Combining learning in simulation and in reality to build robust skills on an arm robot

Mots clés : Array

Directeur de thèse : Stéphane Doncieux
Co-encadrant(s) :
  ● Nicolas PERRIN
  ● Alexandre CONINX
Unité de recherche : Institut des Systèmes Intelligents et de Robotique
Ecole doctorale : École Doctorale Informatique, Télécommunications, Électronique de Paris
Domaine scientifique principal : Sciences pour l’ingénieur

Résumé du projet de recherche (Langue 1)

Learning usually consists in a trial and error approach, and most of the state-of-the-art algorithms are "data hungry". Learning and improving skills with a robot can typically require thousands of hours of interactions between the robot and its environment (a seminal result by Levine et al. involved more than 3000 robot-hours of practice [Levine et al. 2016]).

To avoid such impractical and long training phases, it would be convenient to rely on simulation, where virtual experiments are much faster and easier to perform and parallelize.

However, the well-known "reality gap" makes it difficult to transfer virtual skills to the real robot. We propose to tackle this problem by learning in simulation a large variety of distinct solutions for every task, thus aiming not only at solving them but also understanding the potential variability of solutions. We propose to rely on this variability to then guide a fast and robust transfer to the real robot.

Given a task (defined by a reward to maximize, and constraints), we plan, in a preliminary phase, to use a divergent search process to build many different behaviors that are adapted to many different situations. This will require reasoning on potential constraints that are not known yet, on how behaviors can be adapted if the situation changes, and on how the behaviors can be stored to make selection as easy as possible. To tackle these issues, it will be necessary to develop new specific divergent search algorithms. To this end, we propose to rely on Quality Diversity algorithms [Pugh et al. 2016, Cully and Demiris 2017]. On the basis of a given behavior descriptor that defines a space of behaviors to focus on, Quality Diversity algorithms build a repertoire of skills that cover, as densely as possible, the behaviors that the robot can exhibit. This offers the possibility to later extract from the repertoire the behavior that will fit best to the situation [Cully et al. 2015]. QD algorithm proceed by generating skills, testing them in a particular setup and extracting from this observation a behavior descriptor. The behavior descriptor is then used by the algorithm to determine what to do with this new skill: either adding it if no similar behaviors have been generated so far or comparing it to similar solutions, if any, and keeping it if it performs better. The skills generated are thus adapted to a particular context and may not generalize to a different situation. In the case of an arm robot for instance, a grasping skill can be discovered that works only for a particular object position and orientation.

The goal of this thesis is to propose an approach to extract regularities from the set of examples generated by the QD search in order to build parameterized motion primitives that will work on the real robot and adapt to different contexts, for instance different object positions in the example of a grasping skill. The thesis aims at exploiting the flexibility of simulation to reduce the number of samples required on the real robot to build such robust skills. To reach that goal, the main questions will be to (1) generate diverse sets of skills in simulation, (2) extract regularities to build a parameterized skill that is robust to the spatial and temporal perturbations resulting from a real world experiment.

Informations complémentaires (Langue 1)

Learning usually consists in a trial and error approach, and most of the state-of-the-art algorithms are "data hungry". Learning and improving skills with a robot can typically require thousands of hours of interactions between the robot and its environment (a seminal result by Levine et al. involved more than 3000 robot-hours of practice [Levine et al. 2016]).

To avoid such impractical and long training phases, it would be convenient to rely on simulation, where virtual experiments are much faster and easier to perform and parallelize.

However, the well-known "reality gap" makes it difficult to transfer virtual skills to the real robot. We propose to tackle this problem by learning in simulation a large variety of distinct solutions for every task, thus aiming not only at solving them but also understanding the potential variability of solutions. We propose to rely on this variability to then guide a fast and robust transfer to the real robot.

Given a task (defined by a reward to maximize, and constraints), we plan, in a preliminary phase, to use a divergent search process to build many different behaviors that are adapted to many different situations. This will require reasoning on potential constraints that are not known yet, on how behaviors can be adapted if the situation changes, and on how the behaviors can be stored to make selection as easy as possible. To tackle these issues, it will be necessary to develop new specific divergent search algorithms. To this end, we propose to rely on Quality Diversity algorithms [Pugh et al. 2016, Cully and Demiris 2017]. On the basis of a given behavior descriptor that defines a space of behaviors to focus on, Quality Diversity algorithms build a repertoire of skills that cover, as densely as possible, the behaviors that the robot can exhibit. This offers the possibility to later extract from the repertoire the behavior that will fit best to the situation [Cully et al. 2015]. QD algorithm proceed by generating skills, testing them in a particular setup and extracting from this observation a behavior descriptor. The behavior descriptor is then used by the algorithm to determine what to do with this new skill: either adding it if no similar behaviors have been generated so far or comparing it to similar solutions, if any, and keeping it if it performs better. The skills generated are thus adapted to a particular context and may not generalize to a different situation. In the case of an arm robot for instance, a grasping skill can be discovered that works only for a particular object position and orientation.

The goal of this thesis is to propose an approach to extract regularities from the set of examples generated by the QD search in order to build parameterized motion primitives that will work on the real robot and adapt to different contexts, for instance different object positions in the example of a grasping skill. The thesis aims at exploiting the flexibility of simulation to reduce the number of samples required on the real robot to build such robust skills. To reach that goal, the main questions will be to (1) generate diverse sets of skills in simulation, (2) extract regularities to build a parameterized skill that is robust to the spatial and temporal perturbations resulting from a real world experiment.
References:


