Research review

Model-based planning from scratch

A team around DeepMind in London approaches planning via the creation of models. While this has been proposed before, the team also proposes to add an "imagination" step for the agent. This means, that the agent can either act directly or beforehand imagine what would happen if he would act in a certain way. Both imagination and actual action then are used to inform future imagination and acting steps. With this method, an agent is able to approach problems with continuous state and action spaces that would otherwise pose an infinite search space.

Their model includes a "manager" module that decides if acting or imagining is the more efficient way to approach the next step. Either way, a controller then decides on the next action. If the agent was to imagine the consequences, an imagination module then plays though the possible outcome and puts the result into memory. If the agent decided that an action should take place, the action will be performed and observations from the real world are then fed back into memory. Both memorization events inform following searches.

In first tests, their proposed model is showing promising results.

Some of their work is based on an earlier idea from Joshua Bengio (2013). Bengio looks into various challenges of deep learning and how to overcome them. He proposes conditional computation as a model to reduce the search space by not following certain planning trees based on previous experience. Dropping search paths with no promising outcome and instead focusing on planning the paths that have proven to be beneficial in the past will allow for easier computation.

To circumvent local minima, always following a beaten (proven) path and also reduce the number of good outcomes that need to be memorized, Bengio proposes to include a factor of randomness into the consideration. While the DeepMind's imagination-based paper does not state any randomness, it would be prudent to assume they are making use of this tactic to some degree.

A similar approach was also proposed in the form of "The Predictron" by a DeepMind team. The main difference is that this model is abstract and does not necessarily interact with a real environment. The approach of training, predicting and learning from those predictions is used to get a better understanding of the task at hand. Accurate value prediction is the main goal.

The proposed architecture is different as well. An "accumulator" module generates an estimated value based on the current state, a model of how an action would translate to the next internal state and a value function. The action models are then updated based on real outcome vs. predicted outcome.

These approaches all help generate agents that are less reliant on specific heuristics but can build models of their perceived world more easily and flexibly. It would allow them to react to huge search spaces more efficiently. Also, with this approach it could be possible to specify more general tasks than what is needed with a narrow heuristic approach.

Sources:

Razvan Pascanu , Yujia Li , Oriol Vinyals, Nicolas Heess, Lars Buesing, Sebastien Racanière, David Reichert, Théophane Weber, Daan Wierstra, Peter Battaglia in "Learning model-based planning from scratch" published July 2017 with the reference *arXiv:1707.06170v1*

Yoshua Bengio in "Deep Learning of representations: Looking Forward" published June 2013 with reference *arXiv:1305.0445v2*

David Silver, Hado van Hasselt, Matteo Hessel, Tom Schaul, Arthur Guez, Tim Harley, Gabriel Dulac-Arnold, David Reichert, Neil Rabinowitz, Andre Barreto, Thomas Degris in "The Predictron: End-To-End Learning and Planning" published July 2017 with the reference *arXiv:1612.08810v3*