Think Aloud Imitation Learning

Laurent Orseau
Insa/Irisa
35000 Rennes, France
lorseau@irisa.fr

1. Abstract

Learning by imitation (Schaal, 1999) is an interesting compromise between pure autonomous learning, which can be extremely long, and programming a robot, which is not easy either, and is too specific: the agent has no autonomy in learning. In imitation, since intermediate steps are provided, the complexity of the task becomes much lower. But no explicit rule is given, and the imitated agent may not be aware of being observed by the imitating one.

Learning sequences of actions by imitation can then take place with usual Machine Learning techniques such as Recurrent Neural Networks (RNN) (Billard and Mataric, 2001, Ogata et al., 2003), which are efficient on-line learning devices. However, there remains an issue for more complex tasks than mere sequences of motor actions: some problems require hidden states inside the "head" of the teacher, that the imitator cannot thus see from the simple observation of the actions of the imitated agent. This is known in RNNs as part of the long-term dependency issue (Bengio et al., 1994). In this context, one of the most basic tasks is counting: in a sequence of symbols, the agent must acquire a scheme to count each symbol as it comes, e.g. (Hochreiter and Schmidhuber, 1997). Learning this scheme is more difficult because it requires to take events that might be variably far away in the past into account.

When using such machine learning methods as is, imitating an agent that has already learned such long-term dependencies is as difficult as learning the same problem without imitation, precisely because of the hidden nature of the useful information. But could there be a way to use imitation to ease learning of such internal states encoding long-term dependencies?

What are people internal states while they solve a given task? A field called "think aloud protocols" (van Someren et al., 1994) deals with this issue. A human subject is given a task. While solving it, instead of thinking quietly, he is instructed to talk aloud. The supervisors record the process, do not intervene and can then exploit the results. This is thus very close to imitation, where the subject is the teacher. The only difference is that the supervisor is not learning to solve the task. This protocol requires only little learning from the subject, and does not disturb the solving process. The idea behind think aloud is to make internal states external. This idea can be used for imitation to learn more complex tasks, involving internal states. Note that, for some reason, when people count objects they often do it aloud.

To force the agents to use a think aloud protocol, they must be constrained so that they do not have internal states. In fact, the are made external. How can this be done? This can be easily seen with RNNs: long-term dependencies are encoded in internal states by recurrent nodes, which allow to hold information through long periods of time, e.g. (Hochreiter and Schmidhuber, 1997). The fact that these states remain hidden for the imitator is because such recurrences are internal, and local.

In the Think Aloud Imitation Learning paradigm, in order to make intermediate computations visible by another agent, agents are constrained not to have internal recurrences. The only possible recurrence is the auditory feedback: the agent hears what it itself says. Contrarily to usual RNN learners, this implies that intermediate computations are done on different time steps, which is also a more general mechanism. Internal states can then be externalized with symbols, like words, that are these intermediate computations. With such a paradigm and using the architecture described in (Orseau, 2005), results show that imitation of long term dependencies can now be done very easily.

An interesting scheme follows from this: it is also possible to make the external recurrence become internal, but still keeping it global. Indeed, with experience, when the agent is about to talk, it can predict what it will hear, and therefore does not need to explicitly say something to activate the corresponding auditory area, and muscles need not move. The agent speaks "aloud", but in its head. Beyond or in extension to imitation, our capacity to explain (teach) to others might be related to our capacity to easily make internal states external.

References

Bengio, Y., Simard, P., and Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2):157–166.

Billard, A. and Mataric, M. (2001). Learning human arm movements by imitation: Evaluation of a biologically-inspired connectionist architecture. *Robotics and Autonomous Systems*, 941:1–16.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8):1735–1780.

Ogata, T., Masago, N., Sugano, S., and Tani, J. (2003). Collaboration development through interactive learning between human and robot. In *Proc. of the 3rd Int. Workshop on Epigenetic Robotics (EpiRob 2003)*.

Orseau, L. (2005). Short-term memories and forcing the re-use of knowledge for generalization. In *ICANN 2005*, LNCS 3697, pages 39–44.

Schaal, S. (1999). Is imitation learning the route to humanoid robots? *Trends in Cognitive Sciences*, 3(6):233–242.

van Someren, M. W., Barnard, Y., and Sandberg, J. (1994). *The Think Aloud Method: A Practical Guide to Modeling Cognitive Processes*. London: Academic Press.