolonged period of time, it is provided with the opportunity to build abstract representations in order to simplify the learning of future tasks. The domain is a common structure underlying large families of tasks, and exploiting this affords the agent the potential to not only minimise relearning from scratch, but over time to build better models of the environment. We propose to learn such regularities from the environment, and extract the commonalities between tasks. This thesis aims to address the major question: what are the domain invariances which should be learnt by a long lived agent which encounters a range of different tasks? This question can be decomposed into three dimensions for learning invariances, based on perception, action and interaction. We present novel algorithms for dealing with each of these three factors. Firstly, how does the agent learn to represent the structure of the world? We focus here on learning inter-object relationships from depth information as a concise representation of the structure of the domain. To this end we introduce contact point networks as a topological abstraction of a scene, and present an algorithm based on support vector machine decision boundaries for extracting these from three dimensional point clouds obtained from the agent's experience of a domain. By reducing the specific geometry of an environment into general skeletons based on contact between different objects, we can autonomously learn predicates describing spatial relationships. Secondly, how does the agent learn to acquire general domain knowledge? While the agent attempts new tasks, it requires a mechanism to control exploration, particularly when it has many courses of action available to it. To this end we draw on the fact that many local behaviours are common to different tasks. Identifying these amounts to learning “common sense” behavioural invariances across multiple tasks. This principle leads to our concept of action priors, which are defined as Dirichlet distributions over the action set of the agent. These are learnt from previous behaviours, and expressed as the prior probability of selecting each action in a state, and are used to guide the learning of novel tasks as an exploration policy within a reinforcement learning framework. Finally, how can the agent react online with sparse information? There are times when an agent is required to respond fast to some interactive setting, when it may have encountered similar tasks

previously. To address this problem, we introduce the notion of types, being a latent class variable describing related problem instances. The agent is required to learn, identify and respond to these different types in online interactive scenarios. We then introduce Bayesian policy reuse as an algorithm that involves maintaining beliefs over the current task instance, updating these from sparse signals, and selecting and instantiating an optimal response from a behaviour library. This thesis therefore makes the following contributions. We provide the first algorithm for autonomously learning spatial relationships between objects from point cloud data. We then provide an algorithm for extracting action priors from a set of policies, and show that considerable gains in speed can be achieved in learning subsequent tasks over learning from scratch, particularly in reducing the initial losses associated with unguided exploration. Additionally, we demonstrate how these action priors allow for safe exploration, feature selection, and a method for analysing and advising other agents' movement through a domain. Finally, we introduce Bayesian policy reuse which allows an agent to quickly draw on a library of policies and instantiate the correct one, enabling rapid online responses to adversarial conditions.