Learning highly diverse robot throwing movements through quality diversity search
Seungsu Kim, Stéphane Doncieux

To cite this version:

HAL Id: hal-02987421
https://hal.archives-ouvertes.fr/hal-02987421
Submitted on 3 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Learning Highly Diverse Robot Throwing Movements through Quality Diversity Search

Seungsu Kim
Sorbonne Universités, UPMC Univ Paris 06, CNRS,
Institute of Intelligent Systems and Robotics (ISIR),
F-75005 Paris, France
seungsu.kim@isir.upmc.fr

Stephane Doncieux
Sorbonne Universités, UPMC Univ Paris 06, CNRS,
Institute of Intelligent Systems and Robotics (ISIR),
F-75005 Paris, France
stephane.doncieux@upmc.fr

ACM Reference format:
DOI: 10.1145/nnnnnn.nnnnnn

1 INTRODUCTION
The behavior of a robot depends on its environment. Evolutionary approaches can be adapted to look for controllers that are robust to changing conditions [4, 10], but it is hard to guarantee that a single behavior will adapt to any new situation. An alternative approach consists in building a repertoire of behaviors in which to look for behaviors adapted to the context [1, 2]. Having a large set of behaviors instead of a single one further allows for extracting information about the robot, the environment and the task, thus opening the way to a self-built understanding of what makes sense for the robot [3] and to the acquisition of higher level representations that can solve new tasks faster and more efficiently [14].

Quality Diversity search algorithms (QD) are well suited to build such repertoires of behaviors [5–8, 11, 13]. Instead of looking for a single solution that is optimal with respect to a user defined function describing a task, QD aims at building a set of high-performing solutions that is as large as possible. Conceptually, QD is based on the novelty search algorithm [5, 12] and Novelty Search with Local Competition (NSLC), its extension that takes into account performance [6]. This is a multi-objective approach in which a local performance score balances a novelty objective. It generates many different good quality solutions, introducing the concept of quality and diversity. MAP-Elites is another QD algorithm [7, 13]. The multi-dimensional behavior space is discretized into a set of bins (a behavioral map). The algorithm aims at filling the bins with high quality behaviors. These QD algorithms are all based on the same principles: they look for novel solutions, in terms of the behavior they exhibit, and take into account the performance only at a local scale: an individual succeeds if it is novel or else if it is close to other behaviors, but more fit.

The process of the QD algorithm used here follows the principles of NSLC [6]. After an initial random generation, the parents are selected proportionally to their novelty and local quality score. The off-springs (individuals) are populated from the selected parents. Each individual is evaluated by observing the corresponding behavior in simulation and computing the corresponding quality and diversity scores. If the diversity score of an individual is larger than a threshold, the individual is stored in the archive. Otherwise, its quality is compared with one of the most similar individuals in the archive. If the quality score is larger than the one in the archive, the individual in the archive is replaced with the new one.

In this work, we present an application of QD algorithms to robotics. The proposed experiment consists in learning trajectories that make the robot launch a ball in a basket. The diverse set of solutions found is aimed to be used in a selection process that could discard solutions that do not cross the reality gap [9] or adapt to a new situation. This work focuses on the generation of the repertoire of trajectories. Exploitation of the repertoire is left to future work.

2 EXPERIMENT
The goal of the presented application is to find diverse and feasible robot arm trajectories for throwing an object into a given basket. The experiment relies on the Baxter robot. The robot has two arms with 7 degrees of freedom and a gripper is attached at the end of
each arm. In order to plan the robot throwing trajectory, we use 3rd order polynomials, because of their computational simplicity and smooth trajectory planning capacity. The parameters of the planner are the joint positions of the arm (7-DOF) and their velocities at the releasing time as well as the duration of the motion. These 15 parameters constitute the genotype of a solution.

The behaviour associated to a trajectory is evaluated in the DART physical engine. The evaluation of a genotype starts with a cube placed in the robot gripper. The robot moves its right arm from a pre-defined starting joint configuration to the target joint positions with the target joint velocities contained in the genotype. When it reaches the target joint configuration, the robot releases the cube. The simulation runs until the cube falls on the ground or reaches the basket. The behaviour descriptor $bd(x)$ associated to a genotype $x$ is the arm trajectory of the robot in the Cartesian space, described by five equally distributed intermediate points of the trajectory ($bd(x) \in \mathbb{R}^{15}$). The local quality of a behavior is the minimum distance between the thrown object trajectory and the basket position. Hence, the QD algorithm finds diverse robot arm trajectories that launch the cube as close as possible to the target.

Overall 5000 generations of QD algorithms are performed. In each generation, 200 parent individuals are selected from the archive with a probability proportional to their Pareto rank. Crossover (crossover rate: 0.15) and mutation (mutation rate: 0.15) operators are applied, and the same number of independent offsprings are generated and simulated.

3 RESULT AND DISCUSSION

After 5000 generations, the archive contains 3430 throwing behaviors. Among them, 98 are able to throw the cube into the basket. Half of the obtained throwing behaviors are shown in Figure 1. Two of them are highlighted with some snapshots from the simulation in Figure 2. As can be seen on the figure, the QD algorithm has found diverse throwing behaviors that can throw the cube into the basket.

The next step is to use these trajectories to design primitive behaviours for a real Baxter robot. When the robot faces a new situation (e.g. an obstacle is presented in the environment), through trial and error, its controller can select an appropriate behavior that avoids it [1], thus simultaneously finding out how to face the new situation and crossing the reality gap [9].

ACKNOWLEDGEMENT

This work is supported by the DREAM project 1 through the European Unions Horizon 2020 research and innovation program under grant agreement No 640891.

REFERENCES


1http://www.robotsatthatdream.eu