Safety-Aware Robot Damage Recovery Using Constrained Bayesian Optimization and Simulated Priors

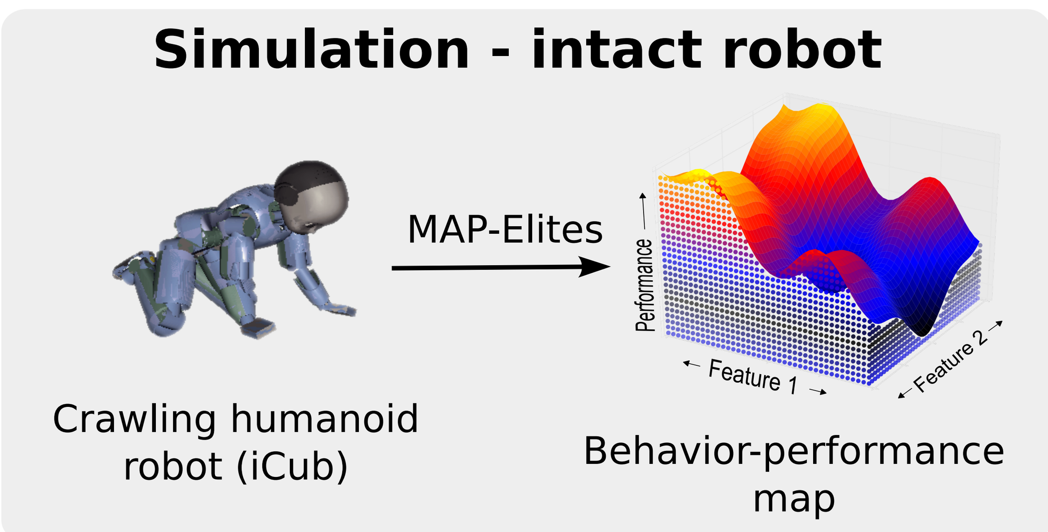
■ **Bayesian Optimization: Black-box Optimization and Beyond**
NIPS Workshop on Bayesian Optimization, 10 December 2016, Barcelona - Spain

Vaios Papaspyros[1,2,3,4] • Konstantinos Chatzilygeroudis[1,2,3] • Vassilis Vassiliades[1,2,3] • Jean-Baptiste Mouret[1,2,3]
papaspyros@ceid.upatras.gr / konstantinos.chatzilygeroudis@inria.fr / vassilis.vassiliades@inria.fr / jean-baptiste.mouret@inria.fr

- A major obstacle to the widespread adoption of robots is their fragility.
- The recently introduced **Intelligent Trial-and-Error algorithm (IT&E)**, by *Cully et al.* [1], allows autonomous robots to adapt to damage in a matter of a few trials. However, trial-and-error approaches **lack any safety constraints/limits** and are likely to damage the robot even more by trying behaviors that are too extreme for the mechanical design.
- We extend the IT&E algorithm to a **safety-aware Intelligent Trial-and-Error algorithm (sIT&E) by: (1)** *introducing safety constraints in the Bayesian optimization procedure* and **(2)** *automatically computing priors over the safety of controller parameters*.

Creating the behavior-performance map with MAP-Elites

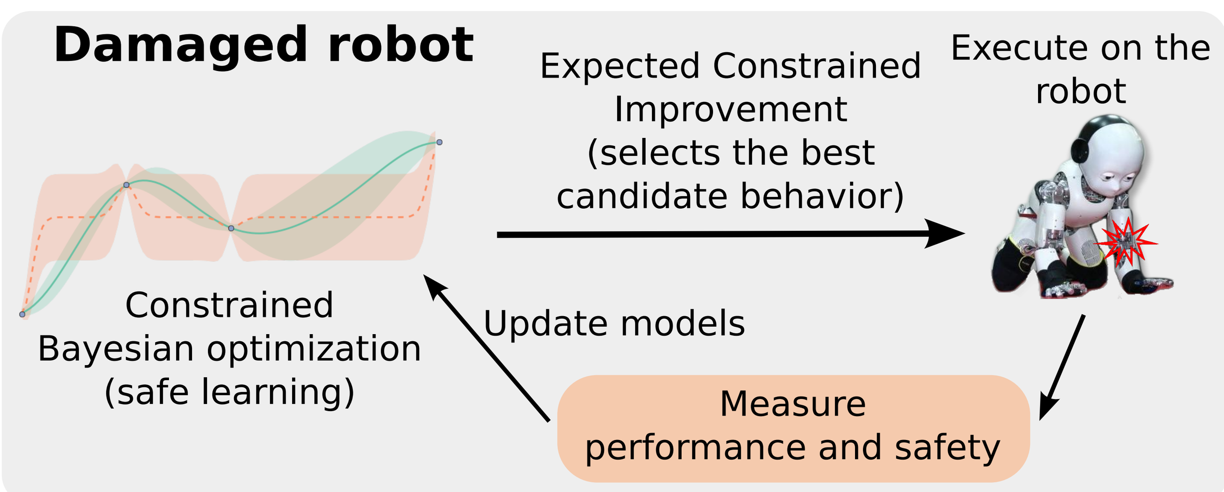
- **Goal:** (1) reduce the dimensionality of the search space (from policy parameters to behavior descriptors), (2) allow to compute kernels in the behavior space, (3) pre-compute priors on performance.



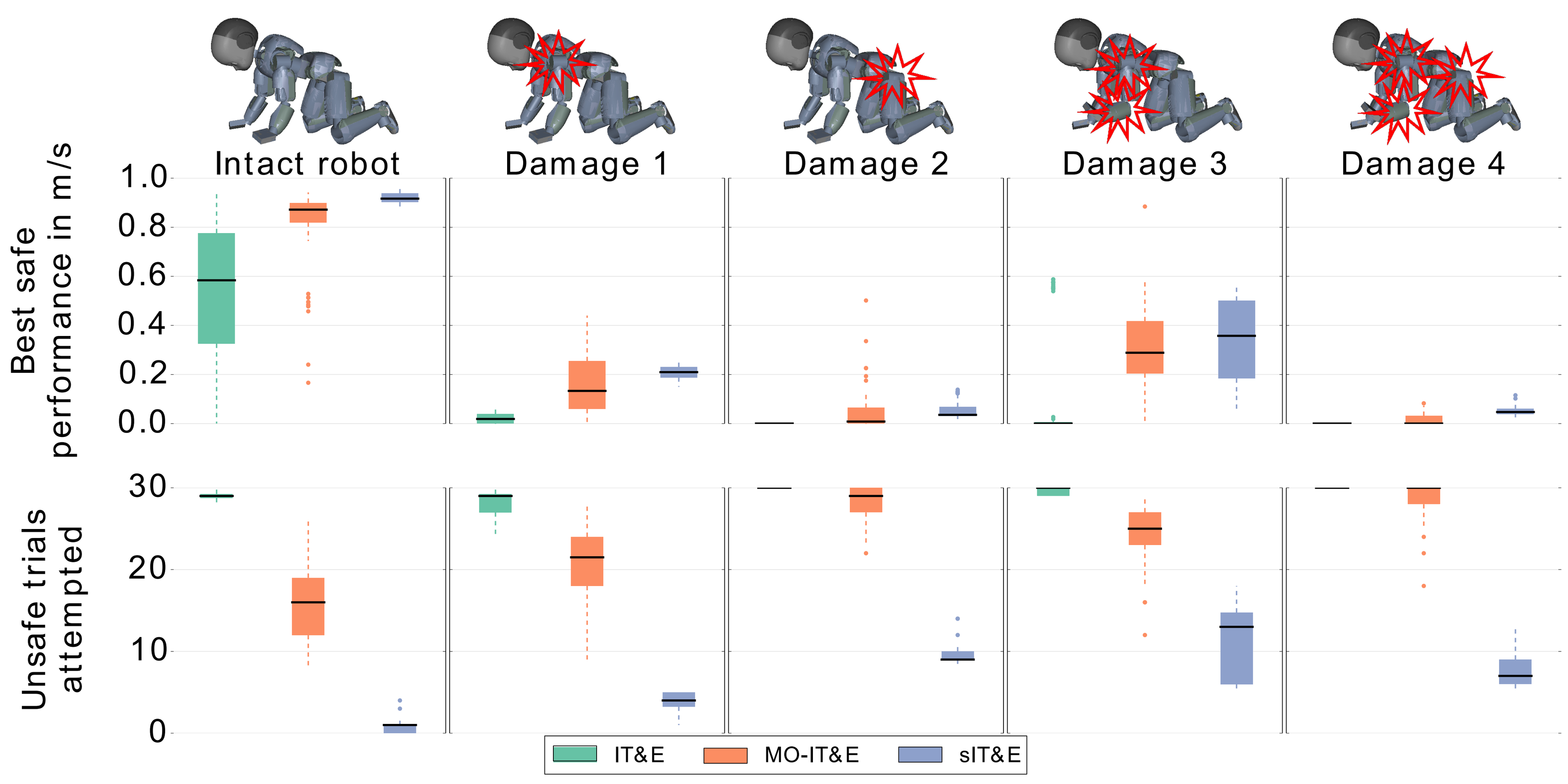
- We create this "behavior-performance" map with additional diversity and priors over user-defined safety constraints.
- In our simulated iCub experiments, the behavior space is 5D: 4 dimensions for crawling diversity and 1 dimension for the sum of contact point forces.

