Relations between Reinforcement Learning, Visual Input, Perception and Action

Martijn van Otterlo
Artificial Intelligence, Radboud University (The Netherlands)
e-mail m.vanotterlo@donders.ru.nl

july 2013

Reinforcement learning, states and actions

The majority of reinforcement learning (RL) models is based on fundamental notions of a set of *states* S and a set of *actions* A. Conceptually states are seen as *input* to the system, while actions are *output*. States (fully or partially observable) usually contain the *relevant or measurable properties* of the current state of information needed to select an action. Furthermore, actions are *atomic choices* and *models* (such as (PO)MDPs) couple them to states by representing how states *change* when actions are applied. RL algorithms learn from *traces* of state, action, state, action, ... sequences, in order to optimize action selection for each state (wrt. a reward criterium).

The field of RL has come up with many algorithms fitting this description [8]. Many such methods scale up to large problems using abstraction and generalization, usually over states but sometimes over actions too. Value function approximation, hierarchical decomposition of actions, continuous action generalization, state abstraction and compact probabilistic models of environmental dynamics are all examples of how to deal with large problems (see also [5]). In my view general reinforcement learning amounts to feedback-based, *interactive experimentation* with particular abstraction levels for states, actions and tasks. However, despite all these efforts, and although many computer vision algorithms are based on similar types of abstraction and generalization over (state) features, direct couplings of RL with complex visual input (e.g. raw images) are still rare. RL is typically employed quite separately from computer vision.

Note that not all RL methods are built per se upon models such as *Markov decision processes*. Among several alternatives, so-called *predictive approaches* ([8], chap.13), or the recent *Horde* architecture [4] are less dogmatic about states and actions but emphasize a tight coupling between actions and series of (predictions of) perceptions. They connect to a large body of literature in cognitive (neuro)science on *sensorimotor loops*, *embodied cognition* and related ideas.

Relational representations in action and vision

In roughly the recent decade, RL has been combined with so-called *relational* knowledge representation for states and languages [6, 7]. For many structured domains, it is useful to explicitly capture the relational structure of the domain. For example, in the well-known *blocks world*, blocks can be *on top* of other blocks, they can be *next to* each other and so on. Utilizing a powerful representation language, strategies can be learned (using RL) that can generalize over *objects* and *relations* they may have among them. This may give rise to policies trained in worlds with few objects that work in worlds with many, similar, objects. Similarly, many forms of *decision-theoretic planning*, using abstract or relational version of Bellman equations, can employ such powerful knowledge representation schemes [5].

An interesting development is that also in the computer vision community, people wish to employ relational generalization over visual input, due to advances in (probabilistic) logical learning.

For example, we recently proposed an architecture for visual perception of structured objects [1]. The use of relational knowledge can be effective in the high-level *interpretation* of low-level visual input. For example, interpreting a structure in an image as a *house* means that one needs to find several parts of the image representing a *front door*, *windows* and possibly a *chimney*, all in a particular relational structure. Similar advances can be found in robotics (e.g. [3])

Towards richer, relational couplings of vision and action

I believe that, in addition to state-action models as typically used, more opportunities exist to couple complex states and actions. In particular, the real potential of relational representations is that states can *share* information with actions (e.g. parameters, or more specifically *objects*) such that integration of states (i.e. visual perception) and actions (e.g. behavior) can become more tight, both in terms of representation and learning algorithms. In addition, this would tune in with recent developments in cognitive (neuro)science on the intricate interactions between action and perception in general cognition as well.

For example, we recently extended existing work on *affordance models* to the relational case. Affordances are (in a nutshell) opportunities to act, for example a chair provides the opportunity to "sit on it". Our models allow a robot to reason about interactions between (visual) state features, dynamic aspects of states, and action features, in one probabilistic model [3]. Such models also allow to process visual input to *plan ahead* in order to *imitate* visually demonstrated object manipulation behavior [2].

I believe there are several other possibilities to define novel languages for *interactive experimentation* with relational abstraction levels in the context of both complex visual input and complex behavioral output. This includes new types of interactions – for example dealing with scarce human feedback, new types of experimentation – for example incorporating visual feedback and physical manipulation, and new types of abstraction levels – such as probabilistic programming languages. The goal would be to couple complex visual input to behavior, in interaction, in various ways. I will present first steps towards a more tight integration of relational vision and relational action for interactive [1] learning settings. In addition, new problem domains are needed to reach the full potential of relational abstraction over states and actions. Several new problem domains are available to illustrate exactly that.

References

[1] L. Antanas, M. van Otterlo, J. Oramas, T. Tuytelaars, and L. De Raedt. There are plenty of places like home: Using relational representations in hierarchy for distance-based image understanding. *Neurocomputing*, 2013. in press (`http://dx.doi.org/10.1016/j.neucom.2012.10.037`).

[2] B. Moldovan, P. Moreno, and M. van Otterlo. On the use of probabilistic relational affordance models for sequential manipulation tasks in robotics. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2013.

[3] B. Moldovan, P. Moreno, M. van Otterlo, J. Santos-Victor, and L. De Raedt. Learning relational affordance models for robots in multi-object manipulation tasks. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 4373 –4378, 2012.

[4] R. S. Sutton, J. Modayil, M. Delp, T. Degris, P. M. Pilarski, A. White, and D. Precup. Horde: A scalable real-time architecture for learning knowledge from unsupervised sensorimotor interaction. In *AAMAS*, pages 762–768, 2011.

[5] M. van Otterlo. Intensional dynamic programming: A Rosetta stone for structured dynamic programming. *Journal of Algorithms*, 64:169–191, 2009.

[6] M. van Otterlo. *The Logic of Adaptive Behavior*. IOS Press, Amsterdam, The Netherlands, 2009.

[7] M. van Otterlo. Solving relational and first-order logical Markov decision processes: A survey. In M. Wiering and M. van Otterlo, editors, *Reinforcement Learning: State-of-the-Art*, chapter 8. Springer, 2012.

[8] M. Wiering and M. van Otterlo. *Reinforcement Learning: State-of-the-Art*. Springer, 2012.

[1] See also the IJCAI Workshop on Machine Learning for Interactive Systems (`mlis-workshop.org/2013/`)