Enhanced POET

Thursday, 15 October 2020 23:19

- Enhancements:
  - algorithmic
    1. Hand-designed, domain specific metric as a measure of environment novelty → fully generic method for identifying meaningfully novel envs. novel environments are those that provide new information about how the behaviors of agents differ within them.
    2. A more computationally efficient heuristic is formalized; for detemining when agents should goal-switch from one environment to another.
  - external
    3. Novel environment encoding generates more complex and diverse environments
    4. Novel measure for quantifying open-endness

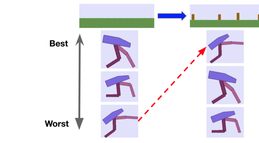
- Algorithmic enhancement:

two important concepts are essential to search through environments

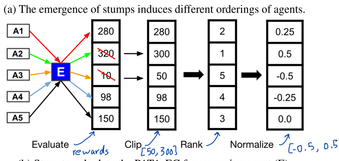
1. The Environmental Encoding (EE): a mapping from a parameter vector to an instance of an environment, creating an environmental search space. EE: $\vec{v} \rightarrow env$
2. The Environment Characterization (EC): describes key attributes of an environment that thereby facilitate calculating distances between environments. EC: $d(c_1, c_2)$

@ - Problem: EE and EC are derived from same set of static, hand-coded features in the original POET
→ the system's output will be bound to exploration only within such prescripted possibilities.

- Solution: a domain-general EC
  - PATA-EC (Performance of All Transferred Agents EC)
    which is grounded by how all agents (in the population and archive) perform in that environment.



(a) The emergence of stumps induces different orderings of agents.



(b) Steps to calculate the PATA-EC for an environment (E).

Figure 2. **PATA-EC, a domain-general distance metric for measuring meaningfully different environments.** (a) An agent that walks with one leg raised is not energy-efficient on flat ground and thus ranks last, but that gaits enables it to step over high stumps and thus ranks highest in a more stumpy environment. (b) The calculation of the PATA-EC for environment E based on the rank of performance of five agents (A1–A5).

- If a newly-generated environment induces a significantly distinct ordering on how agents perform within it (relative to other environments); it likely poses a qualitatively new kind of challenge.
- Performing the normalization allows direct use of the Euclidean distance metric to measure the distance between PATA-ECs.

@ - Problem: the transfer mechanism in original POET has 2 problems:
1. Computationally expensive because it involves an optimization step to compute the fine tuning transfer score.
2. Prone to "false positives" due to stochasticity in RL optimization, and a low bar for replacing more proven incumbents.

- Solution: a more stringent threshold;
(the maximum of the 5 most recent scores of the incumbent)
which both direct and fine-tuning transfer scores of a candidate agent must exceed to qualify as an incoming transfer.
  - This simplification ① smooth out noise from the stochasticity of ES optimization and ② saves computation because the fine-tuning step is performed only if the direct transfer test is passed.

- External enhancement:
  1. More expressive environment encoding:
     - the challenge is how to formalize an encoding that can sustain an environmental space with possibilities beyond the imagination of its designer.
     - Solution: the desired encoding should be highly expressive
       i.e. able to express environmental details with a high degree of granularity and precession to capture ever-more-intricate details.
       Compositional Pattern Producing Networks CPPN [Stanly'07]
       takes as input geometric coordinates, and when queried across such coordinates produces a geometric pattern.

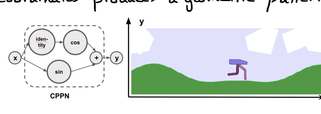


Figure 3. **A sample CPPN (left) and its generated landscape (right).** The CPPN produces y coordinates, given each x coordinate, which are then rendered into a bipedal walker environment for the Bipedal Walker environment in OpenAI Gym (Brockman et al., 2016). An agent, shown in the right figure, is controlled by a different agent neural network to navigate through the generated landscape and is rewarded for quickly moving from left to right.

- As an encoding mechanism, CPPNs offer desirable properties:
  1. They are typically ① initialized with simple topologies (e.g. no hidden layers), and are ② trained with NEAT; a neuroevaluation algorithm that learns both topology and the weights of CPPNs. (more complex landscapes gradually)
  2. CPPNs can evolve arbitrarily complex architectures,
     → in theory, they can express any possible landscape at any conceivable resolution or size.

2. The ANNECS Measure of progress:
How can we tell whether a system continuous generates interesting new things? ANNECS
Accumulated Number of Novel Environments Created and Solved
- The environment created at a particular iteration must:
  1. pass the minimal criterion (neither too hard nor too easy) measured against all agents (in active population and archive)
  2. be eventually solved by the system; the system doesn't receive credit for producing unsolvable environments.

- A POET run takes 60,000 POET iteration
  run in about 12 days
  with 750 CPU cores
  with a population of 40 envs

- For distributed computing Fiber [Zhi'20]