Core RL Behavior Suite: `bsuite` report

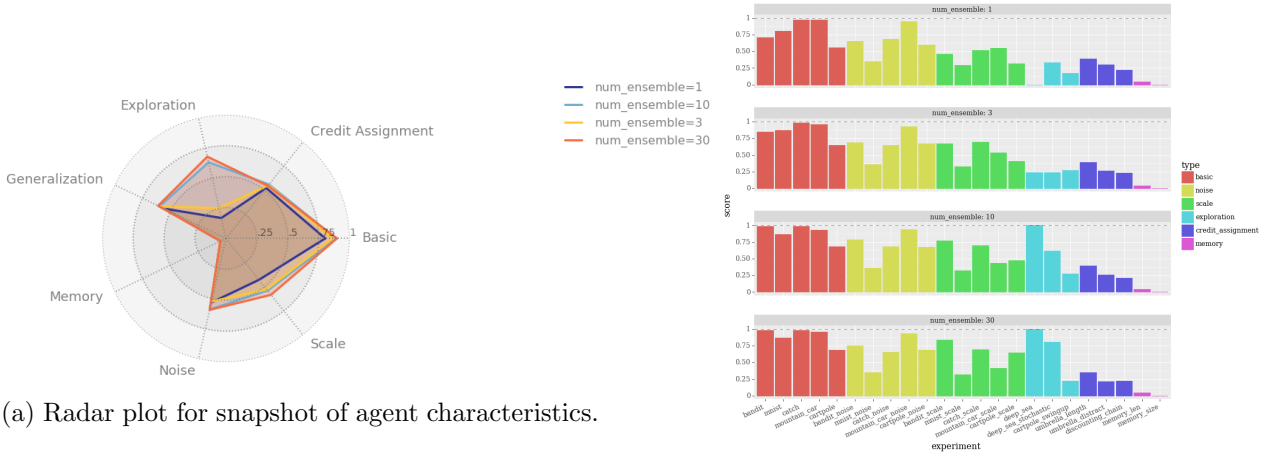
The *Core RL Behavior Suite*, or `bsuite` for short, is a collection of carefully-designed experiments that investigate core capabilities of a reinforcement learning (RL) agent. The aim of the `bsuite` project is to collect clear, informative and scalable problems that capture key issues in the design of efficient and general learning algorithms and study agent behaviour through their performance on these shared benchmarks. This report provides a snapshot of performance on `bsuite2019`, obtained by running the experiments from `github.com/deepmind/bsuite` [2].

1 Agent definition

In this experiment, all agents correspond to different instantiations of a Bootstrapped DQN agent [1], as implemented in `github.com/deepmind/bsuite/baselines/boot_dqn`, and they were trained on `bsuite2019`. We investigate the effect of the number of models ensembled by Bootstrap DQN, sweeping over {1, 3, 10, 30}. We used the default values from `bsuite2019`for the other hyper-parameters of the agent.

2 Summary scores

Each `bsuite` experiment outputs a summary score in [0,1]. We aggregate these scores by according to key experiment type, according to the standard analysis notebook. A detailed analysis of each of these experiments may be found in a notebook hosted on Colaboratory: `ADD-LINK-HERE`.



(a) Radar plot for snapshot of agent characteristics.

(b) Examining score by individual experiment.

Figure 1: Summary output from the `bsuite2019` experiments.

3 Results commentary

As the number of models ensembled by Bootstrap DQN increases the performance on the exploration experiments improves significantly, although there are diminishing returns beyond 10 models. The robustness to the rewards scale also increases mildly with the number of ensembles.

References

[1] Ian Osband, Charles Blundell, Alexander Pritzel, and Benjamin Van Roy. Deep exploration via bootstrapped DQN. In *Advances In Neural Information Processing Systems 29*, pages 4026–4034, 2016.

[2] Ian Osband, Yotam Doron, Matteo Hessel, John Aslanides, Hado Van Hasselt, Eren Sezener, Andre Saraiva, Tor Lattimore, Csaba Szepezvari, Satinder Singh, Benjamin Van Roy, Richard Sutton, and David and Silver. Core RL behaviour suite. 2019.