Adaptation Step with constrained Bayesian optimization

- **Goal:** Find high performing and safe compensatory behaviors using the behavior-performance produced in simulation as prior knowledge.



- We use Gaussian Processes to learn the difference between the performance of the intact and the damaged robot.
- We model each safety constraint as a separate Gaussian Process, which we update as we collect data from the damaged robot.
- We use the Expected Constrained Improvement (ECI) acquisition function: $ECI(\hat{x}) = EI(\hat{x}) \prod_{i=1}^{n} p(c_i(\hat{x}) \geq 0)$
- After damage, the robot (1) selects the most promising behavior in terms of performance but within the safe region defined by the constraints (by optimizing the ECI), (2) executes it, and (3) updates the Gaussian processes.



Crawling Humanoid Experiments (in simulation)

- We select 4 different damage scenarios; (1) locked shoulder joint, (2) locked hip joint, (3) locked shoulder joint & angled elbow, and (4) combination of 2 & 3.
- We compare against the vanilla IT&E algorithm and a Multi-objective IT&E (MO-IT&E) variant based on the Expected Hypervolume Improvement (EHVI).
- sIT&E outperforms both algorithms in terms of safety (we measure safety by the unsafe trials conducted through the optimization procedure), while it also dominates in terms of safe crawling speed (m/s). Specifically, sIT&E averages less than 10 unsafe trials for damages 1, 2, 4 and just over 10 for damage 3 compared to 29 or 30 for IT&E and more than 22 for MO-IT&E.

Future/Current Work

- Non-episodic learning — **"robots should learn while doing"**, not "learn and then do".
  - Reset-free learning algorithms [2] are essential if we want robots to operate in real world scenarios.
- Illumination algorithms that scale better can allow us to produce priors on big policy parameter spaces [3].
- Combining closed-loop policies with learning can lead to safer and more robust results.

Take home messages

- Unconstrained learning can break robots: we need constraints.
- Pre-computing a behavior-performance map can accelerate search in large search spaces.
- Computing kernels in a meaningful space (e.g. the behavior space) can allow BO to work in high-dimensional search spaces.

1 Inria, Villers-lès-Nancy, F-54600, France
2 CNRS, Loria, UMR 7503, Vandœuvre-lès-Nancy, F-54500, France
3 Université de Lorraine, Loria, UMR 7503, Vandœuvre-lès-Nancy, F-54500, France
4 University of Patras, 26500 Rion, Patras, Greece

[1] A. Cully, J. Clune, D. Tarapore, and J.-B. Mouret, "Robots that can adapt like animals", Nature, vol. 521, no. 7553, pp. 503–507, 2015.
[2] K. Chatzilygeroudis, V. Vassiliades and J.B. Mouret, 2016. "Reset-free Trial-and-Error Learning for Data-Efficient Robot Damage Recovery". arXiv preprint arXiv:1610.04213.
[3] V. Vassiliades, K. Chatzilygeroudis, and J.B. Mouret, 2016. "Scaling Up MAP-Elites Using Centroidal Voronoi Tessellations". arXiv preprint arXiv:1610.05729.

Funding: This work received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement number 637972, project "ResiBots"): www.resibots.eu