BEFORE WE CAN FIND A MODEL, WE MUST FORGET ABOUT PERFECTION

Dimiter Dobrev
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
d@dobrev.com

In Reinforcement Learning we look for a model of the world. Typically, we aim to find a model which tells everything or almost everything. In other words, we hunt for a perfect model (a total determinate graph) or for an exhaustive model (Markov Decision Process). Finding such a model is an overly ambitious task and indeed a practically unsolvable problem with complex worlds. In order to solve the problem, we will replace perfect and exhaustive models with Event-Driven models.

Only young children assume they can fully understand the world and give answers to all questions. We, the adults, recognise that no one can understand everything and know everything out there. For that reason, we are prepared to relinquish some of our excessive expectations.

The first of these is that the model cannot be improved any further (an assumption known as *Markov property*).

Second, we will abandon the assumption that the model is determinate – which we already did in Markov Decision Process (MDP). Nevertheless, the MDP assumes that whenever nondeterminism is present, each nondeterministic branch occurs with exact probability. This is the third assumption we will say farewell to. We will assume that the probability in question is not precisely determined, but resides within a certain interval *[a, b]*. It may also be that we know nothing about the probability, in which case the interval is *[0, 1]*.

Fourth, we will assume that the model observes only some of the more important events rather than all events. The MDP takes into account all actions of the agent. This makes it an Action-Driven model. We will replace actions with events and will thus move to an Event-Driven model. Event-Driven models are much more stable as they do not change their state at each step, but only at the occurrence of one of the observed events.

Fifth, we will dispense with having separate descriptions of the world and of the agent. Instead, we will describe them as a composite system. The MDP provides only a description of the world which does not include a description of the agent. There is nothing wrong about separating the description of the world from that of the agent. This separation, regretfully, would take a heavy toll on us and prevent our transition to an Event-Driven model. Accordingly, we will discard that assumption, too.

References

[1] Dimiter Dobrev. Event-Driven Models, viXra:1811.0085. November, 2018. http://vixra.org/abs/1811.0085